DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Deterministic quadrature formulas for SDEs based on simplified weak Ito-Taylor steps

T. Müller-Gronbach, L. Yaroslavtseva

Preprint 167

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Deterministic quadrature formulas for SDEs based on simplified weak Ito-Taylor steps

T. Müller-Gronbach, L. Yaroslavtseva

Preprint 167

DETERMINISTIC QUADRATURE FORMULAS FOR SDES BASED ON SIMPLIFIED WEAK ITÔ-TAYLOR STEPS

THOMAS MÜLLER-GRONBACH AND LARISA YAROSLAVTSEVA

ABSTRACT. We study the problem of approximating the expected value $\mathbb{E}f(X(1))$ of a function f of the solution X of a d-dimensional system of stochastic differential equations (SDE) at time point 1 based on finitely many evaluations of the coefficients of the SDE, the integrand f and their derivatives. We present a deterministic algorithm, which produces a quadrature rule by iteratively applying a Markov transition based on the distribution of a simplified weak Ito-Taylor step together with strategies to reduce the diameter and the size of the support of a discrete measure. We essentially assume that the coefficients of the SDE are s-times continuously differentiable and the diffusion coefficient satisfies a uniform non-degeneracy condition and that the integrand f is r-times continuously differentiable. In the case $r \leq (\lfloor s/2 \rfloor - 1) \cdot 2d/(d+2)$ we almost achieve an error of order $\min(r, s)/d$ in terms of the computational cost, which is optimal in a worst case sense.

1. INTRODUCTION

Consider a d-dimensional system of stochastic differential equations

$$
\begin{aligned}
dX(t) &= a(X(t))\,dt + b(X(t))\,dW(t), \quad t \in [0,1], \\
X(0) &= x_0
\end{aligned}
\tag{1}
$$

with initial value $x_0 \in \mathbb{R}^d$, drift coefficient $a\colon \mathbb{R}^d \to \mathbb{R}^d$, diffusion coefficient $b\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and an m-dimensional driving Brownian motion W as well as a function

$$
f\colon \mathbb{R}^d \to \mathbb{R}.
$$

Our computational task is to approximate the integral

$$
I(x_0, a, b, f) = \mathbb{E}f(X(1)) \tag{2}
$$

by means of a deterministic method that uses x_0 and finitely many evaluations of a, b and f and their derivatives at points in $\mathbb{R}^d$.

In this paper we restrict to the case of bounded coefficients a and b that have bounded partial derivatives up to order $s \in \mathbb{N}$ with $s \geq 4$, and we impose a uniform non-degeneracy condition on the diffusion coefficient b. In particular, we assume $m \geq d$. The integrand f is supposed to have polynomially bounded partial derivatives up to order $r \in \mathbb{N}$.

For every $n \in \mathbb{N}$ we present a method $\widehat{I}_n$ that performs a deterministic walk of n steps in the space $\mathcal{M}_1(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$ with finite support and finally yields a probability measure $Q_n(x_0, a, b) \in \mathcal{M}_1(\mathbb{R}^d)$, which is close to the distribution $\mathbb{P}_{X(1)}$ of the solution of (1) at point 1. The integral $I(x_0, a, b, f)$ is then approximated by the finite weighted sum

$$\widehat{I}_n(x_0, a, b, f) = \int f \, dQ_n(x_0, a, b).$$

The deterministic walk in $\mathcal{M}_1(\mathbb{R}^d)$ starts with the one-point measure δ_{x_0}. In every step, first a Markov transition of the form

$$\mathcal{M}_1(\mathbb{R}^d) \ni \mu \mapsto T(\mu) = \sum_{y \in \mathrm{supp}(\mu)} \mu(\{y\}) \cdot \mathbb{P}_{Y^y} \in \mathcal{M}_1(\mathbb{R}^d) \tag{3}$$

is carried out, where Y^y denotes a simplified Itô-Taylor step with initial value y, and afterwards the diameter and the cardinality of the support of $T(\mu)$ are reduced in order to avoid an explosion of the computational cost. The weak order of the simplified Itô-Taylor steps is adjusted to the smoothness s of the coefficients a and b and the single step-sizes are chosen in a non-uniform way.

Our method is closely related to the Wiener Cubature approach, which was initiated in [10] and further developed in particular in [11, 12, 13, 14]. We use the same type of non-uniform discretization of time and we employ the localized version of a sequential support point elimination procedure of Davis [2] introduced in [12] to reduce the cardinality of the support of a discrete measure. However, the transition step (3) is conceptually much simpler than the procedures used in [10, 11] and in [12, 13, 14]. In particular, we do not need to solve a collection of ODEs along paths of bounded variation as in [12, 13, 14]. Moreover, by additionally performing a support diameter reduction step we obtain a further reduction of the computational cost while preserving the error estimate.

To give a flavour of our results we consider the case that

$$r \le (\lfloor s/2 \rfloor - 1) \cdot 2d/(d+2) \tag{4}$$

and all partial derivatives of f up to order r are bounded. In this case our method achieves an error

$$|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)| \le c \cdot \max_{1 \le |\alpha|_1 \le r} \|f^{(\alpha)}\|_\infty \cdot n^{-\frac{\eta r}{2}}$$

with computational cost

$$\mathrm{cost}(\widehat{I}_n, (x_0, a, b, f)) \le c \cdot n^{\eta d(\frac{1}{2}+\varepsilon)},$$

see Theorem 1. Here, $\eta, \varepsilon > 0$ are parameters of the method $\widehat{I}_n$ and $c > 0$ is a positive constant that neither depends on n nor on f. Furthermore, $\mathrm{cost}(\widehat{I}_n, (x_0, a, b, f))$ is defined by the sum of the number of all evaluations of a, b and f as well as their partial derivatives and the number of all arithmetical operations and evaluations of elementary functions that are carried out by the algorithm $\widehat{I}_n$ to compute the approximation $\widehat{I}_n(x_0, a, b, f)$. In particular,

$$|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)| \leq c \cdot \max_{1 \leq |\alpha|_1 \leq r} \|f^{(\alpha)}\|_\infty \cdot (\mathrm{cost}(\widehat{I}_n, (x_0, a, b, f)))^{-\frac{r}{d(1+2\varepsilon)}},$$

and therefore our method almost achieves the order of convergence r/d in terms of its computational cost since ε can be chosen arbitrarily small.

If the support diameter reduction step is skipped in the definition of $\widehat{I}_n$ then we can only prove the order of convergence

$$\frac{r}{d} \cdot \frac{2(\lfloor s/2 \rfloor - 1)}{2(\lfloor s/2 \rfloor - 1) + r} < \frac{r}{d}$$

for the resulting method in terms of its computational cost, see Remark 2. We conjecture that this estimate is essentially sharp and and we expect a suboptimal order of convergence for the Wiener Cubature approach as well. See also the discussion in [17, Remark 4.1] with respect to the latter method in the case $r = 1$.

In [3, 16, 17] the complexity of approximating the distribution $\mathbb{P}_{X(1)}$ by a probability measure with finite support is studied in a worst case setting with respect to classes of equations (x_0, a, b) and classes of integrands f that are essentially specified by smoothness assumptions on a, b and f. It follows from the results in [17] that up to an arbitrary small power, $\min(r, s)/d$ is the best possible order in terms of the computational cost that can be achieved by any deterministic algorithm for the computational problem (2) in a suitable worst case sense, see Theorem 2. Consequently, if the condition (4) is satisfied then the algorithms $\widehat{I}_n$ perform almost asymptotically optimal, see Theorem 3. If the condition (4) does not hold then our method almost achieves the order $2(\lfloor s/2 \rfloor - 1)/(d + 2)$, which is suboptimal.

We briefly outline the content of the paper. Basic notation is presented in Section 2. In Section 3 we introduce the particular type of a simplified Itô-Taylor step that is used in Section 4 for the construction of the algorithm $\widehat{I}_n$. The computational cost and the error of $\widehat{I}_n$ are analyzed in Section 5, while Section 6 is devoted to lower error bounds for arbitrary deterministic algorithms and optimality properties of $\widehat{I}_n$. In Section 7 we present numerical experiments to illustrate our results. Proofs are postponed to Section 8. An appendix contains the description of the sequential support point elimination procedure of Davis [2] and provides results on smoothness properties of the semigroup of transition operators associated with an equation (x_0, a, b), which are used for the proofs in Section 8.

2. Notation

Throughout the sequel we use $c, c_1, \dots$ and $c(\cdot), c_1(\cdot), \dots$ to denote unspecified positive constants, which may only depend on the parameters eventually specified in brackets or explicitly mentioned in the context.

The minimal eigenvalue of a symmetric non-negative square matrix A is denoted by $\lambda_{\min}(A)$. For a vector $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and a multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ we put $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ as well as $\alpha! = \alpha_1! \cdots \alpha_d!$.

For $p \in [1, \infty]$ we use $|\cdot|_p$ to denote the p-norm on a finite-dimensional vector space and we put $|\cdot| = |\cdot|_\infty$. Furthermore, $\|h\| = \sup_{x \in \mathbb{R}^d} |h(x)|$ denotes the supremum norm of a function $h\colon \mathbb{R}^d \to \mathbb{R}^{l \times k}$.

We use $\mathcal{P}_k(\mathbb{R}^d)$ to denote the set of polynomials on $\mathbb{R}^d$ of total degree at most k. By $C(\mathbb{R}^d)$ and $C^k(\mathbb{R}^d)$ we denote the space of all continuous functions and the space of all k times continuously differentiable functions $h\colon \mathbb{R}^d \to \mathbb{R}$, respectively. For $K > 0$ we put

$$C_K^k = \Big\{ h \in C^k(\mathbb{R}^d)\colon \max_{0 \le |\alpha|_1 \le k} \|h^{(\alpha)}\| \le K \Big\}.$$

Furthermore, for $\beta \ge 0$ we let

$$F^{k,\beta} = \{ h \in C^k(\mathbb{R}^d)\colon \|h\|_{k,\beta} < \infty \},$$

where

$$\|h\|_{k,\beta} = \max_{1 \le |\alpha|_1 \le k} \sup_{x \in \mathbb{R}^d} \frac{|h^{(\alpha)}(x)|}{1 + |x|^\beta},$$

and we put

$$F^k = \bigcup_{\beta \ge 0} F^{k,\beta}.$$

We use $\mathcal{M}(\mathbb{R}^d)$ to denote the set of all Borel measures with finite support on $\mathbb{R}^d$. Moreover, $\mathcal{M}_1(\mathbb{R}^d) \subset \mathcal{M}(\mathbb{R}^d)$ denotes the subset of probability measures in $\mathcal{M}(\mathbb{R}^d)$. For a Borel probability measure μ on $\mathbb{R}^d$ and a μ-integrable function $h\colon \mathbb{R}^d \to \mathbb{R}$ we use

$$\mathbb{E}_\mu h = \int h \, d\mu$$

to denote the integral of h with respect to μ, and we put

$$m_p(\mu) = \int |x|_p^p \, \mu(dx)$$

for $p \ge 1$. Finally, we use $\stackrel{d}{=}$ to denote equality in distribution of two random variables.

3. Simplified weak order γ Itô-Taylor steps

Let $W_0(u) = u$ for $u \geq 0$. For $\ell \in \mathbb{N}$, $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{N}_0^\ell$ and $\tau \in [0, 1]$ we use

$$J_\alpha(\tau) = \int_0^\tau \ldots \int_0^{u_2} 1 \, dW_{\alpha_1}(u_1) \ldots dW_{\alpha_\ell}(u_\ell)$$

to denote the corresponding iterated Itô-integral and we put

$$\deg(\alpha) = \ell + |\{i \colon \alpha_i = 0\}|.$$

For further purposes we note that

$$\big(J_{\alpha^{(1)}}(\tau), \ldots, J_{\alpha^{(n)}}(\tau)\big) \stackrel{d}{=} \big(\tau^{\deg(\alpha^{(1)})/2} \cdot J_{\alpha^{(1)}}(1), \ldots, \tau^{\deg(\alpha^{(n)})/2} \cdot J_{\alpha^{(n)}}(1)\big) \tag{5}$$

for any multi-indices $\alpha^{(1)}, \ldots, \alpha^{(n)}$ and $n \in \mathbb{N}$.

For $\gamma \in \mathbb{N}$ we put

$$\Gamma_\gamma = \bigcup_{\ell=1}^{\gamma} \{0, \ldots, m\}^\ell.$$

Consider a family $\xi = (\xi_\alpha)_{\alpha \in \Gamma_\gamma}$ of real-valued random variables ξ_α on a common probability space. We call ξ a *simplification of order* γ if $\mathbb{P}_{\xi_\alpha} \in \mathcal{M}_1(\mathbb{R})$ for all $\alpha \in \Gamma_\gamma$ and

$$\mathbb{E} \prod_{\ell=1}^{L} \xi_{\alpha^{(\ell)}} = \mathbb{E} \prod_{\ell=1}^{L} J_{\alpha^{(\ell)}}(1) \tag{6}$$

for all $\alpha^{(1)}, \ldots, \alpha^{(L)} \in \Gamma_\gamma$ with $\deg(\alpha^{(1)}) + \ldots + \deg(\alpha^{(L)}) \leq 2\gamma+1$ and $L = 1, \ldots, 2\gamma+1$. We add that in [10] ξ is called a $(2\gamma+1)$-moment similar family if (6) holds with iterated Stratonovich-integrals in place of the iterated Itô-integrals $J_{\alpha^{(\ell)}}(1)$.

By Tchakaloff's theorem, for every $\gamma \in \mathbb{N}$, there exists a simplification ξ of order γ with

$$|\operatorname{supp}(\mathbb{P}_{\xi_\alpha})| \leq \Big| \Big\{ (\alpha^{(1)}, \ldots, \alpha^{(L)}) \in \Gamma_\gamma^L \colon \sum_{\ell=1}^{L} \deg(\alpha^{(\ell)}) \leq 2\gamma + 1, \, L = 1, \ldots, 2\gamma + 1 \Big\} \Big|$$

for all $\alpha \in \Gamma_\gamma$, see [1, Corollary 2]. See Example 1 for examples of simplifications of order $\gamma = 1, 2$.

Consider a drift coefficient $a \in (C^{2(\gamma-1)}(\mathbb{R}^d))^d$ as well as a diffusion coefficient $b \in (C^{2(\gamma-1)}(\mathbb{R}^d))^{d \times m}$ with associated differential operators

$$L_0 \colon C^2(\mathbb{R}^d) \to C(\mathbb{R}^d), \; h \mapsto \sum_{i=1}^{d} a_i \frac{\partial h}{\partial y_i} + \frac{1}{2} \sum_{i,k=1}^{d} \sum_{j=1}^{m} b_{i,j} b_{k,j} \frac{\partial^2 h}{\partial y_i \partial y_k}$$

and

$$L_j \colon C^1(\mathbb{R}^d) \to C(\mathbb{R}^d), \; h \mapsto \sum_{i=1}^{d} b_{i,j} \frac{\partial h}{\partial y_i}$$

for $j = 1, \dots, m$. Put $\pi_i(y) = y_i$ for $y \in \mathbb{R}^d$ and $i = 1, \dots, d$. The Itô-coefficient functions associated with a and b are then given by

$$\psi_\alpha = (L_{\alpha_1} \circ \dots \circ L_{\alpha_\ell} \pi_i)_{i=1,\dots,d} \colon \mathbb{R}^d \to \mathbb{R}^d$$

for multi-indices $\alpha \in \Gamma_\gamma$.

Let $y \in \mathbb{R}^d$. Then

$$\widehat{X}^{y,a,b}(\tau) = y + \sum_{\alpha \in \Gamma_\gamma} \psi_\alpha(y)\, J_\alpha(\tau) \tag{7}$$

is a single step of size τ with initial value y of the weak order γ Itô-Taylor scheme associated with a and b, see e.g. [9, Section 14.2]. Consider a simplification ξ of order γ and put

$$\xi_\alpha(\tau) = \tau^{\deg(\alpha)/2}\, \xi_\alpha$$

for $\alpha \in \Gamma_\gamma$. We define a simplified weak order γ Itô-Taylor step $Y^{y,a,b}(\tau)$ of size τ with initial value y by replacing the iterated Itô-integrals $J_\alpha(\tau)$ in (7) by the discrete variables $\xi_\alpha(\tau)$, i.e.,

$$Y^{y,a,b}(\tau) = y + \sum_{\alpha \in \Gamma_\gamma} \psi_\alpha(y)\, \xi_\alpha(\tau). \tag{8}$$

Example 1. A simplification ξ of order 1 is, e.g., given by $\xi_{(0)} = 1$ and independent random variables $\xi_{(1)}, \dots, \xi_{(m)}$ with

$$\mathbb{P}_{\xi_{(j)}} = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$$

for $j = 1, \dots, m$.

