DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

On the complexity of computing quadrature formulas for marginal distributions of SDEs

T. Müller-Gronbach, K. Ritter, L. Yaroslavtseva

Preprint 156

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"

On the complexity of computing quadrature formulas for marginal distributions of SDEs

T. Müller-Gronbach, K. Ritter, L. Yaroslavtseva

Preprint 156

ON THE COMPLEXITY OF COMPUTING QUADRATURE FORMULAS FOR MARGINAL DISTRIBUTIONS OF SDES

THOMAS MÜLLER-GRONBACH, KLAUS RITTER, AND LARISA YAROSLAVTSEVA

Abstract. We study the problem of approximating the distribution of the solution of a d-dimensional system of stochastic differential equations (SDE) at a single time point by a probability measure with finite support, i.e., by a quadrature formula with positive weights summing up to one. We consider deterministic algorithms that may use finitely many evaluations of the drift coefficient and the diffusion coefficient and we analyze their worst case behavior with respect to classes of SDEs, which are specified in terms of smoothness constraints for the coefficients of the equation. The worst case error of an algorithm is defined in terms of a metric on the space of probability measures on the state space of the solution, which is given in a dual representation in terms of a class of test functions on the state space. For the definition of the worst case cost of an algorithm we either consider the size of the support of the approximating probability measure or the number of evaluations of the coefficients of the equation or the total computational cost, i.e., the total number of operations that are carried out by the algorithm to obtain the output. We show that the order of convergence of the corresponding minimal errors is r/d, $\min(s_1, s_2)/d$ and $\min(r, s_1, s_2)/d$, respectively, up to an arbitrarily small power, where the parameters r, s_1 and s_2 denote the smoothness of the test functions, the drift coefficients and the diffusion coefficients, respectively.

1. Introduction

Let $d \in \mathbb{N}$ and consider a d-dimensional system of autonomous SDEs

$$
\begin{aligned}
dX(t) &= a(X(t))\, dt + b(X(t))\, dW(t), \quad t \in [0, 1], \\
X(0) &= x_0,
\end{aligned}
\tag{1.1}
$$

with a d-dimensional driving Brownian motion W and

$$(x_0, a, b) \in \mathcal{C} = \mathcal{C}_0 \times \mathcal{C}_1 \times \mathcal{C}_2,$$

where $\mathcal{C}_0 \subset \mathbb{R}^d$ and $\mathcal{C}_1$ and $\mathcal{C}_2$ are classes of functions $a\colon \mathbb{R}^d \to \mathbb{R}^d$ and $b\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$, respectively, that are at least Lipschitz continuous. In particular, for every $(x_0, a, b) \in \mathcal{C}$ there exists a unique strong solution X of (1.1), and we have $\mathbb{E}\|X(1)\|^p < \infty$ for every $p \geq 1$.

Let $M(\mathbb{R}^d)$ denote the space of Borel probability measures μ on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} \|x\|^p d\mu(x) < \infty$$

for every $p \geq 1$. Our computational task is to approximate the mapping

$$S \colon \mathcal{C} \to M(\mathbb{R}^d), \quad S(x_0, a, b) = \mathbb{P}_{X(1)},$$

by means of deterministic algorithms $\widehat{S}$ that are based on finitely many evaluations of the coefficients a and b and yield measures in the set

$$M_0(\mathbb{R}^d) = \{\mu \in M(\mathbb{R}^d) \mid |\operatorname{supp}(\mu)| < \infty\}$$

of Borel probability measures on $\mathbb{R}^d$ with finite support. We present a worst case analysis for the error and the cost of any such algorithm, and we aim at algorithms with an optimal relation between error and cost.

The constructive approximation of marginal distributions of SDEs has been studied in [8, 9, 11, 12, 13, 14] with a focus on upper bounds, and we refer to [15] for the analogous problem on the path space. A technique to prove lower bounds has been developed in a particular setting in [21]. The complexity of Feynman-Kac path integration, which is related to the problem studied in this paper, has been analyzed in [10, 22].

The error of any algorithm $\widehat{S}$ for a particular SDE is defined in terms of a metric on $M(\mathbb{R}^d)$, which is induced by a class $\mathcal{F}$ of test functions $f \colon \mathbb{R}^d \to \mathbb{R}$. Basically, $\mathcal{F}$ is characterized by a smoothness parameter $r \in \mathbb{N}$, as it consists of functions f with continuous and polynomially bounded partial derivatives of order up to r. Accordingly, the error of $\widehat{S}$ for the input (x_0, a, b), i.e., for the SDE with initial value x_0, drift coefficient a and diffusion coefficient b, is defined by

$$\rho_{\mathcal{F}}(S(x_0, a, b), \widehat{S}(x_0, a, b)) = \sup_{f \in \mathcal{F}} \left| \int_{\mathbb{R}^d} f \, dS(x_0, a, b) - \int_{\mathbb{R}^d} f \, d\widehat{S}(x_0, a, b) \right|.$$

We also impose smoothness assumptions on a and b. Basically, $\mathcal{C}_1$ and $\mathcal{C}_2$ consist of functions with components having continuous and bounded partial derivatives of order up to $s_1 \in \mathbb{N}$ and $s_2 \in \mathbb{N}$, respectively. Furthermore, $\mathcal{C}_0$ is a bounded subset of $\mathbb{R}^d$. This leads to the definition

$$e(\widehat{S}) = \sup_{(x_0, a, b) \in \mathcal{C}} \rho_{\mathcal{F}}(S(x_0, a, b), \widehat{S}(x_0, a, b))$$

of the worst case error of $\widehat{S}$.

The total cost, $\operatorname{cost}^{\text{total}}(\widehat{S}, (x_0, a, b))$, for computing $\widehat{S}(x_0, a, b)$ consists of

(i) the number of evaluations of a and b and
(ii) the number of arithmetical operations, jumps, etc.,

that are carried out by $\widehat{S}$ for the input (x_0, a, b) plus

(iii) the size of the output, i.e., the size of the support of $\widehat{S}(x_0, a, b)$.

The worst case cost of $\widehat{S}$ is defined by

$$\mathrm{cost}^{\mathrm{total}}(\widehat{S}) = \sup_{(x_0,a,b)\in\mathcal{C}} \mathrm{cost}^{\mathrm{total}}(\widehat{S},(x_0,a,b)).$$

As the key quantity we study the N-th minimal error

$$e_N^{\mathrm{total}}(\mathcal{C},\mathcal{F}) = \inf\{e(\widehat{S}) \mid \mathrm{cost}^{\mathrm{total}}(\widehat{S}) \leq N\}, \quad N \in \mathbb{N},$$

which depends on the smoothness parameters r, s_1, s_2 via $\mathcal{F}$ and $\mathcal{C}$ as well as on the dimension d. In particular, we show that

$$N^{-\eta^{\mathrm{total}}} \preccurlyeq e_N^{\mathrm{total}}(\mathcal{C},\mathcal{F}) \preccurlyeq N^{-\eta^{\mathrm{total}}+\varepsilon}$$

for every $\varepsilon > 0$ with

$$\eta^{\mathrm{total}} = \min(r, s_1, s_2)/d.$$

Thus, up to ε, the order of convergence of the minimal error is the ratio of the minimal smoothness $\min(r, s_1, s_2)$ and the dimension d. Results of this type are common for many linear problems with isotropic smoothness conditions, see [17, 18, 19, 20, 25] for an overview and further references, but the present problem is a non-linear one.

To understand the individual impact of the three smoothness parameters, we also analyze minimal errors that merely take into account the information cost according to (i) or the support size, see (iii). The respective worst case quantities are denoted by $\mathrm{cost}^*(\widehat{S})$ with $* \in \{\mathrm{inf}, \mathrm{supp}\}$. The N-th minimal errors

$$e_N^*(\mathcal{C},\mathcal{F}) = \inf\{e(\widehat{S}) \mid \mathrm{cost}^*(\widehat{S}) \leq N\}, \quad N \in \mathbb{N},$$

satisfy

$$N^{-\eta^*} \preccurlyeq e_N^*(\mathcal{C},\mathcal{F}) \preccurlyeq N^{-\eta^*+\varepsilon}$$

for every $\varepsilon > 0$, where

$$\eta^* = \begin{cases} \min(s_1, s_2)/d, & \text{if } * = \mathrm{inf}, \\ r/d, & \text{if } * = \mathrm{supp}. \end{cases}$$

We sketch how to derive the upper bounds for the minimal errors, and we start with the analysis of $e_N^*(\mathcal{C},\mathcal{F})$ with $* \in \{\mathrm{inf}, \mathrm{supp}\}$. To simplify the exposition we assume for the rest of the introduction that all derivatives up to order r of the test functions $f \in \mathcal{F}$ satisfy a linear growth condition. We establish the upper bound for the minimal error in both cases $* = \mathrm{inf}$ and $* = \mathrm{supp}$ by means of a single sequence of algorithms in the following way. The drift coefficient a and the diffusion coefficient b are evaluated on a uniform grid with mesh-size $1/n$ that covers $[-n^\varepsilon, n^\varepsilon]^d$, which causes an information cost of order $d(1+\varepsilon)$. This information allows to compute approximations $\widetilde{a}$ and $\widetilde{b}$ to a and b such that the solutions $\widetilde{X}$ and X of the corresponding SDEs satisfy

$$(\mathbb{E}\|X(1) - \widetilde{X}(1)\|^2)^{1/2} \preccurlyeq n^{-s}$$

with $s = \min(s_1, s_2)$. In the next step a quantization procedure is applied to $\widetilde{X}(1)$. A general construction due to [4] yields a mapping $T : \mathbb{R}^d \to \mathbb{R}^d$ with a finite range of essentially n^{ds} points such that

$$(\mathbb{E}\|\widetilde{X}(1) - T(\widetilde{X}(1))\|^2)^{1/2} \preccurlyeq n^{-s}.$$