In the case $\gamma = 2$ one can take independent random variables

$$(N_j)_{1 \le j \le m},\ (V_{j_1,j_2})_{1 \le j_1 < j_2 \le m}$$

with distributions given by

$$\mathbb{P}_{N_j} = \frac{1}{6}\delta_{-\sqrt{3}} + \frac{2}{3}\delta_0 + \frac{1}{6}\delta_{\sqrt{3}}, \quad \mathbb{P}_{V_{j_1,j_2}} = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1,$$

and define a simplification ξ of order 2 by $\xi_{(0)} = 1$, $\xi_{(0,0)} = 1/2$ as well as

$$\xi_{(j)} = \xi_{(j,0)} = \xi_{(0,j)} = N_j$$

and

$$\xi_{(j_1,j_2)} = \begin{cases} (N_{j_1}N_{j_2} + V_{j_1,j_2})/2, & \text{if } j_1 < j_2, \\ (N_{j_1}N_{j_2} - V_{j_2,j_1})/2, & \text{if } j_1 > j_2, \\ (N_{j_1}^2 - 1)/2, & \text{if } j_1 = j_2 \end{cases}$$

for $j, j_1, j_2 \in \{1, \dots, m\}$. This family has been introduced in [19] in the case $d = m = 1$ and was extended in [9, Sec. 5.12] to arbitrary dimensions $d, m \in \mathbb{N}$.

4. The Algorithm

Let $r \in \mathbb{N}$ and $\beta \geq 0$ as well as $\gamma \in \mathbb{N}$ and $K \geq 1 \geq \lambda > 0$. We consider integrands f from the class $F^{r,\beta}$ and equations (x_0, a, b) from the class

$$\mathcal{C}^{2\gamma+2}_{K,\lambda} = \Big\{ (x_0, a, b) \in [-K, K]^d \times (C^{2\gamma+2}_K)^d \times (C^{2\gamma+2}_K)^{d\times m} \colon \inf_{x\in\mathbb{R}^d} \lambda_{\min}(bb^T(x)) \geq \lambda \Big\}.$$

Let $n \in \mathbb{N}$. We first introduce a method

$$Q_n \colon \mathcal{C}^{2\gamma+2}_{K,\lambda} \to \mathcal{M}_1(\mathbb{R}^d),$$

which for a given equation $(x_0, a, b) \in \mathcal{C}^{2\gamma+2}_{K,\lambda}$ uses x_0 and finitely many function values and derivative values of a and b to compute a probability measure $Q_n(x_0, a, b)$ on $\mathbb{R}^d$ with finite support that is close to the distribution of the solution of the corresponding equation (1) at time $t = 1$. Our deterministic algorithm is then defined by

$$\widehat{I}_n \colon \mathcal{C}^{2\gamma+2}_{K,\lambda} \times F^{r,\beta} \to \mathbb{R}, \quad (x_0, a, b, f) \mapsto \int f \, dQ_n(x_0, a, b).$$

The method Q_n employs a nonuniform discretization

$$0 = t_0 < \cdots < t_n = 1$$

of the unit time interval $[0, 1]$. For a given equation (x_0, a, b) it starts with the one-point distribution at the initial value x_0 and iteratively applies a Markov transition operator

$$T^{a,b}_{\tau_i} \colon \mathcal{M}_1(\mathbb{R}^d) \to \mathcal{M}_1(\mathbb{R}^d),$$

which is based on simplified weak order γ Itô-Taylor steps of size

$$\tau_i = t_i - t_{i-1}$$

associated with a and b, together with strategies

$$D_{\tau_i}, R_{\tau_i} \colon \mathcal{M}_1(\mathbb{R}^d) \to \mathcal{M}_1(\mathbb{R}^d)$$

to reduce the diameter and the cardinality of the support of a discrete measure, respectively, in order to avoid an explosion of the computational cost. Thus

$$Q_n(x_0, a, b) = R_{\tau_n} \circ D_{\tau_n} \circ T^{a,b}_{\tau_n} \circ \cdots \circ R_{\tau_1} \circ D_{\tau_1} \circ T^{a,b}_{\tau_1}(\delta_{x_0}). \tag{9}$$

We turn to the definition of the time discretization and the operators $T^{a,b}_\tau$, D_τ and R_τ for general $\tau \in (0, 1]$. To this end we employ the smoothness parameter γ and we choose a simplification ξ of order γ as well as further parameters $\eta, \varepsilon, \kappa > 0$. The parameter η determines the size of the single time-steps τ_i, the simplification ξ is used to construct the transition operator $T^{a,b}_\tau$, the parameters ε and κ specify the support diameter reduction strategy D_τ, and the support cardinality reduction strategy R_τ is determined by the parameter γ.

4.1. **The non-uniform time discretization.** We employ a time discretization that was introduced in [10] and is also used in [7] for approximation of certain stochastic integrals. Choose $\eta > 0$ and define

$$t_i = 1 - (1 - i/n)^\eta, \quad i = 0, \dots, n.$$

4.2. **The Itô-Taylor transition.** Choose a simplification ξ of order γ. For coefficients $a \in (C_K^{2\gamma+2})^d$ and $b \in (C_K^{2\gamma+2})^{d\times m}$ consider the associated simplified weak order γ Itô-Taylor steps $Y^{y,a,b}(\tau)$ given by (8) and define

$$T_\tau^{a,b}(\mu) = \sum_{y\in \mathrm{supp}(\mu)} \mu(\{y\})\, \mathbb{P}_{Y^{y,a,b}(\tau)}$$

for $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

4.3. **The support diameter reduction strategy.** Choose $\varepsilon, \kappa > 0$. Let $H_\tau = [-\kappa\tau^{-\varepsilon}, \kappa\tau^{-\varepsilon}]^d$ and project points $y \in \mathbb{R}^d$ onto H_τ by

$$\lfloor y \rfloor_{H_\tau} = \bigl(y_i + (-y_i - \kappa\tau^{-\varepsilon})_+ - (y_i - \kappa\tau^{-\varepsilon})_+\bigr)_{i=1,\dots,d}.$$

Define

$$D_\tau(\mu) = \sum_{y\in \mathrm{supp}(\mu)} \mu(\{y\})\, \delta_{\lfloor y \rfloor_{H_\tau}}$$

for $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

4.4. **The support cardinality reduction strategy.** By a well-known sequential support point elimination procedure, see [2], we obtain an algorithm

$$R\colon \mathcal{M}(\mathbb{R}^d) \to \mathcal{M}(\mathbb{R}^d),$$

which for every $\mu \in \mathcal{M}(\mathbb{R}^d)$ satisfies

$$\mathrm{supp}(R(\mu)) \subset \mathrm{supp}(\mu) \tag{10}$$

as well as

$$|\mathrm{supp}(R(\mu))| \le \dim(\mathcal{P}_{2\gamma+1}(\mathbb{R}^d)) = \binom{2\gamma+1+d}{d} \tag{11}$$

and

$$\int_{\mathbb{R}^d} p\, dR(\mu) = \int_{\mathbb{R}^d} p\, d\mu \tag{12}$$

for every $p \in \mathcal{P}_{2\gamma+1}(\mathbb{R}^d)$. See Section 9.1 in the appendix for a detailed description of the algorithm R.

We employ a localized variant of this method, which is due to [12]. Put

$$A_{j,\tau} = [\sqrt{\tau}\, j,\ \sqrt{\tau}\,(j+1))$$

for $j \in \mathbb{Z}^d$. Let

$$J_{\mu,\tau} = \{j \in \mathbb{Z}^d \colon A_{j,\tau} \cap \operatorname{supp}(\mu) \neq \emptyset\}$$

and define

$$R_\tau(\mu) = \sum_{j \in J_{\mu,\tau}} R\biggl(\sum_{y \in \operatorname{supp}(\mu) \cap A_{j,\tau}} \mu(\{y\})\, \delta_y \biggr)$$

for $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

5. Analysis of Cost and Error

For the analysis of the computational cost of the algorithm $\widehat{I}_n$ we employ the real number model of computation, see [18] for a formal definition. In particular, computations with real numbers and with elementary functions like exp and floor are carried out exactly and for a given input $(x_0, a, b, f) \in \mathcal{C}^{2\gamma+2}_{K,\lambda} \times F^{r,\beta}$ we have access to oracles, which provide function values and derivative values of the components of the coefficients a and b as well as of the integrand f at arbitrary points in $\mathbb{R}^d$.

The cost of applying $\widehat{I}_n$ to an input (x_0, a, b, f) is denoted by $\operatorname{cost}(\widehat{I}_n, (x_0, a, b, f))$ and is given by the sum of

(i) the number of evaluations of a_i, $b_{i,j}$, $\partial a_i/\partial x_k$, $\partial b_{i,j}/\partial x_k$ etc.,
(ii) the number of evaluations of f, $\partial f/\partial x_k$ etc., and
(iii) the number of basic computational operations, i.e. arithmetical operations and evaluations of elementary functions,

that are carried out by $\widehat{I}_n$ for this input.

Let

$$q = \min(2\gamma + 2, r)$$

denote the minimum smoothness of the coefficients a, b and the integrands f and recall that the algorithm $\widehat{I}_n$ is specified by a simplification ξ of order γ and further parameters $\eta, \varepsilon, \kappa > 0$. We then have the following estimates of the cost and the error of $\widehat{I}_n$.

Theorem 1. *For all $(x_0, a, b) \in \mathcal{C}^{2\gamma+2}_{K,\lambda}$, $f \in F^{r,\beta}$ and $n \in \mathbb{N}$,*

(i) $$\operatorname{cost}(\widehat{I}_n, (x_0, a, b, f)) \leq c \cdot \begin{cases} n^{1+d(\frac{1}{2}+\varepsilon)}, & \text{if } \eta < \frac{1}{d(\frac{1}{2}+\varepsilon)} + 1, \\ n^{1+d(\frac{1}{2}+\varepsilon)} \cdot \ln(n), & \text{if } \eta = \frac{1}{d(\frac{1}{2}+\varepsilon)} + 1, \\ n^{\eta d(\frac{1}{2}+\varepsilon)}, & \text{if } \eta > \frac{1}{d(\frac{1}{2}+\varepsilon)} + 1, \end{cases}$$

where c only depends on $d, m, \gamma, \xi, \eta, \varepsilon$ and κ, and

(ii) $$|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)| \leq c \cdot \|f\|_{r,\beta} \cdot \begin{cases} n^{-\frac{\eta q}{2}}, & \text{if } \eta < \frac{2\gamma}{q}, \\ n^{-\frac{\eta q}{2}} \ln(n), & \text{if } \eta = \frac{2\gamma}{q}, \\ n^{-\gamma}, & \text{if } \eta > \frac{2\gamma}{q}, \end{cases}$$

where c only depends on $d, m, r, \beta, \gamma, K, \lambda, \xi, \eta, \varepsilon$ and κ.

We proceed with the analysis of the relation of the error and the cost of the algorithm $\widehat{I}_n$ in a worst case setting with respect to the equations and the integrands.

Put
$$F_K^{r,\beta} = \{f \in F^{r,\beta} \colon \|f\|_{r,\beta} \le K\}$$
and consider the worst case error
$$\operatorname{err}(\widehat{I}_n) = \sup_{(x_0,a,b,f)\in \mathcal{C}_{K,\lambda}^{2\gamma+2}\times F_K^{r,\beta}} |I(x_0,a,b,f) - \widehat{I}_n(x_0,a,b,f)|$$
as well as the worst case cost
$$\operatorname{cost}(\widehat{I}_n) = \sup_{(x_0,a,b,f)\in \mathcal{C}_{K,\lambda}^{2\gamma+2}\times F_K^{r,\beta}} \operatorname{cost}(\widehat{I}_n, (x_0,a,b,f))$$
of the algorithm $\widehat{I}_n$ over the class $\mathcal{C}_{K,\lambda}^{2\gamma+2} \times F_K^{r,\beta}$.

We distinguish the cases
$$\frac{1}{d(1/2+\varepsilon)} + 1 \gtreqless \frac{2\gamma}{q}$$
and we use Theorem 1 to obtain estimates of the worst case errors of the algorithms $\widehat{I}_n$ in terms of their worst case cost, where the time discretization parameter η is chosen in the most favorable way.

Corollary 1. *For all $n \in \mathbb{N}$,*
$$\operatorname{err}(\widehat{I}_n) \le c \cdot \begin{cases} (\operatorname{cost}(\widehat{I}_n))^{-\frac{q}{d(1+2\varepsilon)}}, & \text{if } \frac{1}{d(1/2+\varepsilon)} + 1 < \eta < \frac{2\gamma}{q}, \\ (\operatorname{cost}(\widehat{I}_n))^{-\frac{q}{d(1+2\varepsilon)}} \cdot (\ln(\operatorname{cost}(\widehat{I}_n)))^{1+\frac{q}{d(1+2\varepsilon)}}, & \text{if } \frac{1}{d(1/2+\varepsilon)} + 1 = \eta = \frac{2\gamma}{q}, \\ (\operatorname{cost}(\widehat{I}_n))^{-\frac{2\gamma}{2+d(1+2\varepsilon)}}, & \text{if } \frac{2\gamma}{q} < \eta < \frac{1}{d(1/2+\varepsilon)} + 1, \end{cases}$$
where c only depends on $d, m, r, \beta, \gamma, K, \lambda, \xi, \eta, \varepsilon$ and κ.

Remark 1. It is easy to see that
$$\Big(\forall \varepsilon > 0\colon \frac{1}{d(1/2+\varepsilon)} + 1 < \frac{2\gamma}{q}\Big) \;\Leftrightarrow\; r \le \frac{2\gamma d}{d+2}.$$
Thus, if $r \le 2\gamma d/(d+2)$ then by Corollary 1 we can achieve, up to an arbitrary small exponent, the order
$$q/d = \min(2\gamma+2, r)/d = r/d$$
for the worst case errors of the algorithms $\widehat{I}_n$ in terms of their worst case cost. In Section 6 we will show that under the respective smoothness assumptions the order $\min(2\gamma+2, r)/d$ is in fact best possible for algorithms based on finitely many evaluations of the coefficients a, b, the integrand f and their derivatives.

In the case $r > 2\gamma d/(d+2)$ we almost achieve the order $2\gamma/(d+2)$, which, however, is suboptimal.

Remark 2. We address the question whether the estimates in Theorem 1 and Corollary 1 can also be achieved without employing the support diameter reduction strategy D_τ, see Section 4.3.

We therefore consider the algorithm

$$\widetilde{I}_n \colon \mathcal{C}^{2\gamma+2}_{K,\lambda} \times F^{r,\beta} \to \mathbb{R}, \quad (x_0, a, b, f) \mapsto \int f\, d\widetilde{Q}_n(x_0, a, b)$$

with

$$\widetilde{Q}_n(x_0, a, b) = R_{\tau_n} \circ T^{a,b}_{\tau_n} \circ \cdots \circ R_{\tau_1} \circ T^{a,b}_{\tau_1}(\delta_{x_0}),$$

i.e., the successive application of the operator D_τ in the definition (9) of the method Q_n is skipped.

By an obvious slight modification of the proof of Theorem 1(ii) in Section 8.2 we obtain, up to a constant factor, the same error estimate for $\widetilde{I}_n$ as for $\widehat{I}_n$. However, we only get

$$\text{(13)} \qquad \operatorname{cost}(\widetilde{I}_n, (x_0, a, b, f)) \le c \cdot \begin{cases} n^{1+d}, & \text{if } \eta < \frac{2}{d} + 1, \\ n^{1+d} \cdot \ln(n), & \text{if } \eta = \frac{2}{d} + 1, \\ n^{\frac{(\eta+1)d}{2}}, & \text{if } \eta > \frac{2}{d} + 1, \end{cases}$$

where c may depend on d, m, γ, K, ξ and η, as an upper bound for the cost of applying $\widetilde{I}_n$ to the input (x_0, a, b, f). See Section 8.3 for a proof of (13).

We conjecture that both the error estimate and the cost estimate for $\widetilde{I}_n$ are sharp. If the conjecture is true then the order of the worst case errors of the algorithms $\widetilde{I}_n$ in terms of their worst case cost on the class of inputs $\mathcal{C}^{2\gamma+2}_{K,\lambda} \times F^{r,\beta}_K$ can not exceed the value

$$\sup_{\eta \in [2/d+1,\, 2\gamma/q]} \frac{q}{d} \cdot \frac{\eta}{\eta+1} = \frac{q}{d} \cdot \frac{2\gamma}{2\gamma+q} = \frac{r}{d} \cdot \frac{2\gamma}{2\gamma+r}$$

in the case $r \le 2\gamma d/(d+2)$, which is suboptimal, see Remark 1.

Remark 3. Consider a Lipschitz continuous integrand $f : \mathbb{R}^d \to \mathbb{R}$. Clearly, the upper bounds in Theorem 1(i) and (13) for the cost of applying $\widehat{I}_n$ and $\widetilde{I}_n$ to the input (x_0, a, b, f) still hold. Furthermore, by a slight modification of the proof of Theorem 1(ii) in Section 8.2 one can show that for both errors $|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)|$ and $|I(x_0, a, b, f) - \widetilde{I}_n(x_0, a, b, f)|$ the upper bound in Theorem 1(ii) with $q = r = 1$, $\beta = 0$ and $\|f\|_{1,0}$ replaced by the Lipschitz seminorm of f is valid.

As a consequence, the worst case estimates in Corollary 1 with $q = 1$ and $\beta = 0$ hold true for the class of Lipschitz continuous integrands with Lipschitz seminorm bounded by K.

6. Lower bounds and optimality properties

Recall the results in Corollary 1 on the performance of the algorithms $\widehat{I}_n$ in a worst case setting with respect to the class of equations $\mathcal{C}^{2\gamma+2}_{K,\lambda}$ and the class of integrands $F^{r,\beta}_K$. It is natural to ask whether these results can be improved, i.e., whether it is possible to construct deterministic algorithms for the computational problem (2) that are based on the initial value x_0 and finitely many function values of the integrand f, the coefficients a, b and derivatives of these mappings, and achieve, asymptotically, worst case errors of a higher order in terms of their worst case cost than the orders provided by the bounds from Corollary 1. Motivated by this question we consider general worst case settings defined by smoothness constraints on the coefficients a, b and the integrands f, and we present corresponding lower worst case error bounds for any such deterministic algorithm that satisfies a given worst case cost constraint.

Fix $r \in \mathbb{N}$ and $\beta \geq 0$ as well as $s_1, s_2 \in \mathbb{N}$ and $K \geq 1 \geq \lambda > 0$. We consider the class of integrands $F^{r,\beta}_K$ and the class of equations

$$\mathcal{C}^{s_1,s_2}_{K,\lambda} = \Big\{(x_0, a, b) \in [-K, K]^d \times (C^{s_1}_K)^d \times (C^{s_2}_K)^{d\times m} \colon \inf_{x\in\mathbb{R}^d} \lambda_{\min}(bb^T(x)) \geq \lambda\Big\}.$$

For $(x_0, a, b) \in \mathcal{C}^{s_1,s_2}_{K,\lambda}$ and $f \in F^{r,\beta}_K$ we put

$$\mathbf{a} = \big(a_i^{(\alpha)}\big)_{\substack{1\leq i\leq d\\ 0\leq|\alpha|_1\leq s_1}}, \ \mathbf{b} = \big(b_{i,j}^{(\alpha)}\big)_{\substack{1\leq i\leq d\\ 1\leq j\leq m\\ 0\leq|\alpha|_1\leq s_2}}, \ \mathbf{f} = \big(f^{(\alpha)}\big)_{0\leq|\alpha|_1\leq r}.$$

Thus

$$(\mathbf{a}, \mathbf{b}, \mathbf{f}) \colon \mathbb{R}^d \to \mathbb{R}^L$$

with

$$L = d \cdot \binom{d + s_1}{d - 1} + d \cdot m \cdot \binom{d + s_2}{d - 1} + \binom{d + r}{d - 1}.$$

We use an integer $N \in \mathbb{N}$ as well as mappings

$$\psi_\ell \colon \mathbb{R}^{d+(\ell-1)\cdot L} \to \mathbb{R}^d, \quad \ell = 1, \dots, N,$$

and

$$\varphi \colon \mathbb{R}^{d+N\cdot L} \to \mathbb{R}$$

to describe a (generalized) deterministic algorithm that computes approximations to $I(x_0, a, b, f)$ for $(x_0, a, b) \in \mathcal{C}^{s_1,s_2}_{K,\lambda}$ and $f \in F^{r,\beta}_K$ based on x_0 and on finitely many sequential evaluations of $(\mathbf{a}, \mathbf{b}, \mathbf{f})$ at points in $\mathbb{R}^d$.

The mapping ψ_ℓ is employed to determine the node for the ℓ-th evaluation of $(\mathbf{a}, \mathbf{b}, \mathbf{f})$ based on the data from the previous $\ell-1$ evaluations of $(\mathbf{a}, \mathbf{b}, \mathbf{f})$, and the mapping φ is used to obtain the approximation to $I(x_0, a, b, f)$ based on the data of all N evaluations of $(\mathbf{a}, \mathbf{b}, \mathbf{f})$. More precisely, we define a mapping

$$\mathcal{N}_{\psi_1,\dots,\psi_N} \colon \mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K \to \mathbb{R}^{d+N\cdot L}$$

by

$$\mathcal{N}_{\psi_1,\dots,\psi_N}(x_0,a,b,f) = (x_0, y_1, \dots, y_N)$$

with

$$y_1 = (\mathbf{a},\mathbf{b},\mathbf{f})(\psi_1(x_0)) \in \mathbb{R}^L$$

and

$$y_\ell = (\mathbf{a},\mathbf{b},\mathbf{f})(\psi_\ell(x_0, y_1, \dots, y_{\ell-1})) \in \mathbb{R}^L$$

for $\ell = 2, \dots, N$. Then

$$\varphi \circ \mathcal{N}_{\psi_1,\dots,\psi_N} \colon \mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K \to \mathbb{R} \tag{14}$$

defines a (generalized) deterministic algorithm, which for a given input (x_0, a, b, f) sequentially evaluates the corresponding mapping $(\mathbf{a},\mathbf{b},\mathbf{f})$ at the points

$$\psi_\ell(x_0, y_1, \dots, y_{\ell-1}) \in \mathbb{R}^d, \quad \ell = 1, \dots, N,$$

and finally applies the mapping φ to the resulting data $\mathcal{N}_{\psi_1,\dots,\psi_N}(x_0,a,b,f)$ to obtain the real number $\varphi(\mathcal{N}_{\psi_1,\dots,\psi_N}(x_0,a,b,f))$ as an approximation to $I(x_0,a,b,f)$.

We use

$$\mathcal{I}_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) = \left\{\varphi \circ \mathcal{N}_{\psi_1,\dots,\psi_N} \colon \varphi \colon \mathbb{R}^{d+NL} \to \mathbb{R},\ \psi_\ell \colon \mathbb{R}^{d+(\ell-1)L} \to \mathbb{R}^d, \ell = 1, \dots, N\right\}$$

to denote the class of all deterministic algorithms of the form (14) and we put

$$\mathcal{I}(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) = \bigcup_{N\in\mathbb{N}} \mathcal{I}_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K).$$

The worst case error of $\widehat{I} \in \mathcal{I}(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)$ is defined by

$$\operatorname{err}(\widehat{I}) = \sup_{(x_0,a,b,f)\in\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K} |I(x_0,a,b,f) - \widehat{I}(x_0,a,b,f)|$$

and for every $N \in \mathbb{N}$ we put

$$e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) = \inf\{\operatorname{err}(\widehat{I}) \colon \widehat{I} \in \mathcal{I}_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)\}.$$

Thus, $e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)$ is the smallest possible worst case error on the class of inputs $\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K$ that can be achieved by any deterministic algorithm $\widehat{I}$, which uses the initial value x_0 and at most N sequential evaluations of $(\mathbf{a},\mathbf{b},\mathbf{f})$ to compute an approximation $\widehat{I}(x_0,a,b,f)$ to $I(x_0,a,b,f)$.

From [17] we obtain the following lower bounds for the minimal errors $e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)$.

Theorem 2. *For every $N \in \mathbb{N}$,*

$$e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) \geq c \cdot N^{-\frac{\min(s_1,s_2,r)}{d}},$$

where c only depends on d, s_1, s_2 and r.

Proof. We get $e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) \geq c(d, s_1, s_2) \cdot N^{-\min(s_1,s_2)/d}$ and $e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) \geq c(d, r) \cdot N^{-r/d}$ from [17, Theorem 4.2] and [17, (6.4)], respectively, together with [17, Remark 4.5]. The first two results are formulated in [17] only for the class of (generalized) deterministic algorithms based on function values of the coefficients a, b and the integrands f but are easily seen to be valid as well if, additionally, derivative values may be used. □

We now show that the lower bounds from Theorem 2 can almost be achieved by the algorithms $\widehat{I}_n$ introduced in Section 4 if the smoothness r of the integrands is not too large compared to the minimum smoothness $\min(s_1, s_2)$ of the coefficients a and b.

Assume $s_1, s_2 \geq 4$, let

$$\gamma = \lfloor \min(s_1, s_2)/2 \rfloor - 1$$

and consider the algorithms $\widehat{I}_n\colon \mathcal{C}^{2\gamma+2}_{K,\lambda} \times F^{r,\beta}_K \to \mathbb{R}$ introduced in Section 4 with any choice of the simplification ξ of order γ and of the parameters η, ε and κ. By definition of γ we have $\mathcal{C}^{s_1,s_2}_{K,\lambda} \subset \mathcal{C}^{2\gamma+2,2\gamma+2}_{K,\lambda}$. Thus

$$\widehat{I}_n \in \mathcal{I}_{N_n}(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)$$

with

$$N_n = \operatorname{cost}(\widehat{I}_n)$$

denoting the worst case cost of $\widehat{I}_n$ on the class $\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K$. Using Corollary 1, Remark 1 and Theorem 2 we get the following result, which shows that in the case $r \leq 2\gamma d/(d+2)$ the sequence of methods $\widehat{I}_n$ performs almost optimal if the time discretization parameter η is chosen in an appropriate way.

Theorem 3. *Assume $s_1, s_2 \geq 4$ and $r \leq 2\gamma d/(d + 2)$. If $\eta \in (1/(d(1/2 + \varepsilon)), 2\gamma/r)$ then for every $n \in \mathbb{N}$,*

$$c_1 \cdot N_n^{-\frac{r}{d}} \leq e_{N_n}(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) \leq \operatorname{err}(\widehat{I}_n) \leq c_2 \cdot N_n^{-\frac{r}{d(1+2\varepsilon)}},$$

where c_1 only depends on d and r and c_2 only depends on $d, m, r, \beta, \gamma, K, \lambda, \xi, \eta, \varepsilon$ and κ.

Remark 4. Recall from Theorem 2 the lower bounds $c \cdot N^{-\min(s_1,s_2,r)/d}$ for the minimal errors $e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K)$. By Theorem 3, under the respective restrictions on the smoothness parameters s_1, s_2 and r, these bounds are sharp, up to an arbitrary small power of N. It follows from the upper bound in [17, Theorem 4.1] together with [17, Remark 4.5] that the estimate

$$e_N(\mathcal{C}^{s_1,s_2}_{K,\lambda} \times F^{r,\beta}_K) \leq c \cdot N^{-\frac{\min(s_1,s_2,r)}{d}+\varepsilon},$$

where c only depends on $d, m, s_1, s_2, r, \beta, K$ and ε, is in fact valid for any $s_1, s_2, r \in \mathbb{N}$ and $\varepsilon > 0$. However, the proof of the upper bounds in [17, Theorem 4.1] is based on the construction of algorithms which, in contrast to the algorithms $\widehat{I}_n$, rely on heavy precomputation and are therefore of no practical relevance.

7. Numerical experiments

We illustrate our results by numerical experiments for the scalar equation given by

$$x_0 = 0,\ a = \sin,\ b = 1 + 0.3\cos, \tag{15}$$

and the integrand

$$f(x) = \max(0, 1 - e^x), \quad x \in \mathbb{R}. \tag{16}$$

In particular, $d = m = 1$ and $(x_0, a, b) \in \mathcal{C}^{2\gamma+2}_{K,\lambda}$ for any $\gamma \in \mathbb{N}$, $K \geq 1.3$ and $\lambda \in (0, 0.7]$. Moreover, f is Lipschitz continuous.

We take $\gamma = 2$ and consider the algorithms $\widehat{I}_n$ introduced in Section 4 using the simplification ξ of order 2 presented in Example 1 and the parameters

$$\eta = 3,\ \kappa = 3,\ \varepsilon = 0.05.$$

We employ $\widehat{I}_n$ with $n = 32000$ time steps to obtain the reference value

$$I_{\text{ref}}(x_0, a, b, f) = \widehat{I}_{32000}(x_0, a, b, f) = 0.3374637100...,$$

which is used in place of the unknown quantity $I(x_0, a, b, f)$ in the subsequent error estimates.

Figure 1 shows, on logarithmic scales, plots of the resulting error estimates

$$e_n = |I_{\text{ref}}(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)|$$

and

$$\widetilde{e}_n = |I_{\text{ref}}(x_0, a, b, f) - \widetilde{I}_n(x_0, a, b, f)|$$

for the method $\widehat{I}_n$ with diameter reduction and the method $\widetilde{I}_n$ without diameter reduction, respectively, versus the number n of time-steps for $n = \lfloor 5 \cdot 1.5^i \rfloor$ with $i = 0, \ldots, 17$. In both cases we computed a regression line based on the last 14 data points, i.e. based on (n, e_n) and $(n, \widetilde{e}_n)$ with $n = \lfloor 5 \cdot 1.5^4 \rfloor, \ldots, \lfloor 5 \cdot 1.5^{17} \rfloor$. The corresponding slopes are approximately -2.04 in both cases, which is consistent with an order of convergence of at least 1.5 for the respective true errors $|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)|$ and $|I(x_0, a, b, f) - \widetilde{I}_n(x_0, a, b, f)|$ in terms of n due to Theorem 1(ii), Remark 2 and Remark 3.

Figure 2 shows, on logarithmic scales, plots of the error estimates e_n and $\widetilde{e}_n$ versus the computational cost, $\text{cost}(\widehat{I}_n, (x_0, a, b, f))$ and $\text{cost}(\widetilde{I}_n, (x_0, a, b, f))$, for $n = \lfloor 5 \cdot 1.5^i \rfloor$ with $i = 0, \ldots, 17$ and $i = 0, \ldots, 15$, respectively. In both cases we computed a regression line based on all but the first 2 data points. The respective slopes are approximately -1.27 and -1.08, which is in accordance with an order of convergence of at least $1/(1+2\varepsilon) = 0.909...$ for the true errors $|I(x_0, a, b, f) - \widehat{I}_n(x_0, a, b, f)|$ of the method with diameter reduction and $\eta/(\eta+1) = 0.75$ for the true errors $|I(x_0, a, b, f) - \widetilde{I}_n(x_0, a, b, f)|$ of the method without diameter reduction in terms of the computational cost, see

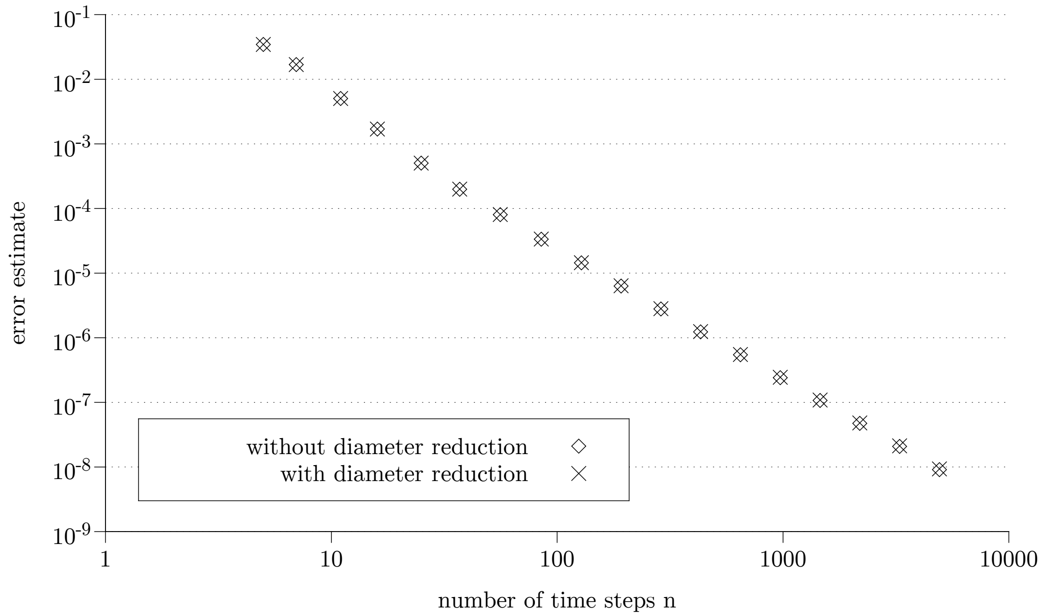

FIGURE 1. Error vs. number of time steps

Corollary 1, Remark 2 and Remark 3. We thus gain a power of almost 0.2 by employing the diameter reduction strategy.