Furthermore, $T(\mathbb{R}^d)$ is contained in a ball of radius $n^{\varepsilon \cdot s/(r+1)}$. All these estimates hold uniformly on $\mathcal{C}$, and for $\widehat{S}(x_0, a, b) \in M_0(\mathbb{R}^d)$ denoting the distribution of $T(\widetilde{X}(1))$ this leads to

$$e(\widehat{S}) \preccurlyeq n^{-s}$$

for smoothness $r = 1$ of the test functions $f \in \mathcal{F}$. To properly handle the remaining case $* = \mathrm{supp}$ and $r > 1$, a support reduction algorithm $\mathcal{R}$ due to [2] is applied to $\widehat{S}(x_0, a, b)$, and this yields a measure $\mathcal{R}(\widehat{S}(x_0, a, b)) \in M_0(\mathbb{R}^d)$ with the following properties. The support size is reduced from order ds, essentially, to the order $d(1 + \varepsilon) \cdot s/r$, while the smoothness of the test functions and the bound on the norm of the support points leads to

$$e(\mathcal{R} \circ \widehat{S}) \preccurlyeq n^{-s}.$$

Unfortunately, $\widehat{S}$ relies on the probability values of a large number of rectangles in $\mathbb{R}^d$ w.r.t. the distributions $S(x_0, \widetilde{a}, \widetilde{b})$ for all $(x_0, a, b) \in \mathcal{C}$, i.e., for infinitely many SDEs. To address this problem we construct a modification of $\widehat{S}$, which is implementable in the real number model and achieves, up to ε, the optimal order η^* of the error in terms of cost* for $* \in \{\mathrm{supp}, \inf, \mathrm{total}\}$. We basically proceed as follows. By suitably rounding the function values of a and b that enter the computation of the approximations $\widetilde{a}$ and $\widetilde{b}$, we end up with a finite number of SDEs that have to be considered further. For every such SDE an iterative quantization of the Euler scheme may be used to compute $\widehat{S}^{\mathrm{mod}}(x_0, a, b) \in M_0(\mathbb{R}^d)$ so that $\widehat{S}^{\mathrm{mod}}$ enjoys all the relevant properties of $\widehat{S}$. Finally, we consider the algorithm $\mathcal{R} \circ \widehat{S}^{\mathrm{mod}}$. In addition to the error bound of order s, and the bounds of orders $d(1 + \varepsilon)$ for the information cost and order $d(1 + \varepsilon) \cdot s/r$ for the support size, $\mathcal{R} \circ \widehat{S}^{\mathrm{mod}}$ satisfies a bound of order $d(1 + \varepsilon) \cdot \max(1, s/r)$ for the total cost. We admit that this result on the algorithm $\mathcal{R} \circ \widehat{S}^{\mathrm{mod}}$ is hardly of a practical relevance, as $\mathcal{R} \circ \widehat{S}^{\mathrm{mod}}$ still relies on heavy precomputation. The main purpose of $\mathcal{R} \circ \widehat{S}^{\mathrm{mod}}$ and its analysis is to show that our lower bounds for the minimal errors are essentially sharp.

Now we turn to the lower bounds for the minimal errors, and here it obviously suffices to consider $* \in \{\inf, \mathrm{supp}\}$. We derive the lower bounds by relating the given computational problem to suitably chosen weighted integration problems on $\mathbb{R}^d$, which represent extremal cases, where either the measure $S(x_0, a, b)$ is known exactly or where $\mathcal{F}$ consists of a single test function. For the case $* = \mathrm{supp}$ the weight will be given by the probability density of a d-dimensional normal distribution, and the proof of the lower bound is straightforward. For the case $* = \inf$ we essentially follow the approach in [21]. We use a construction of fundamental solutions for parabolic initial value problems by means of the parametrix method to represent $\int_{\mathbb{R}^d} f\, dS(x_0, a, b)$ by a series of integrals of increasing dimensions for appropriate subclasses of coefficients a and b. This series is then truncated at the second term to end up with weighted integrals of components of a and b.

We briefly outline the content of the paper. Basic notation and concepts are presented in Sections 2 and 3. Section 4 contains the formulation and discussion of the main results, while the proofs of the upper and the lower bounds for the minimal errors are given in Sections 5 and 6, respectively. An appendix contains the formulation and proof of a comparison result for SDEs and some technical details.

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.

We use $\|\cdot\|_p$ to denote the p-norm for $p \in [1, \infty]$, and we put $\|\cdot\| = \|\cdot\|_\infty$. By $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ we denote the minimal and the maximal eigenvalue of a symmetric non-negative square matrix A, respectively. For $x \in \mathbb{R}^d$ we use $\operatorname{diag}(x)$ to denote the diagonal matrix $(a_{i,j})_{1\le i,j\le d} \in \mathbb{R}^{d\times d}$ with $a_{i,i} = x_i$ for $i = 1, \dots, d$. For a function $h : \mathbb{R}^d \to \mathbb{R}^d$ we use $\operatorname{diag}(h)$ to denote the mapping $x \mapsto \operatorname{diag}(h(x))$. The d-dimensional identity matrix is denoted by E_d, and e_i denotes the i-th unit vector in $\mathbb{R}^d$.

We put

$$\|f\| = \sup_{x\in\mathbb{R}^d} \|f(x)\|, \quad \|f\|_I = \sup_{x\in I} \|f(x)\|$$

for a function $f\colon \mathbb{R}^d \to \mathbb{R}^{m\times k}$ and a set $I \subset \mathbb{R}^d$. Moreover, $\mathcal{L}(d)$ and $\mathcal{L}(d,\kappa)$ denote the classes of functions from $\mathbb{R}^d$ to $\mathbb{R}$ that are Lipschitz continuous and Lipschitz continuous with Lipschitz constant $\kappa \ge 0$ w.r.t. the maximum norm on $\mathbb{R}^d$, respectively, and we put

$$\mathcal{L}^0(d,\kappa) = \{f \in \mathcal{L}(d,\kappa) \mid |f(0)| \le \kappa\}.$$

For $r \in \mathbb{N}$, $K > 0$ and $\beta \ge 0$ we use $\mathcal{F}(d, r, K, \beta)$ to denote the class of functions $f\colon \mathbb{R}^d \to \mathbb{R}$ that have continuous partial derivatives $f^{(\alpha)}$ with

$$|f^{(\alpha)}(x)| \le K \cdot (1 + \|x\|^\beta)$$

for every $x \in \mathbb{R}^d$ and every $\alpha \in \mathbb{N}_0^d$ with $1 \le \|\alpha\|_1 \le r$. Moreover, we put

$$\mathcal{F}^0(d, r, K) = \{f \in \mathcal{F}(d, r, K, 0) \mid |f(0)| \le K\}.$$

For later purposes we note that

$$|f(x) - f(\widetilde{x})| \le c(d, K) \cdot (1 + \|x\|^\beta + \|\widetilde{x}\|^\beta) \cdot \|x - \widetilde{x}\| \tag{2.1}$$

and, consequently,

$$|f(x)| \le c(d, K, f) \cdot (1 + \|x\|^{1+\beta}) \tag{2.2}$$

for every $f \in \mathcal{F}(d,1,K,\beta)$ and all $x, \widetilde{x} \in \mathbb{R}^d$.

Finally, for sequences u_n and v_n of non-negative reals the notion $u_n \preccurlyeq v_n$ means $u_n \leq c\, v_n$ while $u_n \asymp v_n$ is used to denote weak asymptotic equivalence of the two sequences, i.e., $c_1 v_n \leq u_n \leq c_2 v_n$.

3. Algorithms, Error and Cost

We formally describe the class of deterministic algorithms for the computation of probability measures with finite support based on the initial value x_0 and on finitely many sequential evaluations of the coefficients a and b of the corresponding SDE.

Let $N \in \mathbb{N}$ and consider any functions

$$\psi_i \colon \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d, \quad i = 1, \dots, N.$$

Define

$$\mathcal{N}_{\psi_1,\dots,\psi_N} \colon \mathcal{C} \to \mathbb{R}^{d+N(d+d^2)}$$

by

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

with

$$y_1 = (a(\psi_1(x_0)), b(\psi_1(x_0))) \in \mathbb{R}^{d+d^2}$$

and

$$y_\ell = (a(\psi_\ell(x_0, y_1, \dots, y_{\ell-1})), b(\psi_\ell(x_0, y_1, \dots, y_{\ell-1}))) \in \mathbb{R}^{d+d^2}$$

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

$$\psi \colon \mathbb{R}^{d+N(d+d^2)} \to M_0(\mathbb{R}^d).$$

Then

$$\psi \circ \mathcal{N}_{\psi_1,\dots,\psi_N} \colon \mathcal{C} \to M_0(\mathbb{R}^d) \tag{3.1}$$

defines an algorithm, which sequentially evaluates the coefficients a and b from an equation $(x_0, a, b) \in \mathcal{C}$ 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 ψ to the data $\mathcal{N}_{\psi_1,\dots,\psi_N}(x_0, a, b)$ to obtain a probability measure with finite support.

We use

$$\begin{aligned}\mathfrak{S}_N(\mathcal{C}) = \bigl\{\psi \circ \mathcal{N}_{\psi_1,\dots,\psi_N} \mid \psi \colon \mathbb{R}^{d+N(d+d^2)} \to M_0(\mathbb{R}^d),\\ \psi_i \colon \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d, i = 1, \dots, N\bigr\}\end{aligned}$$

to denote the class of all algorithms of the form (3.1) that are based on at most N sequential evaluations of the coefficients of an equation, and we put

$$\mathfrak{S}(\mathcal{C}) = \bigcup_{n \in \mathbb{N}} \mathfrak{S}_N(\mathcal{C}).$$

We use different notions of cost for the comparison of algorithms $\widehat{S} \in \mathfrak{S}(\mathcal{C})$, in particular, the information cost

$$\mathrm{cost}^{\mathrm{inf}}(\widehat{S}) = \min\{N \in \mathbb{N} \mid \widehat{S} \in \mathfrak{S}_N(\mathcal{C})\},$$

i.e., the maximum number of evaluations of the coefficients a, b that are used by $\widehat{S}$, and

$$\mathrm{cost}^{\mathrm{supp}}(\widehat{S}) = \sup\{|\operatorname{supp}(\widehat{S}(x_0, a, b))| \mid (x_0, a, b) \in \mathcal{C}\},$$

i.e., the maximum support size of the probability measures computed by $\widehat{S}$. In the real number model the total cost also takes into account the basic computational operations, i.e., arithmetical operations, assignments, jump instructions and evaluations of elementary functions. Accordingly, we define $\mathrm{cost}^{\mathrm{total}}(\widehat{S})$ as the sum of $\mathrm{cost}^{\mathrm{inf}}(\widehat{S}) + \mathrm{cost}^{\mathrm{supp}}(\widehat{S})$ and the worst case count of all basic computational operations that are carried out by an algorithm $\widehat{S}$ for input equations (x_0, a, b) from the class $\mathcal{C}$.