Additionally, we consider two randomized methods for computing $I(x_0, a, b, f)$, namely the classical Monte Carlo Euler method and the multilevel Monte Carlo Euler method, which has been introduced in [6] for the computation of global solutions of integral equations and, independently, in [8] for quadrature of diffusion processes.

Let $k \in \mathbb{N}$ and let $Z_1, \dots, Z_k$ denote a sequence of independent standard normally distributed random variables. The weak Euler scheme $(\widehat{X}_\ell)_{\ell=0,\dots,k}$ with k equidistant steps is iteratively defined by

$$\begin{aligned} \widehat{X}_0 &= x_0, \\ \widehat{X}_\ell &= \widehat{X}_{\ell-1} + k^{-1} \cdot a(\widehat{X}_{\ell-1}) + k^{-1/2} \cdot b(\widehat{X}_{\ell-1}) \cdot Z_\ell \end{aligned} \tag{17}$$

for $\ell = 1, \dots, k$. Take k independent copies $\widehat{X}_k^{(1)}, \dots, \widehat{X}_k^{(k)}$ of $\widehat{X}_k$ and consider the Monte Carlo Euler approximation of $I(x_0, a, b, f)$ given by

$$\widehat{I}_k^E(x_0, a, b, f) = \frac{1}{k} \sum_{j=1}^{k} f\big(\widehat{X}_k^{(j)}\big).$$

Then

$$\mathbb{E}(I(x_0, a, b, f) - \widehat{I}_k^E(x_0, a, b, f))^2 \leq c \cdot k^{-1} \tag{18}$$

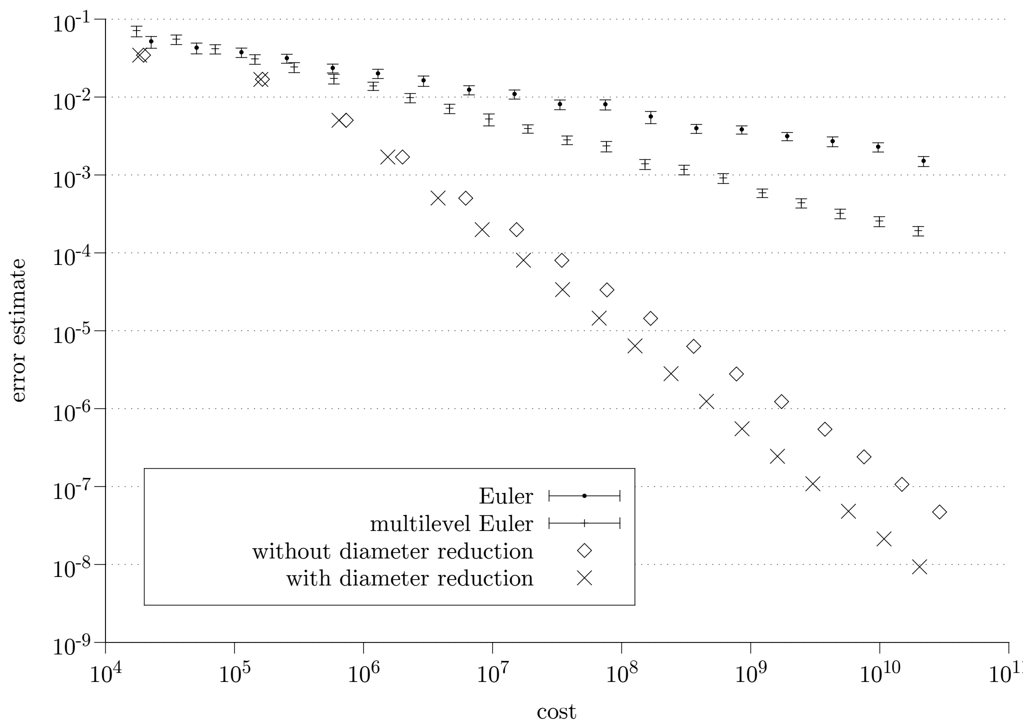

Figure 2. Error vs. computational cost

due to well known results on the Euler scheme. We define the computational cost $\mathrm{cost}(\widehat{I}_k^E, (x_0, a, b, f))$ to compute a single realization of the approximation $\widehat{I}_k^E(x_0, a, b, f)$ in the same way as the computational cost to compute $\widehat{I}_n(x_0, a, b, f)$, see Section 6, up to additionally adding the number k of calls to a random number generator for the standard normal distribution. Then $\mathrm{cost}(\widehat{I}_k^E, (x_0, a, b, f))$ is deterministic and satisfies

$$\mathrm{cost}(\widehat{I}_k^E, (x_0, a, b, f)) \leq c \cdot k^2.$$

As a consequence,

$$\big(\mathbb{E}(I(x_0, a, b, f) - \widehat{I}_k^E(x_0, a, b, f))^2\big)^{1/2} \leq c \cdot \mathrm{cost}(\widehat{I}_k^E, (x_0, a, b, f))^{1/4}. \tag{19}$$

The multilevel approach is based on coupled pairs of weak Euler schemes. For $j \in \mathbb{N}$ we consider the scheme $(\widehat{X}_\ell)_{\ell=0,\dots,2j}$ given by (17) with $k = 2j$ together with the scheme $(\widetilde{X}_\ell)_{\ell=0,\dots,j}$ given by $\widetilde{X}_0 = x_0$ and

$$\widetilde{X}_\ell = \widetilde{X}_{\ell-1} + j^{-1} \cdot a(\widetilde{X}_{\ell-1}) + j^{-1/2} \cdot b(\widetilde{X}_{\ell-1}) \cdot (Z_{2\ell-1} + Z_{2\ell})/\sqrt{2}$$

for $\ell = 1, \dots, j$, and we use

$$(\widehat{X}_{j,1}, \widehat{X}_{j,2}) = (\widetilde{X}_j, \widehat{X}_{2j})$$

to denote the pair of the corresponding final steps.

Let $n \in \mathbb{N}$ with $n \geq 2$. Put $L = \lceil \log_2 n \rceil$ as well as

$$k_\ell = \lceil 2^{L-\ell}/L \rceil$$

for $\ell = 0, \ldots, L$. Take k_0 independent copies $\widehat{X}_{1,1}^{(1)}, \ldots, \widehat{X}_{1,1}^{(k_0)}$ of $\widehat{X}_{1,1}$ and k_ℓ independent copies $(\widehat{X}_{2^\ell,1}^{(1)}, \widehat{X}_{2^\ell,2}^{(1)}), \ldots, (\widehat{X}_{2^\ell,1}^{(k_\ell)}, \widehat{X}_{2^\ell,2}^{(k_\ell)})$ of $(\widehat{X}_{2^\ell,1}, \widehat{X}_{2^\ell,2})$ for $\ell = 1, \ldots, L$, such that the resulting $k_0 + \cdots + k_L$ random variables are independent as well, and consider the multilevel Monte Carlo Euler approximation of $I(x_0, a, b, f)$ given by

$$\widehat{I}_n^{mE} = \frac{1}{k_0} \sum_{j=1}^{k_0} f(\widehat{X}_{1,1}^{(j)}) + \sum_{l=1}^{L} \frac{1}{k_l} \sum_{j=1}^{k_l} \big(f(\widehat{X}_{2^l,2}^{(j)}) - f(\widehat{X}_{2^l,1}^{(j)}) \big).$$

Using (18) it is easy to see that

$$\mathbb{E}(I(x_0, a, b, f) - \widehat{I}_n^{mE}(x_0, a, b, f))^2 \leq c \cdot (\ln n)/n,$$

and furthermore we have

$$\operatorname{cost}(\widehat{I}_n^{mE}, (x_0, a, b, f)) \leq c \cdot n,$$

where the computational cost $\operatorname{cost}(\widehat{I}_n^{mE}, (x_0, a, b, f))$ to compute a realization of the approximation $\widehat{I}_n^{mE}(x_0, a, b, f)$ is defined analogously to $\operatorname{cost}(\widehat{I}_k^{E}, (x_0, a, b, f))$ and is deterministic as well. Thus

$$\begin{aligned} &\big(\mathbb{E}(I(x_0, a, b, f) - \widehat{I}_n^{mE}(x_0, a, b, f))^2\big)^{1/2} \\ &\qquad \leq c \cdot \big(\ln(\operatorname{cost}(\widehat{I}_n^{mE}, (x_0, a, b, f)))\big)^{1/2} \cdot \big(\operatorname{cost}(\widehat{I}_n^{mE}, (x_0, a, b, f))\big)^{-1/2}. \end{aligned} \tag{20}$$

We estimate the root mean squared errors of the Monte Carlo Euler approximation $I_n^E(x_0, a, b, f)$ and the multilevel Monte Carlo Euler approximation $I_n^{mE}(x_0, a, b, f)$ by

$$e_n^E = \Big(\frac{1}{100} \sum_{i=1}^{100} (I_{\text{ref}}(x_0, a, b, f) - I_n^{E,i}(x_0, a, b, f))^2 \Big)^{1/2}$$

and

$$e_n^{mE} = \Big(\frac{1}{100} \sum_{i=1}^{100} (I_{\text{ref}}(x_0, a, b, f) - I_n^{mE,i}(x_0, a, b, f))^2 \Big)^{1/2},$$

where $I_n^{E,1}(x_0, a, b, f), \ldots, I_n^{E,100}(x_0, a, b, f)$ and $I_n^{mE,1}(x_0, a, b, f), \ldots, I_n^{mE,100}(x_0, a, b, f)$ are independent copies of $I_n^E(x_0, a, b, f)$ and $I_n^{mE}(x_0, a, b, f)$, respectively. Figure 2 shows plots of realizations of e_n^E and e_n^{mE} versus the computational cost for $n = \lfloor 50 \cdot 1.5^i \rfloor$ with $i = 0, \ldots, 17$ and $n = 300 \cdot 2^i$ with $i = 0, \ldots, 20$, respectively, together with the resulting asymptotic 95% confidence intervals. The corresponding regression lines have slopes of approximately -0.25 for the Monte Carlo Euler approximation and approximately -0.43 for the multilevel Monte Carlo Euler approximation, which is in accordance with the respective bounds (19) and (20). Thus, both randomized methods perform much worse than the deterministic algorithm $\widehat{I}_n$ for the given equation (15) and the given integrand (16) and, in particular, are far from achieving the error level of 10^{-8} with reasonable computational cost.

8. Proofs

Throughout this section we fix $r \in \mathbb{N}$, $\beta \geq 0$, $\gamma \in \mathbb{N}$ as well as $K \geq 1 \geq \lambda > 0$ and we consider a fixed equation $(x_0, a, b) \in \mathcal{C}^{2\gamma+2}_{K,\lambda}$. We use X^y to denote the solution of the corresponding equation (1) with initial value $y \in \mathbb{R}^d$ in place of x_0. The corresponding semigroup of linear operators

$$P_t \colon F^1 \to F^1, \quad t \in [0, \infty),$$

is given by

$$P_t f(y) = \mathbb{E} f(X^y(t)), \tag{21}$$

see Lemma 8 in the appendix. In particular,

$$I(x_0, a, b, f) = P_1 f(x_0).$$

Furthermore, we fix a simplification ξ of order γ and the parameters $n \in \mathbb{N}$ and $\eta, \varepsilon, \kappa > 0$, which are needed to specify the time-discretization and the operators $T^{a,b}_\tau$, D_τ, R_τ presented in Sections 4.1 to 4.4 and hereby determine the algorithms $\widehat{I}_n$ and $\widetilde{I}_n$.

We write T_τ and $Y^y(\tau)$ in place of $T^{a,b}_\tau$ and $Y^{y,a,b}(\tau)$, respectively. Moreover, we use Q^R_0 to denote the identity operator on $\mathcal{M}_1(\mathbb{R}^d)$ and we iteratively define

$$\begin{aligned} Q^T_\ell &= T_{\tau_\ell} \circ Q^R_{\ell-1}, \\ Q^D_\ell &= D_{\tau_\ell} \circ Q^T_\ell, \\ Q^R_\ell &= R_{\tau_\ell} \circ Q^D_\ell \end{aligned}$$

for $\ell = 1, \dots, n$. Thus, $Q_n(x_0, a, b) = Q^R_n(\delta_{x_0})$ and

$$\widehat{I}_n(x_0, a, b, f) = \mathbb{E}_{Q^R_n(\delta_{x_0})} f.$$

8.1. **Proof of Theorem 1(i).**

In this section unspecified positive constants c may only depend on $d, m, \gamma, \xi, \eta, \varepsilon$ and κ.

Let $\tau \in (0, 1]$ and $\mu \in \mathcal{M}_1$ and consider

$$\nu \in \{T_\tau(\mu), D_\tau(\mu), R_\tau(\mu)\}.$$

It is straightforward to obtain upper bounds for $|\operatorname{supp}(\nu)|$ as well as for the number $\operatorname{op}(\nu)$ of basic computational operations and the number $\operatorname{eval}_{a,b}(\nu)$ of evaluations of the coefficients a, b and derivatives thereof that are needed to compute the support points and the corresponding weights of the probability measure ν as presented in Table 1.

ν	$\lvert\operatorname{supp}(\nu)\rvert$	$\operatorname{op}(\nu)$	$\operatorname{eval}_{a,b}(\nu)$
$T_\tau(\mu)$	$c\cdot\lvert\operatorname{supp}(\mu)\rvert$	$c\cdot\lvert\operatorname{supp}(\mu)\rvert$	$c\cdot\lvert\operatorname{supp}(\mu)\rvert$
$D_\tau(\mu)$	$\lvert\operatorname{supp}(\mu)\rvert$	$c\cdot\lvert\operatorname{supp}(\mu)\rvert$	0
$R_\tau(\mu)$	$c\cdot\min(\lvert\operatorname{supp}(\mu)\rvert, \lvert J_{\mu,\tau}\rvert)$	$c\cdot\lvert\operatorname{supp}(\mu)\rvert$	0

TABLE 1. Upper bounds of the size of support, the numbers of basic computational operations and the number of evaluations of the coefficients and their derivatives for the operators T_τ, D_τ, R_τ.

By definition of D_τ we have $\operatorname{supp}(D_\tau(\mu)) \subset [-\kappa\tau^{-\varepsilon}, \kappa\tau^{-\varepsilon}]$. Hence

$$|J_{D_\tau(\mu),\tau}| \le (2\kappa\tau^{-(1/2+\varepsilon)} + 2)^d.$$

Using the support size estimates from Table 1 it follows

$$|\operatorname{supp}(Q_\ell^R(\delta_{x_0}))| = |\operatorname{supp}(R_{\tau_\ell}(Q_\ell^D(\delta_{x_0})))| \le c\cdot \left|J_{Q_\ell^D(\delta_{x_0}),\tau_\ell}\right| \le c\cdot \tau_\ell^{-d(1/2+\varepsilon)} \tag{22}$$

for $\ell = 1, \dots, n$, and, in particular,

$$|\operatorname{supp}(Q_n(x_0, a, b))| = |\operatorname{supp}(Q_n^R(\delta_{x_0}))| \le c\cdot \tau_n^{-d(1/2+\varepsilon)} = c\cdot n^{d\eta(1/2+\varepsilon)},$$

which provides the claimed bound for the number of evaluations of integrands f.