For the definition of the error of an algorithm $\widehat{S} \in \mathfrak{S}(\mathcal{C})$ we employ metrics on the space $M(\mathbb{R}^d)$ given by

$$\rho_{\mathcal{F}}(\mu_1, \mu_2) = \sup_{f \in \mathcal{F}} \left| \int_{\mathbb{R}^d} f \, d\mu_1 - \int_{\mathbb{R}^d} f \, d\mu_2 \right|,$$

where $\mathcal{F}$ is a class of functions $f \colon \mathbb{R}^d \to \mathbb{R}$ that satisfies

$$\sup_{f \in \mathcal{F}} \sup_{x \in \mathbb{R}^d} \frac{|f(x) - f(0)|}{1 + \|x\|^p} < \infty$$

for some $p \in [1, \infty)$. In particular, $\rho_{\mathcal{F}}$ is called a metric with ζ-structure, if all functions $f \in \mathcal{F}$ are bounded, see [23, p. 72]. The worst case error of $\widehat{S}$ on $\mathcal{C}$ w.r.t. the metric $\rho_{\mathcal{F}}$ is defined by

$$e(\widehat{S}) = \sup_{(x_0, a, b) \in \mathcal{C}} \rho_{\mathcal{F}}(S(x_0, a, b), \widehat{S}(x_0, a, b)).$$

The key quantities in our analysis are the N-th minimal errors

$$e_N^*(\mathcal{C}, \mathcal{F}) = \inf\{e(\widehat{S}) \mid \mathrm{cost}^*(\widehat{S}) \le N\}, \quad N \in \mathbb{N},$$

where $* \in \{\mathrm{inf}, \mathrm{supp}, \mathrm{total}\}$. We determine the decay of these quantities dependent on the smoothness properties of the classes $\mathcal{C}$ and $\mathcal{F}$.

Additionally, we study the related problem of computing the expected values

$$A^f(x_0, a, b) = \int_{\mathbb{R}^d} f \, dS(x_0, a, b), \quad (x_0, a, b) \in \mathcal{C},$$

for a fixed integrand $f \in \mathcal{F}$ by means of deterministic algorithms based on finitely many evaluations of the coefficients a and b. More formally, we consider the class of deterministic algorithms given by

$$\mathfrak{A} = \bigcup_{N \in \mathbb{N}} \mathfrak{A}_N,$$

where

$$\mathfrak{A}_N = \big\{\psi \circ \mathcal{N}_{\psi_1,\dots,\psi_N} \mid \psi\colon \mathbb{R}^{d+N(d+d^2)} \to \mathbb{R},\, \psi_i\colon \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d,\, i = 1,\dots,N\big\},$$

and we define the cost and the error of $\widehat{A} \in \mathfrak{A}$ by

$$\mathrm{cost}(\widehat{A}) = \min\{N \in \mathbb{N} \mid \widehat{A} \in \mathfrak{A}_N\}$$

and

$$e^f(\widehat{A}) = \sup_{(x_0,a,b)\in\mathcal{C}} |A^f(x_0,a,b) - \widehat{A}(x_0,a,b)|.$$

Thus the corresponding N-th minimal errors are given by

$$e_N^f(\mathcal{C}) = \inf\{e^f(\widehat{A}) \mid \mathrm{cost}(\widehat{A}) \le N\}, \quad N \in \mathbb{N}.$$

The study of $e_N^f(\mathcal{C})$ is primarily motivated by the following fact. If $\widehat{S} \in \mathfrak{S}$ then

$$\widehat{A}(x_0,a,b) = \int_{\mathbb{R}^d} f\, d\widehat{S}(x_0,a,b), \quad (x_0,a,b) \in \mathcal{C},$$

defines an algorithm $\widehat{A} \in \mathfrak{A}$ with $\mathrm{cost}(\widehat{A}) \le \mathrm{cost}^{\mathrm{inf}}(\widehat{S})$, and therefore

$$e_N^{\mathrm{inf}}(\mathcal{C},\mathcal{F}) \ge e_N^f(\mathcal{C}) \tag{3.2}$$

for every $N \in \mathbb{N}$ and $f \in \mathcal{F}$.

4. Results

We consider the classes

$$\mathcal{C} = \mathcal{C}(d,s_1,s_2,K) = [-K,K]^d \times (\mathcal{F}^0(d,s_1,K))^d \times (\mathcal{F}^0(d,s_2,K))^{d\times d}$$

of SDEs, and we use the metrics $\rho_{\mathcal{F}}$ with

$$\mathcal{F} = \mathcal{F}(d,r,K,\beta).$$

Theorem 4.1. *Let $d,r,s_1,s_2 \in \mathbb{N}$, $K > 0$ and $\beta \ge 0$. Then*

$$N^{-\eta^*} \preccurlyeq e_N^*(\mathcal{C},\mathcal{F}) \preccurlyeq N^{-\eta^*+\varepsilon}$$

with

$$\eta^* = \begin{cases} \min(s_1,s_2)/d, & \text{if } * = \textit{inf}, \\ r/d, & \text{if } * = \textit{supp}, \\ \min(r,s_1,s_2)/d, & \text{if } * = \textit{total}, \end{cases}$$

for every $\varepsilon > 0$ and $* \in \{\textit{inf}, \textit{supp}, \textit{total}\}$.

We conjecture that the upper bounds in Theorem 4.1 actually hold with $\varepsilon = 0$. Obviously, it suffices to establish the lower bounds for $* \in \{\inf, \mathrm{supp}\}$.

Fix $d, r, s_1, s_2 \in \mathbb{N}$, $K > 0$ and $\beta \geq 0$ as well as $\varepsilon > 0$. For the proof of the corresponding upper bounds we first construct a sequence of algorithms $\widehat{S}_{n,\varepsilon}$ such that $\lim_{n\to\infty} \mathrm{cost}^*(\widehat{S}_{n,\varepsilon}) = \infty$ as well as

$$e(\widehat{S}_{n,\varepsilon}) \preccurlyeq (\mathrm{cost}^*(\widehat{S}_{n,\varepsilon}))^{-\eta^*+\varepsilon} \tag{4.1}$$

simultaneously for $* = \inf$ and $* = \mathrm{supp}$, see Section 5.1. In Section 5.2 we present a modified version $\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}$ of $\widehat{S}_{n,\varepsilon}$, which satisfies $\lim_{n\to\infty} \mathrm{cost}^*(\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}) = \infty$ and achieves (4.1) with $\widehat{S}_{n,\varepsilon}$ replaced by $\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}$ as well as

$$e(\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}) \preccurlyeq (\mathrm{cost}^{\mathrm{total}}(\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}))^{-\eta^{\mathrm{total}}+\varepsilon}.$$

However, implementing $\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}$ requires at least $n^{d^2 n^d \min(s_1,s_2)}$ basic computational operations in a precomputation step, so that these algorithms are implementable within the real number model, but this result is of no practical use.

Remark 4.1. It is natural to ask whether the upper bounds in Theorem 4.1 can be achieved by algorithms that do not rely on heavy precomputation. The answer to this question is positive, at least, if further restrictions are imposed on the input equations. Let

$$r = 1 \quad \text{and} \quad \beta = 0, \tag{4.2}$$

i.e., we essentially take Lipschitz continuous integrands f with Lipschitz constant K, and consider the subclass $\widetilde{\mathcal{C}} \subset \mathcal{C}(d, s_1, s_2, K)$ of equations (x_0, a, b) such that

$$\|a\|, \|b\| \leq K, \ \inf_{x\in\mathbb{R}^d} \lambda_{\min}(bb^{\mathrm{T}}(x)) \geq \delta \tag{4.3}$$

with a fixed $\delta > 0$. Assume further that

$$\min(s_1, s_2) \geq 6.$$

In the forthcoming work [16] algorithms $\widetilde{S}_{n,\varepsilon}$ are presented that iteratively combine non-uniform simplified weak order 2.0 Itô-Taylor steps with a reduction strategy for the size and the diameter of the support of discrete measures. These methods are easy to implement and their worst case errors with respect to the class $\widetilde{\mathcal{C}}$ satisfy

$$e(\widetilde{S}_{n,\varepsilon}) \preccurlyeq (\mathrm{cost}^{\inf}(\widetilde{S}_{n,\varepsilon}))^{-\min(s_1,s_2)/d+\varepsilon}$$

as well as

$$e(\widetilde{S}_{n,\varepsilon}) \preccurlyeq (\mathrm{cost}^{\mathrm{total}}(\widetilde{S}_{n,\varepsilon}))^{-1/d+\varepsilon}.$$

It is easy to see that the lower bounds from Theorem 4.1 are still valid under the restriction (4.3), and therefore the sequence of algorithms $(\widetilde{S}_{n,\varepsilon})_{n\in\mathbb{N}}$ performs asymptotically optimal, up to the arbitrarily small exponent ε, with respect to cost^* for $* \in \{\mathrm{supp}, \inf, \mathrm{total}\}$.

A methodology based on quadrature formulas for integration with respect to the Wiener measure is introduced in [8] and further developed in particular in [9, 11, 12, 13]. These quadrature formulas have nodes of bounded variation and are exact for Stratonovich integrals up to a certain degree m. By (approximately) solving an equation (x_0, a, b) in its Stratonovich formulation along the suitably scaled paths of the quadrature formula, i.e., by solving a collection of ODEs, up to some time instance t_0 one gets, in principle, an approximation to the distribution of the corresponding solution at time t_0. An iteration of this procedure with n non-uniform time-steps combined with a recombination technique yields a discrete approximation to $S(x_0, a, b)$. We discuss the performance of a variant of this method, see [11, Sec. 4.4], in the previous setting, i.e., for r and β given by (4.2) and the subclass $\widetilde{\mathcal{C}} \subset \mathcal{C}(d, s_1, s_2, K)$ of equations specified by (4.3). Assume

$$\min(s_1, s_2) \geq 3.$$

Taking $m = \min(s_1, s_2) - 1$, which is best possible in this case, yields algorithms $\overline{S}_n$ that satisfy

$$e(\overline{S}_n) \preccurlyeq \ln(n) \cdot n^{-(\min(s_1,s_2)-2)/2}$$

as well as

$$\mathrm{cost}^{\mathrm{total}}(\overline{S}_n) \preccurlyeq n^{(\min(s_1,s_2)-1)\cdot d/2}.$$

Thus

$$e(\overline{S}_n) \preccurlyeq \ln(\mathrm{cost}^*(\overline{S}_n)) \cdot (\mathrm{cost}^*(\overline{S}_n))^{-(\min(s_1,s_2)-2)/((\min(s_1,s_2)-1)\cdot d)}$$

for $* \in \{\mathrm{supp}, \mathrm{total}\}$, which is close to optimality for large smoothness parameters s_1 and s_2. We stress however, that the construction of the underlying exact quadrature formulas on the Wiener space is a nontrivial task, see [12, 13].

We turn to results on minimal errors for the problem of computing expected values $A^f(x_0, a, b)$ for a fixed integrand f. By Theorem 4.1 and (3.2) we have the upper bound

$$e_N^f\big(\mathcal{C}(d, s_1, s_2, K)\big) \preccurlyeq N^{-\min(s_1,s_2)/d+\varepsilon}$$

for every $f \in \mathcal{F}(d, r, K, \beta)$ and all $d, r, s_1, s_2 \in \mathbb{N}$, $K > 0$, $\beta \geq 0$, $\varepsilon > 0$. We present lower bounds for two particular classes of SDEs, namely

$$\mathcal{C}^{(1)} = \{0\} \times (\mathcal{F}^0(d, s, K))^d \times \{K \cdot E_d\},$$

i.e., equations with zero initial value, a drift coefficient of smoothness s and additive noise, and

$$\mathcal{C}^{(2)} = \{0\} \times \{0\}^d \times \{\mathrm{diag}(h) \mid h \in (\mathcal{F}^0(d, s, K))^d\},$$

i.e., equations with zero initial value, zero drift coefficient and a diagonal diffusion coefficient of smoothness s.

Theorem 4.2. *Let $d, s \in \mathbb{N}$, $K > 0$ and $\beta \geq 0$.*

(i) *Assume that $f \in \mathcal{F}(d, 1, K, \beta)$ is not constant. Then*

$$e_N^f\big(\mathcal{C}^{(1)}\big) \succcurlyeq N^{-s/d}.$$

(ii) *Assume that $f \in \mathcal{F}(d, 2, K, \beta)$ is not multilinear. Then*

$$e_N^f\big(\mathcal{C}^{(2)}\big) \succcurlyeq N^{-s/d}.$$

Remark 4.2. We briefly comment on the conditions imposed on the integrand f in Theorem 4.2. If $f \in \mathcal{F}(d, 1, K, \beta)$ is constant or $f \in \mathcal{F}(d, 2, K, \beta)$ is multilinear then $A^f(x_0, a, b) = f(0)$, and therefore $e_N^f\big(\mathcal{C}^{(i)}\big) = 0$ for all $N \in \mathbb{N}$ and $i = 1, 2$.

Remark 4.3. Consider the classes $\mathcal{C}(d, s_1, s_2, K)$ and $\mathcal{F}(d, r, K, \beta)$ as in Theorem 4.1. Suppose that we permit algorithms to take values in the space $M(\mathbb{R}^d)$ instead of $M_0(\mathbb{R}^d)$. Then (3.2) still holds, and using Theorem 4.2 we conclude that $M(\mathbb{R}^d)$ is not more powerful, up to an arbitrarily small exponent ε, than $M_0(\mathbb{R}^d)$ with respect to $\mathrm{cost}^{\mathrm{inf}}$.

Remark 4.4. In [21], the asymptotic behavior of minimal errors $e_N^f(\mathcal{C})$ for classes $\mathcal{C}$ of equations with fixed initial value, fixed additive noise and varying non-autonomous drift coefficients $a\colon \mathbb{R}^d \times [0, 1] \to \mathbb{R}^d$ is analyzed. Similar to Theorem 4.2(i), the resulting lower bounds are of the form $c \cdot N^{-s/(d+1)}$, where the parameter $s \in \mathbb{N}$ specifies the smoothness of the drift coefficients with respect to the state variable $x \in \mathbb{R}^d$. It seems, however, that Theorem 4.2(i) may not be deduced from the results in [21].

Remark 4.5. Theorems 4.1 and 4.2 only cover the case that both the solution X and the driving Brownian motion W are d-dimensional. Consider, more generally, the case that W is m-dimensional.

Assume $m > d$. Then Theorems 4.1 and 4.2 still hold true. Indeed, for the upper bounds in Theorem 4.1 for $* \in \{\mathrm{inf}, \mathrm{supp}\}$ the proofs from Section 5.1 carry over almost literally. The upper bound for $* = \mathrm{total}$ is obtained by an obvious modification of the algorithm $\widehat{S}^{\mathrm{mod}}_{n,\varepsilon}$ in Section 5.2 and a straightforward adaptation of the corresponding proofs. The lower bounds follow from the lower bounds for the case $m = d$ and the fact that every system (1.1) can be written as a system driven by an m-dimensional Brownian motion with $m > d$.

Assume $m < d$. Then the upper bounds in Theorem 4.1 are still valid. This follows from the fact that in this case any system (1.1) driven by an m-dimensional Brownian motion may also be considered as a system driven by a d-dimensional Brownian motion and the proven upper bounds for the case $m = d$. The lower bounds for the minimal errors e_N^{supp} in Theorem 4.1 hold true as well. For a proof, consider the equation (x_0, a, b)

given by

$$x_0 = 0,\ a(x) = K \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{d-1} \end{pmatrix} \in \mathbb{R}^d,\ b(x) = \begin{pmatrix} K & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{d \times m},$$

which satisfies

$$S(x_0, a, b) = N(0, \Sigma)$$

with a positive-definite covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$, and proceed as in Section 6.2 with equation (x_0, a, b) in place of equation $(0, 0, K \cdot E_d)$. Lower bounds for the minimal errors e_N^{inf} and for the minimal errors considered in Theorem 4.2 are given in this case by $N^{-\min(s_1,s_2)/m}$ and $N^{-s/m}$, respectively. This follows from the proven lower bounds in Theorem 4.1 and Theorem 4.2 for the case of SDEs with an m-dimensional solution X since the latter obviously constitutes subproblems with smaller minimal errors.

To sum up, the dimension m of the driving Brownian motion W does not alter the order of convergence of the minimal errors under consideration if $m > d$ or if the size of support, i.e., $\text{cost}^{\text{supp}}$, is used as a measure of cost. In the case $m < d$, if the information cost, i.e., cost^{inf}, is considered, we have $N^{-\min(s_1,s_2)/m}$ as a lower bound and $N^{-\min(s_1,s_2)/d+\varepsilon}$ as an upper bound for the respective N-th minimal errors, and the precise asymptotic behavior of the latter quantities is unknown to us.

5. Proof of the Upper Bounds in Theorem 4.1

Fix $d, r, s \in \mathbb{N}$, $K > 0$ and $\beta \geq 0$ and assume w.l.o.g. that

$$s_1 = s_2 = s.$$

Put $\mathcal{C} = \mathcal{C}(d, s, s, K)$ and $\mathcal{F} = \mathcal{F}(d, r, \beta, K)$.

In this section unspecified positive constants c and $c(\cdot)$ may only depend on the parameters d, r, s, K, β listed above and the parameter ε introduced below, additionally to the parameters eventually specified in brackets.

5.1. **The case $* \in \{\textbf{inf}, \textbf{supp}\}$.** For every $n \in \mathbb{N}$ and $\varepsilon \in (0, 1]$ we present an algorithm

$$\widehat{S}_{n,\varepsilon} \colon \mathcal{C} \to M_0(\mathbb{R}^d),$$

which consists of the following ingredients.

1) Evaluation of the coefficients a and b from any input $(x_0, a, b) \in \mathcal{C}$ on the uniform grid

$$\mathcal{X} = \left(1/n \cdot \mathbb{Z} \cap \left[-\lceil n^\varepsilon \rceil, \lceil n^\varepsilon \rceil\right]\right)^d,$$

i.e., the mapping

$$\mathcal{N} \colon \mathbb{R}^{\mathbb{R}^d} \to \mathbb{R}^{|\mathcal{X}|}, \quad f \mapsto (f(x))_{x \in \mathcal{X}},$$

is applied to each of the respective components $a_i, b_{i,j}$. The available information about an input (x_0, a, b) is thus provided by the vector

$$\overline{\mathcal{N}}(x_0, a, b) \in \mathbb{R}^{d+|\mathcal{X}|\cdot(d+d^2)},$$

where

$$\overline{\mathcal{N}} \colon \mathcal{C} \to \mathbb{R}^{d+|\mathcal{X}|\cdot(d+d^2)}, \quad (x_0, a, b) \mapsto (x_0, (\mathcal{N}(a_i))_{i=1,\dots,d}, (\mathcal{N}(b_{i,j}))_{i,j=1,\dots,d}).$$

2) A linear mapping

$$\mathcal{A} \colon \mathbb{R}^{|\mathcal{X}|} \to \mathcal{L}(d),$$

which satisfies

(i) $\mathcal{A}(\mathcal{N}(f)) \in \mathcal{L}^0(d, c(\kappa))$ for every $f \in \mathcal{L}^0(d, \kappa)$ and $\kappa > 0$,

(ii) $\|f - \mathcal{A}(\mathcal{N}(f))\|_{[-n^\varepsilon, n^\varepsilon]^d} \le c \cdot n^{-s}$ for every $f \in \mathcal{F}^0(d, s, K)$.

We use $\mathcal{A}$ to approximate the equation parameters (x_0, a, b) by the parameters

$$\overline{\mathcal{A}}(\overline{\mathcal{N}}(x_0, a, b)) \in \mathbb{R}^d \times (\mathcal{L}(d))^{d+d^2},$$

where

$$\overline{\mathcal{A}} \colon \mathbb{R}^d \times (\mathbb{R}^{|\mathcal{X}|})^{d+d^2} \to \mathbb{R}^d \times (\mathcal{L}(d))^{d+d^2}, \quad (x, y) \mapsto (x, (\mathcal{A}(y_i))_{i=1,\dots,d+d^2}).$$

3) A mapping

$$\mathcal{Q} \colon \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^{d^2} \to M_0(\mathbb{R}^d),$$

which satisfies

(i) $|\operatorname{supp}(\mathcal{Q}(x, g, h))| \le c \cdot n^{d(1+\varepsilon)\cdot s/r}$,

(ii) $\rho_{\mathcal{F}}(S(x, g, h), \mathcal{Q}(x, g, h)) \le c(\kappa) \cdot n^{-s}$

for all $x \in [-\kappa, \kappa]^d$, $g \in (\mathcal{L}^0(d, \kappa))^d$, $h \in (\mathcal{L}^0(d, \kappa))^{d^2}$ and $\kappa > 0$. For every input $(x_0, a, b) \in \mathcal{C}$ the discrete measure

$$\mathcal{Q}(\overline{\mathcal{A}}(\overline{\mathcal{N}}(x_0, a, b))) \in M_0(\mathbb{R}^d)$$

serves as an approximation to $S(x_0, a, b)$.

The construction of $\mathcal{A}$ could be based, e.g., on piecewise polynomial interpolation with total degree of at most $s - 1$ in each coordinate and interpolation nodes given by $\mathcal{X}$. A possible choice of $\mathcal{Q}$ is presented in Sections 5.1.1 and 5.1.2, see also Section 5.2.