By the estimates in columns 3 and 4 of Table 1 we see that the cost to compute the support points and the corresponding weights of $R_\tau(D_\tau(T_\tau(\mu)))$ is bounded by $c\cdot|\operatorname{supp}(\mu)|$. Using (22) we conclude by induction that the cost to compute the support points and the corresponding weights of $Q_n(x_0, a, b)$ is bounded by

$$c\cdot\sum_{\ell=0}^{n-1} |\operatorname{supp}(Q_\ell^R(\delta_{x_0}))| \le c\cdot \sum_{\ell=1}^{n-1} \tau_\ell^{-d(1/2+\varepsilon)}.$$

By definition of τ_ℓ we have

$$\eta(n-\ell)^{\eta-1}\cdot n^{-\eta} \le \tau_\ell \le \eta(n-\ell+1)^{\eta-1}\cdot n^{-\eta} \tag{23}$$

and therefore

$$\begin{aligned}\sum_{\ell=1}^{n-1} \tau_\ell^{-d(1/2+\varepsilon)} &\le \eta^{-d(1/2+\varepsilon)} n^{\eta d(1/2+\varepsilon)} \cdot \sum_{\ell=1}^{n-1} \frac{1}{\ell^{(\eta-1)d(1/2+\varepsilon)}} \\ &\le c\cdot \begin{cases} n^{1+d(1/2+\varepsilon)}, & \text{if } \eta < (d(1/2+\varepsilon))^{-1}+1, \\ n^{1+d(1/2+\varepsilon)}\cdot \ln(n) & \text{if } \eta = (d(1/2+\varepsilon))^{-1}+1, \\ n^{\eta d(1/2+\varepsilon)} & \text{if } \eta > (d(1/2+\varepsilon))^{-1}+1, \end{cases}\end{aligned}$$

which completes the proof of the cost estimate.

8.2. **Proof of Theorem 1(ii).** In this section unspecified positive constants c and $c(\cdot)$ may only depend on the parameters eventually specified in brackets and on $d, m, r, \beta, \gamma, K, \lambda, \xi, \eta, \varepsilon$ and κ.

Let $\tau \in (0,1]$ and $y \in \mathbb{R}^d$. For $\alpha \in \mathbb{N}_0^d$ we put

$$\Delta_\alpha^y(\tau) = \Big| \mathbb{E} \prod_{i=1}^d (X_i^y(\tau) - y_i)^{\alpha_i} - \mathbb{E} \prod_{i=1}^d (Y_i^y(\tau) - y_i)^{\alpha_i} \Big| \tag{24}$$

to compare moments of $X^y(\tau)$ and the simplified weak order γ Itô-Taylor steps $Y^y(\tau)$.

Lemma 1. *For all $\tau \in (0,1]$, $y \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha|_1 \le 2\gamma + 1$ we have*

$$\Delta_\alpha^y(\tau) \le c \cdot (1 + |y|^c) \cdot \tau^{\gamma+1}. \tag{25}$$

Proof. Let $\tau \in (0,1]$, $y \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$. By (5) we have

$$\Big| \mathbb{E} \prod_{\ell=1}^L J_{\alpha^{(\ell)}}(\tau) - \mathbb{E} \prod_{\ell=1}^L \xi_{\alpha^{(\ell)}}(\tau) \Big| = \tau^{\sum_{\ell=1}^L \deg(\alpha^{(\ell)})/2} \cdot \Big| \mathbb{E} \prod_{\ell=1}^L J_{\alpha^{(\ell)}}(1) - \mathbb{E} \prod_{\ell=1}^L \xi_{\alpha^{(\ell)}} \Big| \tag{26}$$

for all $\alpha^{(1)}, \dots, \alpha^{(L)} \in \Gamma_\gamma$ and $L = 1, \dots, 2\gamma+1$. Since ξ is a simplification of order γ, the right hand side of (26) vanishes if $\deg(\alpha^{(1)}) + \ldots + \deg(\alpha^{(L)}) \le 2\gamma + 1$. Otherwise the right hand side of (26) is bounded by $c \cdot \tau^{\sum_{\ell=1}^L \deg(\alpha^{(\ell)})/2} \le c \cdot \tau^{\gamma+1}$.

Hence, by [9, Corollary 5.12.1],

$$\Big| \mathbb{E} \prod_{i=1}^d (X_i^y(\tau))^{\rho_i} - \mathbb{E} \prod_{i=1}^d (Y_i^y(\tau))^{\rho_i} \Big| \le c \cdot (1 + |y|^c) \cdot \tau^{\gamma+1}$$

for all $\rho \in \mathbb{N}_0^d$ with $|\rho|_1 \le 2\gamma + 1$.

It remains to observe that

$$\Delta_\alpha^y(\tau) = \sum_{\rho_1=0}^{\alpha_1} \cdots \sum_{\rho_d=0}^{\alpha_d} \prod_{i=1}^d \binom{\alpha_i}{\rho_i} \cdot y_i^{\alpha_i - \rho_i} \cdot \Big(\mathbb{E} \prod_{i=1}^d (X_i^y(\tau))^{\rho_i} - \mathbb{E} \prod_{i=1}^d (Y_i^y(\tau))^{\rho_i} \Big).$$

□

Next, we estimate the deviation of the simplified weak Itô-Taylor step $Y^y(\tau)$ from its initial value y.

Lemma 2. *For all $\tau \in (0,1]$ and $y \in \mathbb{R}^d$ we have*

$$|Y^y(\tau) - y| \le c \cdot \tau^{1/2}.$$

Proof. Since $a_i, b_{i,j} \in C_K^{2\gamma+2}$ for $i = 1, \dots, d$ and $j = 1, \dots, m$ we have

$$\|\psi_\alpha\| \le c \tag{27}$$

for all of the associated Itô-coefficient functions ψ_α with $\alpha \in \Gamma_\gamma$. Hence

$$|Y^y(\tau) - y| = \Big|\sum_{\alpha\in\Gamma_\gamma} \psi_\alpha(y)\, \tau^{\deg(\alpha)/2}\, \xi_\alpha\Big| \le c \cdot \tau^{1/2} \cdot \sum_{\alpha\in\Gamma_\gamma} |\xi_\alpha|.$$

Clearly,

$$|\xi_\alpha| \le \max\{|x| \colon x \in \mathrm{supp}(\xi_\alpha)\} \le c,$$

which finishes the proof. □

Next, we compare the distributions of $X^y(\tau)$ and $Y^y(\tau)$ with respect to integration of functions $f \in F^{k,\beta}$.

Lemma 3. *Let $\tau \in (0,1]$, $y \in \mathbb{R}^d$ and $k \in \mathbb{N}$. For every $f \in F^{k,\beta}$,*

$$|P_\tau f(y) - \mathbb{E}_{T_\tau(\delta_y)} f| \le c \cdot \|f\|_{k,\beta} \cdot (1 + |y|^c) \cdot \tau^{\min(\gamma+1,k/2)}.$$

Proof. Let $\tau \in (0,1]$, $y \in \mathbb{R}^d$, $k \in \mathbb{N}$ and let $f \in F^{k,\beta}$. Put $k^* = \min(2\gamma+2, k)$. For $z \in \mathbb{R}^d$,

$$f(z) = f(y) + \sum_{\alpha\in\mathbb{N}_0^d \colon 1\le|\alpha|_1\le k^*-1} \frac{f^{(\alpha)}(y)}{\alpha!}\,(z-y)^\alpha + \sum_{\alpha\in\mathbb{N}_0^d \colon |\alpha|_1 = k^*} \frac{f^{(\alpha)}(\theta_z)}{\alpha!}\,(z-y)^\alpha, \tag{28}$$

where $\theta_z \in \mathbb{R}^d$ satisfies $|\theta_z - y| \le |z-y|$. Hence

$$|f^{(\alpha)}(\theta_z)| \le \|f\|_{k,\beta} \cdot (1 + |\theta_z|^\beta) \le c \cdot \|f\|_{k,\beta} \cdot (1 + |z-y|^\beta + |y|^\beta) \tag{29}$$

for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha|_1 = k^*$ and, similarly,

$$|f^{(\alpha)}(y)| \le \|f\|_{k,\beta} \cdot (1 + |y|^\beta)$$

for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha|_1 \le k^* - 1$. Thus

$$\begin{aligned}
|P_\tau f(y) - \mathbb{E}_{T_\tau(\delta_y)} f| &= |\mathbb{E}(f(X^y(\tau)) - f(y)) - \mathbb{E}(f(Y^y(\tau)) - f(y))| \\
&\le \|f\|_{k,\beta} \cdot (1 + |y|^\beta) \cdot \sum_{\alpha\in\mathbb{N}_0^d \colon 1\le|\alpha|_1\le k^*-1} \Delta_\alpha^y(\tau) \\
&\quad + c \cdot \|f\|_{k,\beta} \cdot (1 + |y|^\beta) \cdot \big(\mathbb{E}|X^y(\tau) - y|^{k^*} + \mathbb{E}|Y^y(\tau) - y|^{k^*}\big) \\
&\quad + c \cdot \|f\|_{k,\beta} \cdot \big(\mathbb{E}|X^y(\tau) - y|^{k^*+\beta} + \mathbb{E}|Y^y(\tau) - y|^{k^*+\beta}\big).
\end{aligned}$$

Since $\|a\|, \|b\| \le K$, we have

$$\mathbb{E} \sup_{0\le u\le\tau} |X^y(u) - y|^p \le c(p) \cdot \tau^{p/2}$$

for all $p \in \mathbb{N}$, which jointly with Lemma 1 and Lemma 2 implies

$$|P_\tau f(y) - \mathbb{E}_{T_\tau(\delta_y)} f| \le c \cdot \|f\|_{k,\beta} \cdot (1 + |y|^c) \cdot \tau^{k^*/2}$$

and completes the proof of the lemma. □

We turn to the comparison of μ and $D_\tau(\mu)$ as well as of μ and $R_\tau(\mu)$ for $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

Lemma 4. *Let $\tau \in (0,1]$, $f \in F^{1,\beta}$ and $\mu \in \mathcal{M}_1(\mathbb{R}^d)$. For all $p \in \mathbb{N}$,*

$$|\mathbb{E}_\mu f - \mathbb{E}_{D_\tau(\mu)} f| \le c(p) \cdot \|f\|_{1,\beta} \cdot (1 + m_{1+\beta+p/\varepsilon}(\mu)) \cdot \tau^p.$$

Proof. Let $\tau \in (0,1]$, $f \in F^{1,\beta}$, $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $p \in \mathbb{N}$. By definition of D_τ,

$$\mathbb{E}_\mu f - \mathbb{E}_{D_\tau(\mu)} f = \int_{\mathbb{R}^d \setminus H_\tau} \big(f(y) - f(\lfloor y \rfloor_{H_\tau})\big)\, \mu(dy)$$

with $H_\tau = [-\kappa\tau^{-\varepsilon}, \kappa\tau^{-\varepsilon}]$. Let $y \notin H_\tau$. Then

$$f(y) - f(\lfloor y \rfloor_{H_\tau}) = \sum_{\alpha \in \mathbb{N}_0^d \colon |\alpha|_1 = 1} f^{(\alpha)}(\theta_y) \big(y - \lfloor y \rfloor_{H_\tau}\big)^\alpha,$$

where $\theta_y \in \mathbb{R}^d$ satisfies $|\theta_y| \le |y|$. Moreover, $|y - \lfloor y \rfloor_{H_\tau}| \le |y|$ and $|y|^{p/\varepsilon} > \kappa^{p/\varepsilon} \cdot \tau^{-p}$. Hence

$$\begin{aligned}
&\int_{\mathbb{R}^d \setminus H_\tau} \big|f(y) - f(\lfloor y \rfloor_{H_\tau})\big|\, \mu(dy) \\
&\qquad \le c \cdot \|f\|_{1,\beta} \cdot \int_{\mathbb{R}^d \setminus H_\tau} (1 + |y|^\beta) \cdot |y|\, \mu(dy) \\
&\qquad \le c \cdot \|f\|_{1,\beta} \cdot \kappa^{-p/\varepsilon} \cdot \tau^p \cdot \int_{\mathbb{R}^d \setminus H_\tau} (1 + |y|^{1+\beta+p/\varepsilon})\, \mu(dy) \\
&\qquad \le c(p) \cdot \|f\|_{1,\beta} \cdot \tau^p \cdot \Big(1 + \int_{\mathbb{R}^d} |y|^{1+\beta+p/\varepsilon}\, \mu(dy)\Big),
\end{aligned}$$

which finishes the proof. □

Lemma 5. *Let $\tau \in (0,1]$, $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $k \in \mathbb{N}$. For every $f \in F^{k,\beta}$,*

$$|\mathbb{E}_\mu f - \mathbb{E}_{R_\tau(\mu)} f| \le c \cdot \|f\|_{k,\beta} \cdot (1 + m_\beta(\mu) + m_\beta(R_\tau(\mu)) \cdot \tau^{\min(\gamma+1, k/2)}.$$

Proof. Let $\tau \in (0,1]$, $\mu \in \mathcal{M}_1(\mathbb{R}^d)$, $k \in \mathbb{N}$ and $f \in F^{k,\beta}$. Let $\mu_j = \mu_{|A_{j,\tau}}$ denote the restriction of μ to the cube $A_{j,\tau}$ and choose $z^{(j)} \in A_{j,\tau}$ for $j \in J_{\mu,\tau}$. Thus

$$\mu = \sum_{j \in J_{\mu,\tau}} \mu_j, \;\; R_\tau(\mu) = \sum_{j \in J_{\mu,\tau}} R(\mu_j).$$

Put $k^* = \min(2\gamma + 2, k)$. By the properties (10) and (12) of the operator R we have

$$\int_{A_{j,\tau}} (z - z^{(j)})^\alpha d\mu_j(z) = \int_{A_{j,\tau}} (z - z^{(j)})^\alpha dR(\mu_j)(z)$$

for all $j \in J_{\mu,\tau}$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha|_1 \le k^* - 1$. Using the Taylor expansion (28) of f at the points $y = z^{(j)}$ for $j \in J_{\mu,\tau}$ we get

$$\mathbb{E}_\mu f - \mathbb{E}_{R_\tau(\mu)} f = \sum_{j \in J_{\mu,\tau}} \sum_{\alpha \in \mathbb{N}_0^d \colon |\alpha|_1 = k^*} \int_{A_{j,\tau}} \frac{f^{(\alpha)}(\theta_z^{(j)})}{\alpha!} (z - z^{(j)})^\alpha \, (d\mu_j(z) - dR(\mu_j)(z)),$$

where $\theta_z^{(j)} \in \mathbb{R}^d$ satisfies $|z-\theta_z^{(j)}| \le |z-z^{(j)}| \le \tau^{1/2}$ for all $j \in J_{\mu,\tau}$ and $z \in A_{j,\tau}$. Employing the first inequality in (29) we thus conclude that

$$\begin{aligned} &\Big| \sum_{\alpha\in\mathbb{N}_0^d\colon |\alpha|_1=k^*} \int_{A_{j,\tau}} \frac{f^{(\alpha)}(\theta_z^{(j)})}{\alpha!} (z-z^{(j)})^\alpha \, (d\mu_j(z) - dR(\mu_j)(z)) \Big| \\ &\qquad \le c\cdot \|f\|_{k,\beta} \int_{A_{j,\tau}} |z-z^{(j)}|^{k^*} \cdot (1+|z-z^{(j)}|^\beta + |z|^\beta)\, (d\mu_j(z) + dR(\mu_j)(z)) \\ &\qquad \le c\cdot \|f\|_{k,\beta}\cdot \tau^{k^*/2} \cdot \int_{A_{j,\tau}} (1+|z|^\beta)(d\mu_j(z) + dR(\mu_j)(z)) \end{aligned}$$

for all $j \in J_{\mu,\tau}$. Hence

$$|\mathbb{E}_\mu f - \mathbb{E}_{R_\tau(\mu)} f| \le c\cdot \|f\|_{k,\beta}\cdot \tau^{k^*/2}\cdot \int_{\mathbb{R}^d} (1+|z|^\beta)\, (d\mu(z) + dR_\tau(\mu)(z)),$$

which completes the proof of the lemma. □

We provide estimates of absolute moments of $T_\tau(\mu)$, $D_\tau(\mu)$ and $R_\tau(\mu)$ in terms of absolute moments of $\mu \in \mathcal{M}_1(\mathbb{R}^d)$.

Lemma 6. *For all $\tau \in (0,1]$, $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $p \in 2\mathbb{N}$ we have*

(i) $m_p(T_\tau(\mu)) \le m_p(\mu) + c(p)\cdot\tau\cdot(1+m_p(\mu))$,
(ii) $m_p(D_\tau(\mu)) \le m_p(\mu)$,
(iii) $m_p(R_\tau(\mu)) \le m_p(\mu) + c(p)\cdot\tau\cdot(1+m_p(\mu))$.