We define

$$\widehat{S}_{n,\varepsilon} = \mathcal{Q} \circ \overline{\mathcal{A}} \circ \overline{\mathcal{N}}. \tag{5.1}$$

Theorem 5.1. *The algorithm* $\widehat{S}_{n,\varepsilon}$ *satisfies*

$$e(\widehat{S}_{n,\varepsilon}) \le c \cdot n^{-s}$$

and

$$\mathrm{cost}^*(\widehat{S}_{n,\varepsilon}) \le c \cdot \begin{cases} n^{d(1+\varepsilon)}, & \text{if } * = \textit{inf}, \\ n^{d(1+\varepsilon)\cdot s/r}, & \text{if } * = \textit{supp}. \end{cases}$$

Clearly, Theorem 5.1 implies the upper bounds in Theorem 4.1, except for $* = \text{total}$.

Proof of Theorem 5.1. We have

$$|\mathcal{X}| \le c \cdot n^{d(1+\varepsilon)}, \tag{5.2}$$

which yields the bound for $\mathrm{cost}^{\mathrm{inf}}(\widehat{S}_{n,\varepsilon})$. Property 3)(i) of the mapping $\mathcal{Q}$ implies the bound for $\mathrm{cost}^{\mathrm{supp}}(\widehat{S}_{n,\varepsilon})$.

Let $(x_0, a, b) \in \mathcal{C}$ and put

$$(x_0, \widetilde{a}, \widetilde{b}) = \overline{\mathcal{A}}(\overline{\mathcal{N}}(x_0, a, b)).$$

Thus

$$\widetilde{a}_i = \mathcal{A}(\mathcal{N}(a_i)),\ \widetilde{b}_{i,j} = \mathcal{A}(\mathcal{N}(b_{i,j}))$$

for $i, j = 1, \dots, d$. By assumption, $a_i, b_{i,j} \in \mathcal{L}^0(d, K)$. Hence

$$\widetilde{a}_i, \widetilde{b}_{i,j} \in \mathcal{L}^0(d, c) \tag{5.3}$$

follows from property 2)(i) of $\mathcal{A}$. Furthermore, by property 2)(ii) of $\mathcal{A}$,

$$\|a_i - \widetilde{a}_i\|_{[-n^\varepsilon, n^\varepsilon]^d} + \|b_{i,j} - \widetilde{b}_{i,j}\|_{[-n^\varepsilon, n^\varepsilon]^d} \le c \cdot n^{-s}. \tag{5.4}$$

Let X and $\widetilde{X}$ denote the solutions of the SDEs corresponding to (x_0, a, b) and $(x_0, \widetilde{a}, \widetilde{b})$, respectively. By the properties of $a, \widetilde{a}, b, \widetilde{b}$ we have

$$\mathbb{E}\Big(\sup_{t\in[0,1]} (\|X(t)\|^p + \|\widetilde{X}(t)\|^p)\Big) \le c(p) \cdot (1 + \|x_0\|^p) \tag{5.5}$$

for every $p \ge 1$, see, e.g., [6, Theorem 5.2.3], and using (5.4) we obtain

$$(\mathbb{E}\|X(1) - \widetilde{X}(1)\|^2)^{1/2} \le c \cdot n^{-s} \tag{5.6}$$

by a comparison result for SDEs, see Lemma 7.1 in Section 7.1.

Let $f \in \mathcal{F}$. By (2.1), (5.5) and (5.6),

$$\begin{aligned} &|\mathbb{E}(f(X(1))) - \mathbb{E}(f(\widetilde{X}(1))| \\ &\qquad \le c \cdot \big(1 + (\mathbb{E}\|X(1)\|^{2\beta})^{1/2} + (\mathbb{E}\|\widetilde{X}(1)\|^{2\beta})^{1/2}\big) \cdot \big(\mathbb{E}\|X(1) - \widetilde{X}(1)\|^2\big)^{1/2} \le c \cdot n^{-s}. \end{aligned}$$

It follows

$$\rho_{\mathcal{F}}(S(x_0, a, b), S(x_0, \widetilde{a}, \widetilde{b})) \le c \cdot n^{-s}.$$

By (5.3) and property 3)(ii) of the mapping $\mathcal{Q}$ we have

$$\rho_{\mathcal{F}}(S(x_0, \widetilde{a}, \widetilde{b}), \mathcal{Q}(x_0, \widetilde{a}, \widetilde{b})) \le c \cdot n^{-s},$$

which completes the proof. □

We construct $\mathcal{Q}$ in two steps, i.e.,

$$\mathcal{Q} = \mathcal{Q}_2 \circ \mathcal{Q}_1,$$

where the mapping

$$\mathcal{Q}_1 \colon \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^{d^2} \to M_0(\mathbb{R}^d)$$

is obtained by applying a quantization procedure for random variables with finite moments of any order, see Section 5.1.1, and satisfies

$$\begin{aligned} \max_{x \in \operatorname{supp}(\mathcal{Q}_1(x_0,a,b))} \|x\| &\leq n^{\varepsilon \cdot s/(r+\beta)}, \\ \rho_{\mathcal{F}(d,1,K,\beta)}(S(x_0,a,b), \mathcal{Q}_1(x_0,a,b)) &\leq c(\kappa) \cdot n^{-s} \end{aligned} \tag{5.7}$$

for all $x_0 \in [-\kappa,\kappa]^d$, $a \in (\mathcal{L}^0(d,\kappa))^d$, $b \in (\mathcal{L}^0(d,\kappa))^{d^2}$ and $\kappa > 0$, and the mapping

$$\mathcal{Q}_2 \colon M_0(\mathbb{R}^d) \to M_0(\mathbb{R}^d)$$

is obtained by applying a measure support reduction algorithm, see Section 5.1.2, and satisfies

$$\begin{aligned} |\operatorname{supp}(\mathcal{Q}_2(\mu))| &\leq c \cdot n^{d(1+\varepsilon)\cdot s/r}, \\ \rho_{\mathcal{F}(d,r,K,\beta)}(\mu, \mathcal{Q}_2(\mu)) &\leq c \cdot n^{-(1+\varepsilon)s} \cdot (1 + \max_{x \in \operatorname{supp}(\mu)} \|x\|^{r+\beta}) \end{aligned} \tag{5.8}$$

for all $\mu \in M_0(\mathbb{R}^d)$. Clearly, (5.7) and (5.8) imply properties 3)(i) and 3)(ii).

5.1.1. *The mapping $\mathcal{Q}_1$.* Let $\gamma > 1$ and $N \in \mathbb{N}$. Put

$$J = \lfloor N^{1/\gamma} \rfloor$$

and define

$$N_j = \lfloor N/j^\gamma \rfloor,$$

as well as

$$y_{j,k} = j - 1 + (k-1)/N_j$$

for $k = 1, \ldots, N_j + 1$ and $j = 1, \ldots, J$. Let

$$T_{\gamma,N} = \sum_{j=1}^{J} \sum_{k=1}^{N_j} \bigl(y_{j,k} \cdot 1_{[y_{j,k}, y_{j,k+1})} - y_{j,k} \cdot 1_{(-y_{j,k+1}, -y_{j,k}]} \bigr)$$

and consider the mapping

$$T^{(d)}_{\gamma,N} \colon \mathbb{R}^d \to \mathbb{R}, \quad (x_1, \ldots, x_d) \mapsto (T_{\gamma,N}(x_1), \ldots, T_{\gamma,N}(x_d)).$$

Thus

$$T^{(d)}_{\gamma,N}(\mathbb{R}^d) \subset [-N^{1/\gamma}, N^{1/\gamma}]^d \tag{5.9}$$

and we have

$$(2N-1)^d \le |T^{(d)}_{\gamma,N}(\mathbb{R}^d)| = \left(2\sum_{j=1}^{J} N_j - 1\right)^d \le \big(2N\gamma/(\gamma-1)\big)^d. \tag{5.10}$$

Note further that

$$\|T^{(d)}_{\gamma,N}(x)\|_p \le \|x\|_p \tag{5.11}$$

for every $x \in \mathbb{R}^d$ and $p \in [1,\infty]$.

We use the following error bound for quantization based on $T^{(d)}_{\gamma,N}$.

Lemma 5.1. *For every $q \ge 1$, every $p > 1 + (1+\gamma)q$ and every $\mathbb{R}^d$-valued random variable Z we have*

$$\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q \le c(q,p)\max\big(1, \mathbb{E}\|Z\|^p\big)N^{-q}.$$

For convenience of the reader we provide a proof of this fact. See [4, Lemma 7.1] for a more general result.

Proof. For $j \in \mathbb{N}$ put $B_j = [j-1, j)$. Then

$$\begin{aligned}\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q &= \sum_{j=1}^{\infty} \mathbb{E}\big(1_{B_j}(\|Z\|)\cdot\|Z - T^{(d)}_{\gamma,N}(Z)\|^q\big) \\ &\le \sum_{j=1}^{J} N_j^{-q}\cdot\mathbb{P}(\|Z\| \in B_j) + \sum_{j>J} j^q \cdot \mathbb{P}(\|Z\| \in B_j).\end{aligned}$$

We have

$$N_j^{-q}\cdot\mathbb{P}(\|Z\| \in B_j) \le c(q)N^{-q}j^{\gamma q}\cdot\mathbb{P}(\|Z\| \ge j-1) \le c(q,p)N^{-q}j^{\gamma q - p}\cdot\mathbb{E}\|Z\|^p$$

for $j = 2, \dots, J$, and

$$j^q\cdot\mathbb{P}(\|Z\| \in B_j) \le (j^\gamma/N)^q j^q \cdot\mathbb{P}(\|Z\| \ge j-1) \le c(p)N^{-q}j^{(1+\gamma)q-p}\cdot\mathbb{E}\|Z\|^p$$

for $j > J$. Hence

$$\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q \le c(q,p)N^{-q}\left(1 + \mathbb{E}\|Z\|^p\cdot\sum_{j\ge 2} j^{(1+\gamma)q-p}\right),$$

which finishes the proof. □

Take

$$\gamma = (r+\beta)\cdot\varepsilon^{-1},\ \ N = n^s$$

and define the mapping $\mathcal{Q}_1$ by

$$\mathcal{Q}_1(x_0, a, b) = \mathbb{P}_{T^{(d)}_{\gamma,N}(X(1))},$$

where $X(1)$ denotes the solution of the SDE given by (x_0, a, b) at time $t = 1$.

Lemma 5.2. *Let $\kappa > 0$ and assume $\|x_0\| \le \kappa$ and $a_i, b_{i,j} \in \mathcal{L}^0(d, \kappa)$ for $i, j = 1, \ldots, d$. Then the measure $\mathcal{Q}_1(x_0, a, b)$ satisfies* (5.7).

Proof. By definition,

$$\operatorname{supp}(\mathcal{Q}_1(x_0, a, b)) \subset T_{(r+\beta)\cdot\varepsilon^{-1}, n^s}(\mathbb{R}^d).$$

Thus

$$\max_{x \in \operatorname{supp}(\mathcal{Q}_1(x_0,a,b))} \|x\| \le n^{\varepsilon \cdot s/(r+\beta)}$$

is a consequence of (5.9).

For the proof of the error estimate in (5.7) let X denote the solution of an SDE given by (x_0, a, b) and put

$$V = T_{(r+\beta)\cdot\varepsilon^{-1}, n^s}(X(1)).$$

By the properties of x_0, a and b and by (5.11) we have

$$\mathbb{E}\|V\|^q \le \mathbb{E}\|X(1)\|^q \le c(\kappa, q)$$

for every $q \ge 1$. Furthermore, by Lemma 5.1,

$$\mathbb{E}\|X(1) - V\|^q \le c(\kappa, q) \cdot n^{-sq}$$

for every $q \ge 1$. Use (2.1) to conclude that

$$\begin{aligned} &|\mathbb{E}(f(X(1))) - \mathbb{E}(f(V))| \\ &\qquad \le c \cdot \mathbb{E}\big((1 + \|X(1)\|^\beta + \|V\|^\beta) \cdot \|X(1) - V\|\big) \le c(\kappa) \cdot n^{-s} \end{aligned}$$

for all $f \in \mathcal{F}(d, 1, K, \beta)$, which completes the proof. □

5.1.2. *The mapping $\mathcal{Q}_2$.* Let $\tau \in \mathbb{N}_0$ and let $\mathcal{P}_\tau$ denote the set of polynomials on $\mathbb{R}^d$ of total degree at most τ. By a well-known sequential support point elimination procedure due to [2] we obtain an algorithm

$$\mathcal{R}_\tau \colon M_0(\mathbb{R}^d) \to M_0(\mathbb{R}^d),$$

which satisfies

$$\operatorname{supp}(\mathcal{R}_\tau(\mu)) \subset \operatorname{supp}(\mu) \tag{5.12}$$

as well as

$$|\operatorname{supp}(\mathcal{R}_\tau(\mu))| \le \dim(\mathcal{P}_\tau) \tag{5.13}$$

and

$$\int_{\mathbb{R}^d} p \, d\mathcal{R}_\tau(\mu) = \int_{\mathbb{R}^d} p \, d\mu \tag{5.14}$$

for every $p \in \mathcal{P}_\tau$.

We define

$$\mathcal{Q}_2 = \mathcal{R}_\tau$$

with

$$\tau = \max\bigl(r-1, \lceil n^{(1+\varepsilon)\cdot s/r} \rceil\bigr).$$

Lemma 5.3. *The mapping $\mathcal{Q}_2$ satisfies* (5.8).

Proof. Let $\mu \in M_0(\mathbb{R}^d)$. By (5.13) we have

$$|\operatorname{supp}(\mathcal{Q}_2(\mu))| \le \binom{\tau+d}{d} \le 2^d \cdot \tau^d,$$

which yields the first estimate in (5.8).

Put

$$\eta = \max_{x \in \operatorname{supp}(\mu)} \|x\|.$$

Let $f \in \mathcal{F}(d,r,K,\beta)$. By (5.12) and (5.14) we have

$$\begin{aligned}\left|\int_{\mathbb{R}^d} f\, d\mu - \int_{\mathbb{R}^d} f\, d\mathcal{Q}_2(\mu)\right| &\le \int_{\mathbb{R}^d} |f-p|\, d\mu + \int_{\mathbb{R}^d} |f-p|\, d\mathcal{Q}_2(\mu)\\ &\le 2 \max_{z \in \operatorname{supp}(\mu)} |f(z)-p(z)|\\ &\le 2\|f-p\|_{[-\eta,\eta]^d}\end{aligned}$$

for every $p \in \mathcal{P}_\tau$. Using a Jackson-type theorem, see, e.g., [24, Theorem 3.4], we get

$$\begin{aligned}\inf_{p \in \mathcal{P}_\tau} \|f-p\|_{[-\eta,\eta]^d} &\le c \cdot \tau^{-r} \cdot \eta^r \cdot \max_{|\alpha|=r} \|f^{(\alpha)}\|_{[-\eta,\eta]^d}\\ &\le c \cdot n^{-(1+\varepsilon)\cdot s} \cdot (1+\eta^{r+\beta}),\end{aligned}$$

which completes the proof. □

5.2. **The case $*=$ total.** The algorithm $\widehat{S}_{n,\varepsilon} = \mathcal{Q}_2 \circ \mathcal{Q}_1 \circ \overline{\mathcal{A}} \circ \overline{\mathcal{N}}$ is hardly implementable since the computation of the discrete measures $\mathcal{Q}_1(\overline{\mathcal{A}}(\overline{\mathcal{N}}(x_0,a,b)))$ requires the knowledge of the marginal probabilities

$$S(x_0,\widetilde{a},\widetilde{b})(R) \tag{5.15}$$

for certain rectangles $R \subset \mathbb{R}^d$ and all of the equations $(x_0,\widetilde{a},\widetilde{b}) = \overline{\mathcal{A}}(\overline{\mathcal{N}}(x_0,a,b))$ with $(x_0,a,b) \in \mathcal{C}$, see Section 5.1.1. We present a modified version $\widehat{S}^{\text{mod}}_{n,\varepsilon}$ of $\widehat{S}_{n,\varepsilon}$, which is implementable in the real number model and achieves, up to the arbitrarily small exponent ε, the optimal order η^* in terms of cost* for $* \in \{\text{supp}, \text{inf}, \text{total}\}$, see Theorem 4.1. The idea is, essentially, to replace the mapping $\mathcal{Q}_1$ by a mapping based on iterative quantization of the Euler scheme, which leads to sufficiently good approximations of the marginal probabilities (5.15), and to approximate the set $\mathcal{C}$ of equations by a finite set of equations such that the computation of all of the approximating discrete measures can be carried out in a precomputation step.

To be more precise let $Z_1, Z_2, \dots$ denote an independent sequence of d-dimensional standard normally distributed random variables. Fix $x_0 \in \mathbb{R}^d$ as well as $a\colon \mathbb{R}^d \to \mathbb{R}^d$

and $b\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$. Let $m \in \mathbb{N}$ and consider the corresponding weak Euler scheme with step-size $1/m$ given by

$$\begin{aligned} U_0 &= x_0, \\ U_\ell &= U_{\ell-1} + a(U_{\ell-1}) \cdot m^{-1} + b(U_{\ell-1}) \cdot m^{-1/2} \cdot Z_\ell \end{aligned}$$

for $\ell = 1, \dots, m$. We obtain a quantization $(\widetilde{U}_\ell)_{\ell=0,\dots,m}$ of this scheme by iteratively applying the mapping $T^{(d)}_{\gamma,N}$ to the consecutive single Euler steps, i.e., we put

$$\overline{U}_0 = \widetilde{U}_0 = x_0$$

and we define recursively

$$\begin{aligned} \overline{U}_\ell &= \widetilde{U}_{\ell-1} + a(\widetilde{U}_{\ell-1}) \cdot m^{-1} + b(\widetilde{U}_{\ell-1}) \cdot m^{-1/2} \cdot Z_\ell, \\ \widetilde{U}_\ell &= T^{(d)}_{\gamma,N}(\overline{U}_\ell) \end{aligned}$$

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

We take

$$\gamma = 3(r+\beta)\cdot\varepsilon^{-1}, \quad N = n^{3s}, \quad m = n^{2s}$$

and we define a mapping

$$\mathcal{Q}_1^{\mathrm{mod}}\colon \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^{d^2} \to M_0(\mathbb{R}^d)$$

by

$$\mathcal{Q}_1^{\mathrm{mod}}(x_0, a, b) = \mathbb{P}_{\widetilde{U}_m}.$$

In order to replace $\mathcal{C}$ by a finite set of equations we employ a projection

$$\mathcal{P}\colon \mathbb{R} \to \mathcal{Y}$$

into the uniform grid

$$\mathcal{Y} = 1/n^s \cdot \mathbb{Z} \cap \big[-2K \cdot \lceil n^\varepsilon \rceil, 2K \cdot \lceil n^\varepsilon \rceil\big]$$

such that

$$|x - \mathcal{P}(x)| = \min_{y\in\mathcal{Y}} |x-y|$$

for all $x \in \mathbb{R}$. Hereby the available data $\overline{\mathcal{N}}(x_0,a,b)$ from an input $(x_0,a,b) \in \mathcal{C}$ is approximated by the vector

$$\overline{\mathcal{P}}(\overline{\mathcal{N}}(x_0,a,b)) \in \mathcal{Y}^{d+|\mathcal{X}|\cdot(d+d^2)},$$

where

$$\overline{\mathcal{P}}\colon \mathbb{R}^{d+|\mathcal{X}|\cdot(d+d^2)} \to \mathcal{Y}^{d+|\mathcal{X}|\cdot(d+d^2)}, \quad y \mapsto (\mathcal{P}(y_i))_{i=1,\dots,d+|\mathcal{X}|\cdot(d+d^2)}.$$

Finally we require that the linear mapping $\mathcal{A}$ satisfies

$$\forall\, x \in \mathbb{R}^{|\mathcal{X}|}\colon\ \|\mathcal{A}(x)\|_{[-n^\varepsilon,n^\varepsilon]^d} \le c \cdot \|x\|, \tag{5.16}$$

additionally to the conditions 2)(i) and 2)(ii), and we define

$$\widehat{S}_{n,\varepsilon}^{\mathrm{mod}} = \mathcal{Q}_2 \circ \mathcal{Q}_1^{\mathrm{mod}} \circ \overline{\mathcal{A}} \circ \overline{\mathcal{P}} \circ \overline{\mathcal{N}}.$$

Using Lemma 5.1 and standard error estimates for the Euler scheme one can show that Lemma 5.2 still holds true with $\mathcal{Q}_1$ replaced by $\mathcal{Q}_1^{\mathrm{mod}}$, and therefore the mapping $\mathcal{Q}_2 \circ \mathcal{Q}_1^{\mathrm{mod}}$ enjoys the same properties 3)(i) and 3)(ii) as the mapping $\mathcal{Q} = \mathcal{Q}_2 \circ \mathcal{Q}_1$. Put $(\overline{x}_0, \overline{a}, \overline{b}) = \overline{\mathcal{A}}(\overline{\mathcal{P}}(\overline{\mathcal{N}}(x_0, a, b)))$ for $(x_0, a, b) \in \mathcal{C}$. Using properties 2)(i), 2)(ii) and (5.16) of $\mathcal{A}$ one can show that (5.3) as well as (5.4) hold true with $\overline{a}_i$ and $\overline{b}_{i,j}$ in place of $\widetilde{a}_i$ and $\widetilde{b}_{i,j}$, respectively. We conclude that Theorem 5.1 holds true with $\widehat{S}_{n,\varepsilon}^{\mathrm{mod}}$ in place of $\widehat{S}_{n,\varepsilon}$.