Proof. Let $\tau \in (0,1]$, $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $p \in 2\mathbb{N}$.

We start with the proof of (i). By definition of T_τ,

$$m_p(T_\tau(\mu)) = \sum_{y\in \mathrm{supp}(\mu)} \mu(\{y\})\cdot \mathbb{E}|Y^y(\tau)|_p^p.$$

Since p is even we have

$$|Y_i^y(\tau)|^p = \sum_{r=0}^p \binom{p}{r} \cdot y_i^{p-r}\cdot (Y_i^y(\tau)-y_i)^r$$

for $i = 1,\dots,d$. Using Lemma 2 we conclude that

$$\begin{aligned} |Y^y(\tau)|_p^p &\le |y|_p^p + p\cdot \sum_{i=1}^d y_i^{p-1}\cdot (Y_i^y(\tau)-y_i) + \sum_{r=2}^p \binom{p}{r} \sum_{i=1}^d |y_i|^{p-r}\cdot |Y_i^y(\tau)-y_i|^r \\ &\le |y|_p^p + p\cdot \sum_{i=1}^d y_i^{p-1}\cdot (Y_i^y(\tau)-y_i) + c\cdot\tau\cdot \sum_{r=2}^p \binom{p}{r}\cdot |y|_p^{p-r}. \end{aligned}$$

Note that $\mathbb{E}(\xi_{(j)}) = 0$ for all $j \in \{1, \dots, m\}$ and observe (27) to obtain

$$|\mathbb{E}(Y_i^y(\tau) - y_i)| = \Big| \sum_{\alpha \in \Gamma_\gamma} \psi_{\alpha,i}(y)\, \tau^{\deg(\alpha)/2}\, \mathbb{E}(\xi_\alpha) \Big| \le c \cdot \tau \cdot \sum_{\alpha \in \Gamma_\gamma} |\mathbb{E}(\xi_\alpha)|$$

for $i = 1, \dots, d$. Hence

$$\mathbb{E}|Y^y(\tau)|_p^p \le |y|_p^p + c \cdot \tau \cdot \sum_{r=1}^{p} \binom{p}{r} \cdot |y|_p^{p-r} \le |y|_p^p + c(p) \cdot \tau \cdot (1 + |y|_p^p),$$

which completes the proof of the estimate (i).

The estimate (ii) is obvious. Let us finally turn to the proof of the estimate (iii). Consider the restriction μ_j of μ to the cube $A_{j,\tau}$ for $j \in J_{\mu,\tau}$ as in the proof of Lemma 5. Clearly,

$$m_p(R_\tau(\mu)) = \sum_{j \in J_{\mu,\tau}} \int_{A_{j,\tau}} |z|_p^p \, dR(\mu_j)(z). \tag{30}$$

Fix $j \in J_{\mu,\tau}$, put

$$\theta = \left(\frac{1}{\mu(A_{j,\tau})} \int_{A_{j,\tau}} z_i \, dR(\mu_j)(z) \right)_{i=1,\dots,d}$$

and note that by properties (10) and (12) of the operator R,

$$\theta = \left(\frac{1}{\mu(A_{j,\tau})} \int_{A_{j,\tau}} z_i \, d\mu_j(z) \right)_{i=1,\dots,d} \in A_{j,\tau}$$

as well as

$$|\theta|_p^p \le \frac{1}{\mu(A_{j,\tau})} \int_{A_{j,\tau}} |z|_p^p \, d\mu_j(z) = \frac{m_p(\mu_j)}{\mu(A_{j,\tau})}.$$

Let $z \in A_{j,\tau}$. Since p is even we have

$$|z_i|^p = \sum_{r=0}^{p} \binom{p}{r} \cdot \theta_i^{p-r} \cdot (z_i - \theta_i)^r$$

for all $i = 1, \dots, d$, and therefore,

$$\begin{aligned}
|z|_p^p &\le |\theta|_p^p + p \cdot \sum_{i=1}^{d} \theta_i^{p-1} \cdot (z_i - \theta_i) + \sum_{r=2}^{p} \binom{p}{r} \sum_{i=1}^{d} |\theta_i|^{p-r} \cdot |z_i - \theta_i|^r \\
&\le |\theta|_p^p + p \cdot \sum_{i=1}^{d} \theta_i^{p-1} \cdot (z_i - \theta_i) + d \cdot \tau \cdot \sum_{r=2}^{p} \binom{p}{r} \cdot |\theta|_p^{p-r} \\
&\le \frac{m_p(\mu_j)}{\mu(A_{j,\tau})} + p \cdot \sum_{i=1}^{d} \theta_i^{p-1} \cdot (z_i - \theta_i) + c(p) \cdot \tau \cdot \Big(1 + \frac{m_p(\mu_j)}{\mu(A_{j,\tau})}\Big).
\end{aligned}$$

Since

$$\int_{A_{j,\tau}} (z_i - \theta_i)\, dR(\mu_j)(z) = 0$$

for $i = 1, \dots, d$ we conclude that

$$\int_{A_{j,\tau}} |z|_p^p \, dR(\mu_j)(z) \leq m_p(\mu_j) + c(p) \cdot \tau \cdot (\mu(A_{j,\tau}) + m_p(\mu_j)),$$

which finishes the proof of part (iii) of the lemma. □

Based on Lemma 6 we can now provide estimates of absolute moments of the probability measures $Q_\ell^R(\delta_{x_0})$, $Q_\ell^T(\delta_{x_0})$ and $Q_\ell^D(\delta_{x_0})$.

Lemma 7. *For all $p \in \mathbb{N}$ we have*

(i) $\max_{\ell=0,\dots,n} m_p(Q_\ell^R(\delta_{x_0})) \leq c(p) \cdot (1 + |x_0|_p^p)$,

(ii) $\max_{\ell=1,\dots,n} m_p(Q_\ell^T(\delta_{x_0})) \leq c(p) \cdot (1 + |x_0|_p^p)$,

(iii) $\max_{\ell=1,\dots,n} m_p(Q_\ell^D(\delta_{x_0})) \leq c(p) \cdot (1 + |x_0|_p^p)$.

Proof. Without loss of generality we may assume $p \in 2\mathbb{N}$. Clearly,

$$m_p(Q_0^R(\delta_{x_0})) = m_p(\delta_{x_0}) = |x_0|_p^p. \tag{31}$$

Moreover, by Lemma 6,

$$m_p(Q_\ell^R(\delta_{x_0})) \leq m_p(Q_{\ell-1}^R(\delta_{x_0})) \cdot (1 + c(p) \cdot \tau_\ell) + c(p) \cdot \tau_\ell \tag{32}$$

and

$$m_p(Q_\ell^D(\delta_{x_0})) \leq m_p(Q_\ell^T(\delta_{x_0})) \leq c(p) \cdot (1 + m_p(Q_{\ell-1}^R(\delta_{x_0}))) \tag{33}$$

for $\ell = 1, \dots, n$.

Using Gronwalls inequality we obtain the estimate in part (i) from (31) and (32). The estimates in part (ii) and (iii) are immediate consequences of (33) together with part (i). □

We are ready to prove part (ii) of Theorem 1. Let $f \in F^{r,\beta}$ and put

$$g_\ell = P_{1-t_\ell} f$$

for $\ell = 0, \dots, n$. Note that $g_n = f$. By Lemma 8 in the appendix we have

$$g_1, \dots, g_{n-1} \in F^{k,\beta}$$

with

$$\|g_\ell\|_{k,\beta} \leq c \cdot \|f\|_{r,\beta} \cdot (1 - t_\ell)^{-(k-\min(k,r))/2} \tag{34}$$

for $k = 1, \dots, 2\gamma + 3$.

We have

$$P_1 f(x_0) - \mathbb{E}_{Q_n^R(\delta_{x_0})} f = \sum_{\ell=1}^{n} E_\ell^T + \sum_{\ell=1}^{n} E_\ell^D + \sum_{\ell=1}^{n} E_\ell^R,$$

where

$$\begin{aligned} E_\ell^T &= \mathbb{E}_{Q_{\ell-1}^R(\delta_{x_0})} g_{\ell-1} - \mathbb{E}_{Q_\ell^T(\delta_{x_0})} g_\ell, \\ E_\ell^D &= \mathbb{E}_{Q_\ell^T(\delta_{x_0})} g_\ell - \mathbb{E}_{Q_\ell^D(\delta_{x_0})} g_\ell, \\ E_\ell^R &= \mathbb{E}_{Q_\ell^D(\delta_{x_0})} g_\ell - \mathbb{E}_{Q_\ell^R(\delta_{x_0})} g_\ell. \end{aligned}$$

We separately estimate the quantities E_ℓ^T, E_ℓ^D and E_ℓ^R. Put

$$k_\ell = \begin{cases} r, & \text{if } \ell = n, \\ 2\gamma + 2, & \text{if } \ell \in \{1, \ldots, n-1\}. \end{cases}$$

By definition of Q_ℓ^T,

$$\begin{aligned} E_\ell^T &= \mathbb{E}_{Q_{\ell-1}^R(\delta_{x_0})} P_{\tau_\ell} g_\ell - \mathbb{E}_{T_{\tau_\ell} \circ Q_{\ell-1}^R(\delta_{x_0})} g_\ell \\ &= \int_{\mathbb{R}^d} \left(P_{\tau_\ell} g_\ell(y) - \mathbb{E}_{T_{\tau_\ell}(\delta_y)} g_\ell \right) Q_{\ell-1}^R(\delta_{x_0})(dy). \end{aligned}$$

Hence, by Lemma 3, Lemma 7(i) and (34),

$$\begin{aligned} |E_\ell^T| &\le c \cdot \|g_\ell\|_{k_\ell,\beta} \cdot \tau_\ell^{\min(\gamma+1, k_\ell/2)} \cdot m_c(Q_{\ell-1}^R(\delta_{x_0})) \\ &\le c \cdot (1 + |x_0|^c) \cdot \|f\|_{r,\beta} \cdot \frac{\tau_\ell^{\min(\gamma+1,k_\ell/2)}}{(1-t_\ell)^{(k_\ell - \min(k_\ell, r))/2}}. \end{aligned} \tag{35}$$

Use Lemma 4 with $p = \gamma + 1$, Lemma 7(ii) and observe (34) to obtain

$$\begin{aligned} |E_\ell^D| &= |\mathbb{E}_{Q_\ell^T(\delta_{x_0})} g_\ell - \mathbb{E}_{D_{\tau_\ell} \circ Q_\ell^T(\delta_{x_0})} g_\ell| \\ &\le c \cdot \|g_\ell\|_{1,\beta} \cdot (1 + m_{1+\beta+(\gamma+1)/\varepsilon}(Q_\ell^T(\delta_{x_0}))) \cdot \tau_\ell^{\gamma+1} \\ &\le c \cdot \|f\|_{r,\beta} \cdot (1 + |x_0|^c) \cdot \tau_\ell^{\gamma+1}. \end{aligned} \tag{36}$$

Finally, by Lemma 5, Lemma 7(i),(iii) and (34),

$$\begin{aligned} |E_\ell^R| &= |\mathbb{E}_{Q_\ell^D(\delta_{x_0})} g_\ell - \mathbb{E}_{R_{\tau_\ell} \circ Q_\ell^D(\delta_{x_0})} g_\ell| \\ &\le c \cdot \|g_\ell\|_{k_\ell,\beta} \cdot (1 + m_\beta(Q_\ell^D(\delta_{x_0})) + m_\beta(Q_\ell^R(\delta_{x_0}))) \cdot \tau_\ell^{\min(\gamma+1,k_\ell/2)} \\ &\le c \cdot (1 + |x_0|^\beta) \cdot \|f\|_{r,\beta} \cdot \frac{\tau_\ell^{\min(\gamma+1,k_\ell/2)}}{(1-t_\ell)^{(k_\ell - \min(k_\ell,r))/2}}. \end{aligned} \tag{37}$$

Combine (35) to (37) to get

$$|P_1 f(x_0) - \mathbb{E}_{Q_n^R(\delta_{x_0})} f| \leq c \cdot \|f\|_{r,\beta} \cdot (1 + |x_0|^c) \cdot \Big(\sum_{\ell=1}^{n-1} \frac{\tau_\ell^{\gamma+1}}{(1-t_\ell)^{\gamma+1-\min(\gamma+1,r/2)}} + \tau_n^{\min(\gamma+1,r/2)} \Big).$$

Note that $\tau_n = 1/n^\eta$. Furthermore, by (23),

$$\frac{\tau_\ell^{\gamma+1}}{(1-t_\ell)^{\gamma+1-\min(\gamma+1,r/2)}} \leq \eta^{\gamma+1} \cdot n^{-\eta \min(\gamma+1,r/2)} \cdot \frac{(n-\ell+1)^{(\gamma+1)(\eta-1)}}{(n-\ell)^{\eta(\gamma+1-\min(\gamma+1,r/2))}}$$

for $\ell \in \{1, \ldots, n-1\}$. Thus

$$\begin{aligned} &\sum_{\ell=1}^{n-1} \frac{\tau_\ell^{\gamma+1}}{(1-t_\ell)^{\gamma+1-\min(\gamma+1,r/2)}} + \tau_n^{\min(\gamma+1,r/2)} \\ &\qquad \leq c \cdot n^{-\eta \min(\gamma+1,r/2)} \cdot \sum_{\ell=1}^{n-1} \frac{1}{\ell^{\gamma+1-\eta\min(\gamma+1,r/2)}} \\ &\qquad \leq c \cdot \begin{cases} n^{-\eta\min(\gamma+1,r/2)}, & \text{if } \eta < \gamma/\min(\gamma+1,r/2), \\ n^{-\eta\min(\gamma+1,r/2)} \ln(n), & \text{if } \eta = \gamma/\min(\gamma+1,r/2), \\ n^{-\gamma}, & \text{if } \eta > \gamma/\min(\gamma+1,r/2), \end{cases} \end{aligned}$$

which completes the proof of part (ii) of Theorem 1.

8.3. **Proof of the cost estimate** (13)**.** In this section unspecified positive constants c may only depend on d, m, γ, K, ξ and η.

We proceed analogously to the proof of Theorem 1(i). Let $\widetilde{Q}_0^R$ denote the identity operator on $\mathcal{M}_1(\mathbb{R}^d)$ and define iteratively

$$\begin{aligned} \widetilde{Q}_\ell^T &= T_{\tau_\ell} \circ \widetilde{Q}_{\ell-1}^R, \\ \widetilde{Q}_\ell^R &= R_{\tau_\ell} \circ \widetilde{Q}_\ell^T \end{aligned}$$

for $\ell = 1, \ldots, n$. Thus $\widetilde{Q}_n(x_0, a, b) = \widetilde{Q}_n^R(\delta_{x_0})$.

For $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and $\tau \in (0, 1]$ we have

$$\operatorname{supp}(T_\tau(\mu)) \subset \big[-\max_{y \in \operatorname{supp}(\mu)} |y| - c \cdot \tau^{1/2}, \max_{y \in \operatorname{supp}(\mu)} |y| + c \cdot \tau^{1/2} \big]^d$$

due to Lemma 2 as well as $\operatorname{supp}(R_\tau(\mu)) \subset \operatorname{supp}(\mu)$ by the definition of R_τ. Hence

$$\operatorname{supp}(\widetilde{Q}_\ell^R(\delta_{x_0})) \subset \operatorname{supp}(\widetilde{Q}_\ell^T(\delta_{x_0})) \subset \Big[-|x_0| - c \cdot \sum_{j=1}^{\ell} \tau_j^{1/2}, \, |x_0| + c \cdot \sum_{j=1}^{\ell} \tau_j^{1/2} \Big]^d$$

for $\ell = 1, \dots, n$, which implies

$$|J_{\widetilde{Q}_\ell^R(\delta_{x_0}),\tau_\ell}| \le c \cdot \Big(\tau_\ell^{-1/2} \cdot \big(1 + \sum_{j=1}^{\ell} \tau_j^{1/2}\big)\Big)^d$$

for $\ell = 1, \dots, n$.

Due to (23) we have

$$\sum_{j=1}^{\ell} \tau_j^{1/2} \le c \cdot n^{-\eta/2} \cdot \sum_{j=1}^{n} (n-j+1)^{(\eta-1)/2} \le c \cdot n^{1/2}.$$

Using the support size estimates from Table 1 we thus conclude that

$$|\operatorname{supp}(\widetilde{Q}_\ell^R(\delta_{x_0}))| \le c \cdot |J_{\widetilde{Q}_\ell^T(\delta_{x_0}),\tau_\ell}| \le c \cdot \tau_\ell^{-d/2} \cdot n^{d/2}$$

for $\ell = 1, \dots, n$.

As in the proof of Theorem 1(i) we therefore get by (23)

$$\begin{aligned} \operatorname{cost}(\widetilde{I}_n, (x_0, a, b, f)) &\le |\operatorname{supp}(\widetilde{Q}_n(x_0, a, b))| + c \cdot \sum_{\ell=1}^{n-1} |\operatorname{supp}(\widetilde{Q}_\ell^R(\delta_{x_0}))| \\ &\le c \cdot n^{d/2} \cdot \sum_{\ell=1}^{n} \tau_\ell^{-d/2} \\ &\le c \cdot n^{(\eta+1)d/2} \cdot \sum_{\ell=1}^{n-1} \ell^{-(\eta-1)d/2} \\ &\le c \cdot n^{(\eta+1)d/2} \cdot \begin{cases} n^{-(\eta+1)d/2+d+1}, & \text{if } \eta < \frac{2}{d}+1, \\ \ln(n), & \text{if } \eta = \frac{2}{d}+1, \\ 1, & \text{if } \eta > \frac{2}{d}+1, \end{cases} \end{aligned}$$

which completes the proof of (13).

9. Appendix

In this section we present the sequential support point elimination procedure R of Davis [2] that constitutes the main ingredient of the support cardinality reduction operator R_τ constructed in Section 4.4, and we discuss its computational cost. Furthermore, we provide results on smoothness properties of the semigroup of the linear operators $P_t \colon F^1 \to F^1$ associated with equations $(x_0, a, b) \in \mathcal{C}_{K,\lambda}^{s,s}$, see (21).

9.1. **A sequential support point elimination procedure.** Fix $k \in \mathbb{N}$ and k functions

$$h_1, \dots, h_k \colon \mathbb{R}^d \to \mathbb{R}.$$

Let $N \in \mathbb{N}$ with $N > k$ and consider a measure

$$\mu = \sum_{i=1}^{N} w_i \cdot \delta_{x_i} \in \mathcal{M}(\mathbb{R}^d)$$

with $|\operatorname{supp}(\mu)| = N$. Choose any $u \in \mathbb{R}^{k+1} \setminus \{0\}$ that satisfies

$$\sum_{i=1}^{k+1} u_i \cdot h_j(x_i) = 0, \qquad j = 1, \dots, k, \tag{38}$$

and $\{i \colon u_i < 0\} \neq \emptyset$, and put

$$\alpha = \min_{i \colon u_i < 0} \Big(-\frac{w_i}{u_i} \Big).$$

Let

$$\widetilde{w}_i = \begin{cases} w_i + \alpha \cdot u_i, & \text{if } i \in \{1, \dots, k+1\}, \\ w_i, & \text{if } i \in \{k+2, \dots, N\}, \end{cases}$$

and define

$$\widetilde{\mu} = \sum_{i=1}^{N} \widetilde{w}_i \cdot \delta_{x_i}.$$

Clearly, $\widetilde{w}_1, \dots, \widetilde{w}_N \geq 0$ and $\widetilde{w}_i = 0$ for some $i \in \{1, \dots, k+1\}$. Hence $\widetilde{\mu} \in \mathcal{M}(\mathbb{R}^d)$ and

$$\operatorname{supp}(\widetilde{\mu}) \subset \operatorname{supp}(\mu), \, |\operatorname{supp}(\widetilde{\mu})| \leq N - 1.$$

Furthermore, by (38),

$$\sum_{i=1}^{N} \widetilde{w}_i \cdot h_j(x_i) = \sum_{i=1}^{N} w_i \cdot h_j(x_i), \qquad j = 1, \dots, k.$$

Iteratively repeating this procedure we obtain a measure $R(\mu) \in \mathcal{M}(\mathbb{R}^d)$ with

$$\operatorname{supp}(R(\mu)) \subset \operatorname{supp}(\mu), \, |\operatorname{supp}(R(\mu))| \leq k$$

and

$$\int_{\mathbb{R}^d} h_j \, d\mu = \int_{\mathbb{R}^d} h_j \, dR(\mu), \qquad j = 1, \dots, k.$$

Clearly, at most $N - k$ steps are needed to obtain $R(\mu)$ and in each step the number of basic computational operations needed to compute the support points and the corresponding weights of the actual measure $\widetilde{\mu}$ is bounded by $c \cdot k^3$. Therefore, the total number $\operatorname{op}(R(\mu))$ of basic computational operations needed to compute $R(\mu)$ satisfies

$$\operatorname{op}(R(\mu)) \leq c \cdot |\operatorname{supp}(\mu)| \cdot k^3.$$

Let $\gamma \in \mathbb{N}$. The particular choice

(39) $$k = \binom{2\gamma+1+d}{d}, \quad \{h_1, \dots, h_k\} = \{p_\alpha \colon \alpha \in \mathbb{N}_0^d, |\alpha|_1 \le 2\gamma+1\}$$

with the monomials

$$p_\alpha(x) = x^\alpha, \quad x \in \mathbb{R}^d,$$

yields the operator R employed in Section 4.4.

Remark 5. In the case of (39) with $d = 1$, the system of linear equations (38) is given by

(40) $$Au = 0,$$

where A is a $(2\gamma+2)\times(2\gamma+3)$ Vandermonde matrix with entries $A_{j,i} = x_i^{j-1}$. Consider the $(2\gamma+3)\times(2\gamma+3)$ Vandermonde matrix $\bar{A} = (x_i^{j-1})_{1\le j,i\le 2\gamma+3}$. Then $\bar{A}$ is invertible and the last column of $\bar{A}^{-1}$ solves the equation (40). Using an explicit representation of $\bar{A}^{-1}$, see, e.g., [15, Section 14(c)], one obtains

$$\ker(A) = \Big\{ \big(z \cdot \prod_{k\neq i}(x_k - x_i)^{-1}\big)_{i=1,\dots,2\gamma+3} \colon z \in \mathbb{R} \Big\}$$

for the null space $\ker(A)$ of A.

In the case of (39) with $d > 1$, explicit formulas for a solution u of equation (38) seem to be unknown up to now, such that suitable numerical methods have to be employed.

9.2. **Smoothness properties of the associated semigroup.** Let $s, r \in \mathbb{N}$, $\beta \ge 0$ and $K \ge 1 \ge \lambda > 0$. We fix an equation

(41) $$(x_0, a, b) \in \mathcal{C}^{s,s}_{K,\lambda}$$

and we consider the semigroup of linear operators $(P_t)_{t\in[0,\infty)}$ associated with a and b, see (21).

Lemma 8. *For every $f \in F^{r,\beta}$ and $t \in (0,1]$ we have*

$$P_t f \in C^{s+1}(\mathbb{R}^d)$$

and for all $y \in \mathbb{R}^d$ and all $\alpha \in \mathbb{N}_0^d$ with $0 \le |\alpha|_1 \le s+1$,

(42) $$|(P_t f)^{(\alpha)}(y)| \le c \cdot \|f\|_{r,\beta} \cdot (1 + |y|^\beta) \cdot \frac{1}{t^{(|\alpha|_1 - \min(|\alpha|_1, r))/2}},$$

where $c = c(d, m, s, r, \beta, K, \lambda)$.

Proof. In the sequel unspecified positive constants c may only depend on the parameters d, m, s, r, β, K and λ.

By (41) the distribution of $X^y(t)$ has a Lebesgue-density for all $y \in \mathbb{R}^d$ and $t \in (0,1]$. More precisely, put $\sigma = bb^T$, let

$$V = \{(v, y, t, z) \mid 0 \le v < t \le 1,\ y, z \in \mathbb{R}^d\}$$

and consider the partial differential equation

$$\frac{1}{2}\sum_{i,j=1}^{d} \sigma_{i,j}(y) \cdot \frac{\partial^2 u}{\partial y_i \partial y_j}(v,y) + \sum_{i=1}^{d} a_i(y) \cdot \frac{\partial u}{\partial y_i}(v,y) + \frac{\partial u}{\partial v}(v,y) = 0 \tag{43}$$

for $(v,y) \in [0,1) \times \mathbb{R}^d$. By (41) the equation (43) has a unique fundamental solution

$$G\colon V \to [0,\infty)$$

and we have

$$\mathbb{P}_{X^y(t)}(A) = \int_A G(0,y,t,z)\,dz \tag{44}$$

for every Borel set $A \subset \mathbb{R}^d$ and all $y \in \mathbb{R}^d$ and $t \in (0,1]$. See [5, Theorem 6.5.4].

Moreover, by [4, Chapter 9, Theorem 7] we have $G(0,\cdot,t,z) \in C^{s+1}(\mathbb{R}^d)$ for all $t \in (0,1]$ and $z \in \mathbb{R}^d$ with

$$\Big|\frac{\partial^\alpha G}{\partial y^\alpha}(0,y,t,z)\Big| \le \frac{c}{t^{(|\alpha|_1+d)/2}} \cdot \exp\Big(-c \cdot \frac{|z-y|^2}{t}\Big) \tag{45}$$

for all $\alpha \in \mathbb{N}_0^d$ with $0 \le |\alpha|_1 \le s+1$ and $y \in \mathbb{R}^d$.

Using (44) and (45) we obtain $P_t f \in C^{s+1}(\mathbb{R}^d)$ with

$$(P_t f)^{(\alpha)}(y) = \int_{\mathbb{R}^d} \frac{\partial^\alpha G}{\partial y^\alpha}(0,y,t,z) \cdot f(z)\,dz \tag{46}$$

for all $f \in F^{r,\beta}$, $t \in (0,1]$, $y \in \mathbb{R}^d$ and $\alpha \in \mathbb{N}_0^d$ with $0 \le |\alpha|_1 \le s+1$.

Fix f, t, y of the latter type and $\alpha \in \mathbb{N}_0^d$ with $1 \le |\alpha|_1 \le s+1$. Put

$$q = \min(|\alpha|_1, r)$$

as well as

$$I = \int_{\mathbb{R}^d} \Big(f(z) - \sum_{\rho \in \mathbb{N}_0^d\colon 0 \le |\rho|_1 \le q-1} \frac{f^{(\rho)}(y)}{\rho!} \cdot (z-y)^\rho\Big) \cdot \frac{\partial^\alpha G}{\partial y^\alpha}(0,y,t,z)\,dz$$

and

$$I_\rho = \frac{f^{(\rho)}(y)}{\rho!} \cdot \int_{\mathbb{R}^d} (z-y)^\rho \cdot \frac{\partial^\alpha G}{\partial y^\alpha}(0,y,t,z)\,dz$$

for $\rho \in \mathbb{N}_0^d$ with $0 \le |\rho|_1 \le q-1$.

Using integration by parts as well as (45) and (46) we obtain

$$I_\rho = (-1)^{|\rho|_1} \cdot f^{(\rho)}(y) \cdot \int_{\mathbb{R}^d} \frac{\partial^{\alpha-\rho} G}{\partial y^{\alpha-\rho}}(0,y,t,z)\,dz = (-1)^{|\rho|_1} \cdot f^{(\rho)}(y) \cdot (P_t 1)^{(\alpha-\rho)}(y) = 0$$

for all $\rho \in \mathbb{N}_0^d$ with $0 \le |\rho|_1 \le q-1$. Hence

$$(P_t f)^{(\alpha)}(y) = I.$$

Clearly,

$$I = \sum_{\rho \in \mathbb{N}_0^d\colon |\rho|_1 = q} \int_{\mathbb{R}^d} f^{(\rho)}(\theta_z) \cdot (z-y)^q \cdot \frac{\partial^\alpha G}{\partial y^\alpha}(0, y, t, z)\, dz,$$

where $\theta_z \in \mathbb{R}^d$ satisfies $|\theta_z - y| \le |z-y|$. Since $f \in F^{r,\beta}$ we get, similarly to (29),

$$|f^{(\rho)}(\theta_z)| \le c \cdot \|f\|_{r,\beta} \cdot (1 + |y|^\beta + |z-y|^\beta)$$

for all $\rho \in \mathbb{N}_0^d$ with $|\rho|_1 = q$ and all $z \in \mathbb{R}^d$. Using (45) we therefore obtain

$$\begin{aligned} |I| &\le \frac{c}{t^{(|\alpha|_1+d)/2}} \cdot \|f\|_{r,\beta} \cdot \int_{\mathbb{R}^d} (1 + |y|^\beta + |z-y|^\beta) \cdot |z-y|^q \cdot \exp\Big(-c \cdot \frac{|z-y|^2}{t}\Big) dz \\ &\le \frac{c}{t^{(|\alpha|_1+d-q)/2}} \cdot \|f\|_{r,\beta} \cdot \int_{\mathbb{R}^d} (1 + |y|^\beta + |z-y|^\beta) \cdot \exp\Big(-c \cdot \frac{|z-y|^2}{t}\Big) dz \\ &\le \frac{c}{t^{(|\alpha|_1-q)/2}} \cdot \|f\|_{r,\beta} \cdot (1 + |y|^\beta), \end{aligned}$$

which completes the proof of (42) in the case $1 \le |\alpha|_1 \le s+1$. The case $\alpha = 0$ is treated analogously. □

Acknowledgement

We are thankful to Fabian Knorr for carrying out the numerical experiments presented in Section 7.

References

[1] Bayer, C., Teichmann, J. (2006), The proof of Tchakaloff's theorem, Proc. Amer. Math. Soc. **134**, 3035–3040.

[2] Davis, P. (1967), A construction of nonnegative approximate quadratures, Math. Comp. **21**, 578–582.

[3] Dereich, S., Müller-Gronbach, T. Ritter, K. (2013), On the complexity of computing quadrature formulas for SDEs, in: Foundations of computational mathematics, Budapest 2011, London Math. Soc. Lecture Note Ser., Vol. 403, pp. 72–92, Cambridge Univ. Press, Cambridge.

[4] Friedman, A. (1964), *Partial differential equations of parabolic type*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

[5] Friedman, A. (1975), *Stochastic Differential Equatios and Applications. Vol. 1.*, Academic Press, New York – London.

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

[7] Geiss, S. (2002), Quantitative approximation of certain stochastic integrals, Stoch. Stoch. Rep. **73**, 241–270.

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

[9] Kloeden, P. E., Platen, E. (1999), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

[10] Kusuoka, S. (2001), Approximation of expectation of diffusion process and mathematical finance, Adv. Stud. Pure Math. **31**, 147–165.

[11] Kusuoka, S. (2004), Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus, Adv. Math. Econ. **6**, 69–83.
[12] Litterer, C., Lyons, T. (2012), High order recombination and an application to cubature on wiener space, Ann. Appl. Probab.**22**, 1301–1327.
[13] Lyons, T., Gyurkó, L.G. (2011), Efficient and practical implementations of cubature on Wiener space, Stochastic Analysis **2010**, 73–111.
[14] Lyons, T., Victoir, N. (2004), Cubature on Wiener space, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **460**, 169–198.
[15] Miller, K. S.(1987), *Some eclectic matrix theory*, Robert E. Krieger Publishing Co., Inc., Melbourne.
[16] Müller-Gronbach, T., Ritter, K., Yaroslavtseva, L. (2012), Derandomization of the Euler scheme for scalar stochastic differential equations, J. Complexity **28**, 139–153.
[17] Müller-Gronbach, T., Ritter, K., Yaroslavtseva, L. (2014), On the complexity of computing quadrature formulas for marginal distributions of SDEs, DFG SPP 1324 Preprint 157.
[18] Novak, E. (1995), The real number model in numerical analysis, J. Complexity **11**, 57–73.
[19] Talay, D. (1984), Efficient numerical schemes for the approximation of expectations of functionals of the solution of an SDE and applications, in: Filtering and control of Random Processes, Lect. Notes Control Inf. Sci., Vol. 61, pp. 294–313, Springer, Berlin.

Fakultät für Informatik und Mathematik, Universität Passau, Innstr. 33, 94032 Passau, Germany
E-mail address: thomas.mueller-gronbach@uni-passau.de

Fakultät für Informatik und Mathematik, Universität Passau, Innstr. 33, 94032 Passau, Germany
E-mail address: larisa.yaroslavtseva@uni-passau.de

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak and H. Woźniakowski. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.