Next, we consider the total cost of $\widehat{S}_{n,\varepsilon}^{\mathrm{mod}}$. Put

$$\mathcal{V} = \mathcal{Y}^{d+|\mathcal{X}|\cdot(d+d^2)}.$$

In the real number model we may assume that all of the discrete measures

$$\mathcal{Q}_2(\mathcal{Q}_1^{\mathrm{mod}}(\overline{\mathcal{A}}(\nu))), \qquad \nu \in \mathcal{V},$$

which constitute the possible outcomes of the algorithm $\widehat{S}_{n,\varepsilon}^{\mathrm{mod}}$, are available via precomputing in the form of $|\mathcal{V}|$ stored lists of support points and corresponding probability weights. As a consequence the mapping $\mathcal{Q}_2 \circ \mathcal{Q}_1^{\mathrm{mod}} \circ \overline{\mathcal{A}}$ can be implemented in such a way that the number of basic computational operations needed to compute $\mathcal{Q}_2(\mathcal{Q}_1^{\mathrm{mod}}(\overline{\mathcal{A}}(\nu)))$ at $\nu \in \mathcal{V}$ is bounded by $c \cdot (|\mathcal{X}| + |\operatorname{supp}(\mathcal{Q}_2(\mathcal{Q}_1^{\mathrm{mod}}(\overline{\mathcal{A}}(\nu))))|)$. Furthermore, the projection $\mathcal{P} \colon \mathbb{R} \to \mathcal{Y}$ can be implemented in such a way that the number of basic computational operations needed for a single evaluation of $\overline{\mathcal{P}} \circ \overline{\mathcal{N}}$ is bounded by $c \cdot (d + |\mathcal{X}|(d+d^2))$. In total we obtain the upper bound $c \cdot (|\mathcal{X}| + |\operatorname{supp}(\mathcal{Q}_2(\mathcal{Q}_1^{\mathrm{mod}}(\overline{x}_0, \overline{a}, \overline{b})))|)$ for the number of basic computational operations needed for the evaluation of the algorithm $\widehat{S}_{n,\varepsilon}^{\mathrm{mod}}$ at the input (x_0, a, b). Using Property 3)(i) of $\mathcal{Q}_2 \circ \mathcal{Q}_1^{\mathrm{mod}}$ and (5.2) we conclude that

$$\operatorname{cost}^{\mathrm{total}}(\widehat{S}_{n,\varepsilon}^{\mathrm{mod}}) \leq c \cdot n^{d(1+\varepsilon)\cdot \max(1,s/r)}.$$

This completes the proof of the upper bound in Theorem 4.1 in the case $* = \text{total}$.

Note, however, that the precomputation step involves at least $|\mathcal{V}| \geq n^{d^2 s n^d}$ basic computational operations.

6. Proof of Theorem 4.2 and the Lower Bounds in Theorem 4.1

Clearly, Theorem 4.2 together with (3.2) implies the lower bound in Theorem 4.1 for $* = \inf$. It therefore remains to prove Theorem 4.2 and the lower bound in Theorem 4.1 for $* = \operatorname{supp}$.

We derive these estimates by relating the corresponding minimal errors to minimal errors for suitably chosen weighted integration problems on $\mathbb{R}^d$. For the case $* = \operatorname{supp}$ the weight will be given by the probability density of a d-dimensional normal distribution, and the proof is straightforward. For the proof of Theorem 4.2 we essentially follow the approach in [21]. We use a construction of fundamental solutions for parabolic initial value problems by means of the parametrix method to represent $\int_{\mathbb{R}^d} f\, dS(x_0, a, b)$ by a series of integrals of increasing dimensions for appropriate subclasses of coefficients a and b. This series is then truncated at the second term to end up with weighted integrals of components of a and b.

Fix $d, r, s, s_1, s_2 \in \mathbb{N}$, $K > 0$ and $\beta \geq 0$. In this section unspecified positive constants c and $c(\cdot)$ may only depend on the parameters eventually specified in brackets, the parameters $d, r, s, s_1, s_2, K, \beta$ listed above, as well as on parameters B, Λ, a rectangle $R \subset \mathbb{R}^d$ and a weight function w to be introduced below. In Sections 6.4 and 6.5 such constants may also depend on a function $f\colon \mathbb{R}^d \to \mathbb{R}$ to be specified. Moreover, $q \in \{r, s\}$.

6.1. **Weighted integration on $\mathbb{R}^d$.** Let $w\colon \mathbb{R}^d \to \mathbb{R}$ be continuous on $\mathbb{R}^d \setminus \{0\}$ and non-zero at some $x \in \mathbb{R}^d \setminus \{0\}$. Take a compact rectangle $R \subset \mathbb{R}^d$ with positive volume and $0 \notin R$ such that

$$\inf_{x \in R} |w(x)| > 0.$$

Let H be a class of continuous functions $h\colon \mathbb{R}^d \to \mathbb{R}$ such that $\{h \neq 0\} \subset R$. We consider the problem of computing

$$I^w(h) = \int_{\mathbb{R}^d} h(x) \cdot w(x)\, dx \tag{6.1}$$

by a deterministic algorithm based on finitely many function values of $h \in H$.

For $N \in \mathbb{N}$ and functions $\eta_i\colon \mathbb{R}^{i-1} \to \mathbb{R}^d$ for $i = 1, \ldots, N$ define $\mathcal{R}_{\eta_1,\ldots,\eta_N}\colon H \to \mathbb{R}^N$ by

$$\mathcal{R}_{\eta_1,\ldots,\eta_N}(h) = (y_1, \ldots, y_N)$$

with $y_1 = h(\eta_1)$ and

$$y_\ell = h(\eta_\ell(y_1, \ldots, y_{\ell-1}))$$

for $\ell = 2, \ldots, N$. Put

$$\mathcal{I}_N(H) = \{\eta \circ \mathcal{R}_{\eta_1,\ldots,\eta_N} \mid \eta\colon \mathbb{R}^N \to \mathbb{R},\, \eta_i\colon \mathbb{R}^{i-1} \to \mathbb{R}^d,\, i = 1, \ldots, N\}.$$

Then

$$\overline{e}_N^w(H) = \inf_{\widehat{I} \in \mathcal{I}_N(H)} \sup_{h \in H} |I^w(h) - \widehat{I}(h)|$$

is the N-th minimal error of deterministic methods for the weighted integration problem given by (6.1).

The following theorem is a particular instance of a classical result due to [1, 17]. See [17, Prop. 2.2.4.1] for a proof.

Theorem 6.1. *Let $N \in \mathbb{N}$ and $\varepsilon > 0$, and assume that there exist continuous functions*

$$h_1, \dots, h_{2N} \colon \mathbb{R}^d \to [0, +\infty)$$

such that

(a) *the sets $\{h_i \neq 0\}$ are pairwise disjoint and contained in R,*

(b) *for all $\sigma_1, \dots, \sigma_{2N} \in \{-1, 1\}$*

$$\sum_{i=1}^{2N} \sigma_i \cdot h_i \in H,$$

(c) *for every $i = 1, \dots, 2N$*

$$\int_R h_i(x)\, dx \geq \varepsilon.$$

Then

$$\overline{e}_N^w(H) \geq c \cdot N \cdot \varepsilon.$$

For every $N \in \mathbb{N}$ we take functions $h_1, \dots, h_{2N} \in \mathcal{F}^0(d, q, K)$ with the following property. Assumptions (a) and (c) with

$$\varepsilon = c / N^{(d+q)/d}$$

from Theorem 6.1 are satisfied, and additionally we have

$$\|h_i^{(\alpha)}\| \leq \frac{K}{N^{(q - \|\alpha\|_1)/d}} \tag{6.2}$$

for all $i = 1, \dots, 2N$ and $\alpha \in \mathbb{N}_0^d$ with $\|\alpha\|_1 \leq q$.

Let

$$H_N = \Big\{ \sum_{i=1}^{2N} \sigma_i \cdot h_i \mid \sigma_1, \dots, \sigma_{2N} \in \{-1, 1\} \Big\}, \tag{6.3}$$

and note that by (6.2),

$$H_N \subset \mathcal{F}^0(d, q, K).$$

Theorem 6.1 yields

$$\overline{e}_N^w(H_N) \geq c \cdot N^{-q/d}. \tag{6.4}$$

6.2. **Proof of the lower bound in Theorem 4.1 for $*$ = supp.** Let w be the density of the centered normal distribution $S(0, 0, K \cdot E_d)$. Take $R = [1, 2]^d$ and $q = r$.

Let $N \in \mathbb{N}$. Obviously $(0, 0, K \cdot E_d) \in \mathcal{C}(d, s_1, s_2, K)$ and $H_N \subset \mathcal{F}(d, r, K, \beta)$. Hence

$$e_N^{\text{supp}}(\mathcal{C}(d, s_1, s_2, K), \mathcal{F}(d, r, K, \beta)) \geq e_N^{\text{supp}}(\{(0, 0, K \cdot E_d)\}, H_N).$$

Consider an algorithm $\widehat{S} \colon \mathcal{C}(d, s_1, s_2, K) \to M_0(\mathbb{R}^d)$ that satisfies $\text{cost}^{\text{supp}}(\widehat{S}) \leq N$. Then

$$\widehat{I} \colon H_N \to \mathbb{R}, \quad h \mapsto \int_{\mathbb{R}^d} h \, d\widehat{S}(0, 0, K \cdot E_d),$$

defines an algorithm $\widehat{I} \in \mathcal{I}_N(H_N)$ and we have

$$\rho_{H_N}(S(0,0,K\cdot E_d), \widehat{S}(0,0,K\cdot E_d)) = \sup_{h\in H_N} \left| \int_{\mathbb{R}^d} h(x)\cdot w(x)\,dx - \widehat{I}(h)\right|.$$

Thus

$$e_N^{\mathrm{supp}}(\{(0,0,K\cdot E_d)\}, H_N) \geq \overline{e}_N^{w}(H_N) \geq c\cdot N^{-r/d}$$

by (6.4), which completes the proof.