[69] G. Kutyniok and W.-Q Lim. Image Separation Using Shearlets. Preprint 69, DFG-SPP 1324, November 2010.

[70] B. Jin and P. Maass. An Analysis of Electrical Impedance Tomography with Applications to Tikhonov Regularization. Preprint 70, DFG-SPP 1324, December 2010.

[71] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT Format. Preprint 71, DFG-SPP 1324, December 2010.

[72] T. Müller-Gronbach and K. Ritter. A Local Refinement Strategy for Constructive Quantization of Scalar SDEs. Preprint 72, DFG-SPP 1324, December 2010.

[73] T. Rohwedder and R. Schneider. An Analysis for the DIIS Acceleration Method used in Quantum Chemistry Calculations. Preprint 73, DFG-SPP 1324, December 2010.

[74] C. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. Preprint 74, DFG-SPP 1324, December 2010.

[75] C. Bender. Primal and Dual Pricing of Multiple Exercise Options in Continuous Time. Preprint 75, DFG-SPP 1324, December 2010.

[76] H. Harbrecht, M. Peters, and R. Schneider. On the Low-rank Approximation by the Pivoted Cholesky Decomposition. Preprint 76, DFG-SPP 1324, December 2010.

[77] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Adaptive Wavelet Methods for Elliptic Stochastic Partial Differential Equations. Preprint 77, DFG-SPP 1324, January 2011.

[78] G. Plonka, S. Tenorth, and A. Iske. Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform. Preprint 78, DFG-SPP 1324, January 2011.

[79] A. Mugler and H.-J. Starkloff. On Elliptic Partial Differential Equations with Random Coefficients. Preprint 79, DFG-SPP 1324, January 2011.

[80] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. A Derandomization of the Euler Scheme for Scalar Stochastic Differential Equations. Preprint 80, DFG-SPP 1324, January 2011.

[81] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. Preprint 81, DFG-SPP 1324, January 2011.

[82] K. Grella and C. Schwab. Sparse Tensor Spherical Harmonics Approximation in Radiative Transfer. Preprint 82, DFG-SPP 1324, January 2011.

[83] D.A. Lorenz, S. Schiffler, and D. Trede. Beyond Convergence Rates: Exact Inversion With Tikhonov Regularization With Sparsity Constraints. Preprint 83, DFG-SPP 1324, January 2011.

[84] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. Preprint 84, DFG-SPP 1324, January 2011.

[85] S. Dahlke and W. Sickel. On Besov Regularity of Solutions to Nonlinear Elliptic Partial Differential Equations. Preprint 85, DFG-SPP 1324, January 2011.

[86] S. Dahlke, U. Friedrich, P. Maass, T. Raasch, and R.A. Ressel. An adaptive wavelet method for parameter identification problems in parabolic partial differential equations. Preprint 86, DFG-SPP 1324, January 2011.

[87] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and Variational Stabilization for Convection-Diffusion Equations. Preprint 87, DFG-SPP 1324, January 2011.

[88] T. Jahnke. On Reduced Models for the Chemical Master Equation. Preprint 88, DFG-SPP 1324, January 2011.

[89] P. Binev, W. Dahmen, R. DeVore, P. Lamby, D. Savu, and R. Sharpley. Compressed Sensing and Electron Microscopy. Preprint 89, DFG-SPP 1324, March 2011.

[90] P. Binev, F. Blanco-Silva, D. Blom, W. Dahmen, P. Lamby, R. Sharpley, and T. Vogt. High Quality Image Formation by Nonlocal Means Applied to High-Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). Preprint 90, DFG-SPP 1324, March 2011.

[91] R. A. Ressel. A Parameter Identification Problem for a Nonlinear Parabolic Differential Equation. Preprint 91, DFG-SPP 1324, May 2011.

[92] G. Kutyniok. Data Separation by Sparse Representations. Preprint 92, DFG-SPP 1324, May 2011.

[93] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to Compressed Sensing. Preprint 93, DFG-SPP 1324, May 2011.

[94] H.-C. Kreusler and H. Yserentant. The Mixed Regularity of Electronic Wave Functions in Fractional Order and Weighted Sobolev Spaces. Preprint 94, DFG-SPP 1324, June 2011.

[95] E. Ullmann, H. C. Elman, and O. G. Ernst. Efficient Iterative Solvers for Stochastic Galerkin Discretizations of Log-Transformed Random Diffusion Problems. Preprint 95, DFG-SPP 1324, June 2011.

[96] S. Kunis and I. Melzer. On the Butterfly Sparse Fourier Transform. Preprint 96, DFG-SPP 1324, June 2011.

[97] T. Rohwedder. The Continuous Coupled Cluster Formulation for the Electronic Schrödinger Equation. Preprint 97, DFG-SPP 1324, June 2011.

[98] T. Rohwedder and R. Schneider. Error Estimates for the Coupled Cluster Method. Preprint 98, DFG-SPP 1324, June 2011.

[99] P. A. Cioica and S. Dahlke. Spatial Besov Regularity for Semilinear Stochastic Partial Differential Equations on Bounded Lipschitz Domains. Preprint 99, DFG-SPP 1324, July 2011.

[100] L. Grasedyck and W. Hackbusch. An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples. Preprint 100, DFG-SPP 1324, August 2011.

[101] N. Chegini, S. Dahlke, U. Friedrich, and R. Stevenson. Piecewise Tensor Product Wavelet Bases by Extensions and Approximation Rates. Preprint 101, DFG-SPP 1324, September 2011.

[102] S. Dahlke, P. Oswald, and T. Raasch. A Note on Quarkonial Systems and Multilevel Partition of Unity Methods. Preprint 102, DFG-SPP 1324, September 2011.

[103] A. Uschmajew. Local Convergence of the Alternating Least Squares Algorithm For Canonical Tensor Approximation. Preprint 103, DFG-SPP 1324, September 2011.

[104] S. Kvaal. Multiconfigurational time-dependent Hartree method for describing particle loss due to absorbing boundary conditions. Preprint 104, DFG-SPP 1324, September 2011.

[105] M. Guillemard and A. Iske. On Groupoid C*-Algebras, Persistent Homology and Time-Frequency Analysis. Preprint 105, DFG-SPP 1324, September 2011.

[106] A. Hinrichs, E. Novak, and H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Preprint 106, DFG-SPP 1324, September 2011.

[107] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander. Efficient Analysis of High Dimensional Data in Tensor Formats. Preprint 107, DFG-SPP 1324, September 2011.

[108] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider. Optimization Problems in Contracted Tensor Networks. Preprint 108, DFG-SPP 1324, October 2011.

[109] S. Dereich, T. Müller-Gronbach, and K. Ritter. On the Complexity of Computing Quadrature Formulas for SDEs. Preprint 109, DFG-SPP 1324, October 2011.

[110] D. Belomestny. Solving optimal stopping problems by empirical dual optimization and penalization. Preprint 110, DFG-SPP 1324, November 2011.

[111] D. Belomestny and J. Schoenmakers. Multilevel dual approach for pricing American style derivatives. Preprint 111, DFG-SPP 1324, November 2011.

[112] T. Rohwedder and A. Uschmajew. Local convergence of alternating schemes for optimization of convex problems in the TT format. Preprint 112, DFG-SPP 1324, December 2011.

[113] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. Preprint 113, DFG-SPP 1324, December 2011.

[114] Y. Dong, T. Görner, and S. Kunis. An iterative reconstruction scheme for photoacoustic imaging. Preprint 114, DFG-SPP 1324, December 2011.

[115] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along generated sets. Preprint 115, DFG-SPP 1324, February 2012.

[116] H. Chen and R. Schneider. Numerical analysis of augmented plane waves methods for full-potential electronic structure calculations. Preprint 116, DFG-SPP 1324, February 2012.

[117] J. Ma, G. Plonka, and M.Y. Hussaini. Compressive Video Sampling with Approximate Message Passing Decoding. Preprint 117, DFG-SPP 1324, February 2012.

[118] D. Heinen and G. Plonka. Wavelet shrinkage on paths for scattered data denoising. Preprint 118, DFG-SPP 1324, February 2012.

[119] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. Preprint 119, DFG-SPP 1324, March 2012.

[120] C. Bender and J. Steiner. A-posteriori estimates for backward SDEs. Preprint 120, DFG-SPP 1324, April 2012.

[121] M. Espig, W. Hackbusch, A. Litvinenkoy, H.G. Matthiesy, and P. Wähnert. Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Preprint 121, DFG-SPP 1324, May 2012.

[122] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. Preprint 122, DFG-SPP 1324, June 2012.

[123] A. Mugler and H.-J. Starkloff. On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. Preprint 123, DFG-SPP 1324, June 2012.

[124] P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R.L. Schilling. On the convergence analysis of Rothe's method. Preprint 124, DFG-SPP 1324, July 2012.

[125] P. Binev, A. Cohen, W. Dahmen, and R. DeVore. Classification Algorithms using Adaptive Partitioning. Preprint 125, DFG-SPP 1324, July 2012.

[126] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation of hierarchical Tucker and Tensor-Train tensors. Preprint 126, DFG-SPP 1324, July 2012.

[127] M. Kovács, S. Larsson, and K. Urban. On Wavelet-Galerkin methods for semilinear parabolic equations with additive noise. Preprint 127, DFG-SPP 1324, August 2012.

[128] M. Bachmayr, H. Chen, and R. Schneider. Numerical analysis of Gaussian approximations in quantum chemistry. Preprint 128, DFG-SPP 1324, August 2012.

[129] D. Rudolf. Explicit error bounds for Markov chain Monte Carlo. Preprint 129, DFG-SPP 1324, August 2012.

[130] P.A. Cioica, K.-H. Kim, K. Lee, and F. Lindner. On the $L_q(L_p)$-regularity and Besov smoothness of stochastic parabolic equations on bounded Lipschitz domains. Preprint 130, DFG-SPP 1324, December 2012.

[131] M. Hansen. $n-$term Approximation Rates and Besov Regularity for Elliptic PDEs on Polyhedral Domains. Preprint 131, DFG-SPP 1324, December 2012.

[132] R. E. Bank and H. Yserentant. On the H^1 -stability of the L_2 -projection onto finite element spaces. Preprint 132, DFG-SPP 1324, December 2012.

[133] M. Gnewuch, S. Mayer, and K. Ritter. On Weighted Hilbert Spaces and Integration of Functions of Infinitely Many Variables. Preprint 133, DFG-SPP 1324, December 2012.

[134] D. Crisan, J. Diehl, P.K. Friz, and H. Oberhauser. Robust Filtering: Correlated Noise and Multidimensional Observation. Preprint 134, DFG-SPP 1324, January 2013.

[135] Wolfgang Dahmen, Christian Plesken, and Gerrit Welper. Double Greedy Algorithms: Reduced Basis Methods for Transport Dominated Problems. Preprint 135, DFG-SPP 1324, February 2013.

[136] Aicke Hinrichs, Erich Novak, Mario Ullrich, and Henryk Wozniakowski. The Curse of Dimensionality for Numerical Integration of Smooth Functions. Preprint 136, DFG-SPP 1324, February 2013.

[137] Markus Bachmayr, Wolfgang Dahmen, Ronald DeVore, and Lars Grasedyck. Approximation of High-Dimensional Rank One Tensors. Preprint 137, DFG-SPP 1324, March 2013.

[138] Markus Bachmayr and Wolfgang Dahmen. Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations. Preprint 138, DFG-SPP 1324, April 2013.

[139] Felix Lindner. Singular Behavior of the Solution to the Stochastic Heat Equation on a Polygonal Domain. Preprint 139, DFG-SPP 1324, May 2013.

[140] Stephan Dahlke, Dominik Lellek, Shiu Hong Lui, and Rob Stevenson. Adaptive Wavelet Schwarz Methods for the Navier-Stokes Equation. Preprint 140, DFG-SPP 1324, May 2013.

[141] Jonas Ballani and Lars Grasedyck. Tree Adaptive Approximation in the Hierarchical Tensor Format. Preprint 141, DFG-SPP 1324, June 2013.

[142] Harry Yserentant. A short theory of the Rayleigh-Ritz method. Preprint 142, DFG-SPP 1324, July 2013.

[143] M. Hefter and K. Ritter. On Embeddings of Weighted Tensor Product Hilbert Spaces. Preprint 143, DFG-SPP 1324, August 2013.

[144] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts. Preprint 144, DFG-SPP 1324, August 2013.

[145] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate functions by trigonometric polynomials based on rank-1 lattice sampling. Preprint 145, DFG-SPP 1324, September 2013.

[146] C. Bender, N. Schweizer, and J. Zhuo. A primal-dual algorithm for BSDEs. Preprint 146, DFG-SPP 1324, October 2013.

[147] D. Rudolf. Hit-and-run for numerical integration. Preprint 147, DFG-SPP 1324, October 2013.

[148] D. Rudolf and M. Ullrich. Positivity of hit-and-run and related algorithms. Preprint 148, DFG-SPP 1324, October 2013.

[149] L. Grasedyck, M. Kluge, and S. Krämer. Alternating Directions Fitting (ADF) of Hierarchical Low Rank Tensors. Preprint 149, DFG-SPP 1324, October 2013.

[150] F. Filbir, S. Kunis, and R. Seyfried. Effective discretization of direct reconstruction schemes for photoacoustic imaging in spherical geometries. Preprint 150, DFG-SPP 1324, November 2013.

[151] E. Novak, M. Ullrich, and H. Woźniakowski. Complexity of Oscillatory Integration for Univariate Sobolev Spaces. Preprint 151, DFG-SPP 1324, November 2013.

[152] A. Hinrichs, E. Novak, and M. Ullrich. A Tractability Result for the Clenshaw Curtis Smolyak Algorithm. Preprint 152, DFG-SPP 1324, November 2013.

[153] M. Hein, S. Setzer, L. Jost, and S. Rangapuram. The Total Variation on Hypergraphs - Learning on Hypergraphs Revisited. Preprint 153, DFG-SPP 1324, November 2013.

[154] M. Kovács, S. Larsson, and F. Lindgren. On the Backward Euler Approximation of the Stochastic Allen-Chan Equation. Preprint 154, DFG-SPP 1324, November 2013.

[155] S. Dahlke, M. Fornasier, U. Friedrich, and T. Raasch. Multilevel preconditioning for sparse optimization of functionals with nonconvex fidelity terms. Preprint 155, DFG-SPP 1324, December 2013.

[156] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. On the complexity of computing quadrature formulas for marginal distributions of SDEs. Preprint 156, DFG-SPP 1324, January 2014.

[157] M. Giles, T. Nagapetyan, and K. Ritter. Multi-Level Monte Carlo Approximation of Distribution Functions and Densities. Preprint 157, DFG-SPP 1324, February 2014.

[158] F. Dickmann and N. Schweizer. Faster Comparison of Stopping Times by Nested Conditional Monte Carlo . Preprint 158, DFG-SPP 1324, February 2014.

[159] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on sampling along rank-1 lattice with generating vector of Korobov form . Preprint 159, DFG-SPP 1324, February 2014.

[160] S. Dereich and T. Müller-Gronbach. Quadrature for Self-Affine Distributions on $\mathbb{R}^d$. Preprint 160, DFG-SPP 1324, March 2014.

[161] S. Dereich and S. Li. Multilevel Monte Carlo for Lévy-driven SDEs: central limit theorems for adaptive Euler schemes. Preprint 161, DFG-SPP 1324, March 2014.

[162] M. Kluge. Sampling Rules for Tensor Reconstruction in Hierarchical Tucker Format. Preprint 162, DFG-SPP 1324, April 2014.

[163] D. Rudolf and N. Schweizer. Error Bounds of MCMC for Functions with Unbounded Stationary Variance. Preprint 163, DFG-SPP 1324, April 2014.

[164] E. Novak and D. Rudolf. Tractability of the approximation of high-dimensional rank one tensors. Preprint 164, DFG-SPP 1324, April 2014.

[165] J. Dick and D. Rudolf. Discrepancy estimates for variance bounding Markov chain quasi-Monte Carlo. Preprint 165, DFG-SPP 1324, April 2014.

[166] L. Kämmerer, S. Kunis, I. Melzer, D. Potts, and T. Volkmer. Computational Methods for the Fourier Analysis of Sparse High-Dimensional Functions. Preprint 166, DFG-SPP 1324, April 2014.

[167] T. Müller-Gronbach and L. Yaroslavtseva. Deterministic quadrature formulas for SDEs based on simplified weak Ito-Taylor steps. Preprint 167, DFG-SPP 1324, June 2014.