6.3. **A series expansion of $\int_{\mathbb{R}^d} f\,dS(x_0,a,b)$.** In this section we fix

$$x_0 \in \mathbb{R}^d,\ a\colon \mathbb{R}^d \to \mathbb{R}^d,\ b\colon \mathbb{R}^d \to \mathbb{R}^{d\times d},$$

and we put

$$\sigma = bb^{\mathrm{T}}.$$

We furthermore fix numbers

$$B \geq 1,\ 0 < \Lambda \leq 1$$

and we assume

(A1) a is Lipschitz continuous and bounded,

(A2) b is continuously differentiable and satisfies

$$\|b\| \leq B, \quad \|b'\| = \max_{1\leq i,j\leq d} \|\nabla b_{i,j}\| < \infty,$$

(A3) the second-order differential operator associated with a and b is uniformly elliptic with

$$\inf_{x\in\mathbb{R}^d} \lambda_{\min}(\sigma(x)) \geq \Lambda.$$

Moreover, we put

$$\|b\|_* = \sup_{x,y\in\mathbb{R}^d} \|b(x)-b(y)\|.$$

It is known that under the assumptions (A1) - (A3) the distribution $S(x_0,a,b)$ has a Lebesgue-density. More precisely, let

$$V = \{(t,x,\tau,\xi) \mid 0\leq t<\tau\leq 1,\ x,\xi\in\mathbb{R}^d\}$$

and consider the partial differential equation

$$\frac{1}{2}\sum_{i,j=1}^{d} \sigma_{i,j}(x)\cdot\frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{d} a_i(x)\cdot\frac{\partial u}{\partial x_i}(t,x) + \frac{\partial u}{\partial t}(t,x) = 0 \tag{6.5}$$

for $(t,x)\in[0,1)\times\mathbb{R}^d$. If the conditions (A1) - (A3) are satisfied then the equation (6.5) has a unique fundamental solution

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

and we have

$$S(x_0, a, b)(A) = \int_A G(0, x_0, 1, \xi)\, d\xi$$

for every Borel set $A \subset \mathbb{R}^d$. See [6, Thm. 6.5.4].

We employ a series representation of G from [3, 7]. Let φ_Σ denote the Lebesgue density of the centered d-dimensional normal distribution with positive definite covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$. Define

$$Z\colon V \to \mathbb{R}, \quad (t, x, \tau, \xi) \mapsto \varphi_{(\tau-t)\sigma(\xi)}(x - \xi),$$

as well as a sequence of functions $\Phi_m\colon V \to \mathbb{R}$ by

$$\Phi_1(t, x, \tau, \xi) = \frac{1}{2}\sum_{i,j=1}^{d} (\sigma_{i,j}(x) - \sigma_{i,j}(\xi)) \cdot \frac{\partial^2}{\partial x_i\, \partial x_j} Z(t, x, \tau, \xi) + \sum_{i=1}^{d} a_i(x) \cdot \frac{\partial}{\partial x_i} Z(t, x, \tau, \xi)$$

and

$$\Phi_m(t, x, \tau, \xi) = \int_{(t,\tau)\times\mathbb{R}^d} \Phi_{m-1}(t, x, t', y) \cdot \Phi_1(t', y, \tau, \xi)\, d(t', y)$$

for $m \in \mathbb{N}$ with $m \geq 2$. Furthermore, define $G_m\colon V \to \mathbb{R}$ by

$$G_m(t, x, \tau, \xi) = \int_{(t,\tau)\times\mathbb{R}^d} Z(t, x, t', y) \cdot \Phi_m(t', y, \tau, \xi)\, d(t', y)$$

for $m \in \mathbb{N}$ and put $G_0 = Z$. Then all of the functions Φ_m and G_m are well-defined and we have

$$\sum_{m=0}^{\infty} |G_m(t, x, \tau, \xi)| < \infty$$

for all $(t, x, \tau, \xi) \in V$. Moreover, the fundamental solution of (6.5) is given by

$$G = \sum_{m=0}^{\infty} G_m. \tag{6.6}$$

See [7, Thm's. 1.7, 1.12] for a proof of these facts.

We use (6.6) to obtain a series representation of

$$\int_{\mathbb{R}^d} f\, dS(x_0, a, b) = \int_{\mathbb{R}^d} f(\xi) \cdot G(0, x_0, 1, \xi)\, d\xi \tag{6.7}$$

for $f \in \mathcal{F}(d, 1, K, \beta)$, which is used in the subsequent sections to derive the lower bounds in Theorem 4.2.

Lemma 6.1. *Let $\gamma \in (0, 1]$ and let $f \in \mathcal{F}(d, 1, K, \beta)$.*

(i) *For all* $m \in \mathbb{N}_0$

$$J_m(f) = \int_{\mathbb{R}^d} f(\xi) \cdot G_m(0, x_0, 1, \xi)\, d\xi$$

is well-defined and satisfies

$$|J_m(f)| \le (c(x_0, f, \gamma))^{m+1} (\|a\| + \|b'\|^\gamma \|b\|_*^{1-\gamma})^m m^{-m\gamma/4}.$$

(ii) *We have*

$$\int_{\mathbb{R}^d} f\, dS(x_0, a, b) = \sum_{m=0}^{\infty} J_m(f).$$

See Section 7.2 in the appendix for a proof of Lemma 6.1. In the particular case of $b = E_d$ the statement is already shown in [21].

6.4. **Proof of Theorem 4.2(i).** Fix a function $f \in \mathcal{F}(d, 1, K, \beta)$, which is not constant. For convenience we assume that

$$\frac{\partial}{\partial \xi_1} f \neq 0. \tag{6.8}$$

Put $q = s$, consider a non-trivial rectangle $R \subset \mathbb{R}^d$ and recall the definition (6.3) of the corresponding function classes $H_N \subset \mathcal{F}^0(d, s, K)$. We consider the classes of SDEs

$$\mathcal{C}_N = \{(0, (a_1, 0, \ldots, 0), K \cdot E_d) \mid a_1 \in H_N\}, \quad N \in \mathbb{N}.$$

Clearly,

$$e_N^f(\mathcal{C}^{(1)}) \ge e_N^f(\mathcal{C}_N).$$

It thus suffices to show that

$$e_N^f(\mathcal{C}_N) \ge c \cdot N^{-s/d} \tag{6.9}$$

for an appropriately chosen rectangle R.

Note that for $(x_0, a, b) \in \mathcal{C}_N$ the conditions (A1)-(A3) in Section 6.3 are satisfied with

$$B = K,\ \Lambda = \min(1, K^2).$$

In particular, we have

$$\|a\| = \|a_1\| \le K/N^{s/d},$$

see (6.2), and

$$\|b'\| = \|b\|_* = 0.$$

Moreover, the functions Z and Φ_1 are given by

$$Z(t, x, \tau, \xi) = \varphi_{K^2(\tau - t) \cdot E_d}(x - \xi) \tag{6.10}$$

and

$$\Phi_1(t, x, \tau, \xi) = a_1(x) \cdot \frac{\partial}{\partial x_1} Z(t, x, \tau, \xi)$$

for all $(t, x, \tau, \xi) \in V$, respectively, and therefore the integrals J_m depend on the drift coefficient a only via the component $a_1 \in H_N$. We write $J_m(f, a_1)$ in place of $J_m(f)$ in the sequel.

We first show that all integrals $J_m(f, a_1)$ with $m \geq 2$ are negligible, asymptotically.

Lemma 6.2. *We have*

$$\forall\, N \in \mathbb{N}\ \forall\, a_1 \in H_N\colon\ \sum_{m=2}^{\infty} |J_m(f, a_1)| \leq c \cdot N^{-2s/d}.$$

Proof. Let $N \in \mathbb{N}$ and $a_1 \in H_N$. Apply Lemma 6.1 with $\gamma = 1$ to get

$$|J_m(f, a_1)| \leq (c \cdot N^{-s/d} m^{-1/4})^m$$

for every $m \in \mathbb{N}_0$. Thus

$$\sum_{m=2}^{\infty} |J_m(f, a_1)| \leq N^{-2s/d} \cdot \sum_{m=2}^{\infty} (c \cdot m^{-1/4})^m \leq c \cdot N^{-2s/d},$$

which completes the proof. □

We turn to the analysis of $J_1(f, a_1)$. Define $w\colon \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ by

$$w(y) = \int_{(0,1)\times\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0, 0, t, y) \cdot Z(t, y, 1, \xi)\, d(t, \xi). \tag{6.11}$$

Lemma 6.3. *The function w is well-defined, continuous and non-zero. Moreover,*

$$\forall N \in \mathbb{N}\ \forall\, a_1 \in H_N\colon\ J_1(f, a_1) = \int_{\mathbb{R}^d} a_1(y) \cdot w(y)\, dy.$$

Proof. Let $y \in \mathbb{R}^d \setminus \{0\}$. Then by (6.10)

$$Z(0, 0, t, y) \leq \frac{1}{\|y\|_2^d} \cdot \frac{\|y\|_2^d}{(2K^2 t)^{d/2}} \exp\Big(-\frac{\|y\|_2^2}{2K^2 t}\Big) \leq \frac{c}{\|y\|_2^d}. \tag{6.12}$$

Hence

$$\begin{aligned}
&\int_{(0,1)\times\mathbb{R}^d} \Big|\frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0, 0, t, y) \cdot Z(t, y, 1, \xi)\Big|\, d(t, \xi) \\
&\qquad \leq \frac{c}{\|y\|_2^d} \int_{(0,1)\times\mathbb{R}^d} (1 + \|\xi\|_2^\beta) \cdot \varphi_{K^2(1-t)\cdot E_d}(y - \xi)\, d(t, \xi) \\
&\qquad = \frac{c}{\|y\|_2^d} \int_{(0,1)\times\mathbb{R}^d} (1 + \|K(1-t)^{1/2}\xi + y\|_2^\beta) \cdot \varphi_{E_d}(\xi)\, d(t, \xi) \\
&\qquad \leq c \cdot \frac{1 + \|y\|_2^\beta}{\|y\|_2^d},
\end{aligned}$$

which shows that w is well-defined.

In order to prove continuity of w note that

$$w(y) = \int_{(0,1)\times\mathbb{R}^d} \frac{\partial}{\partial\xi_1} f((1-t)^{1/2}\xi + y) \cdot Z(0,0,t,y) \cdot Z(0,0,1,\xi)\, d(t,\xi) \tag{6.13}$$

and let $y_0 \in \mathbb{R}^d \setminus \{0\}$. If $\|y-y_0\|_2 \le \|y_0\|_2/2$ then $\|y_0\|_2/2 \le \|y\|_2 \le 3\|y_0\|_2/2$ and therefore

$$\sup_{t\in(0,1)} \sup_{\|y-y_0\|_2\le\|y_0\|_2/2} Z(0,0,t,y) \le c(y_0),$$

see (6.12). Moreover,

$$\begin{aligned}
&\sup_{t\in(0,1)} \sup_{\|y-y_0\|_2\le\|y_0\|_2/2} \left|\partial/\partial\xi_1 f((1-t)^{1/2}\xi + y)\right| \\
&\qquad \le \sup_{t\in(0,1)} \sup_{\|y-y_0\|_2\le\|y_0\|_2/2} K(1 + \|(1-t)^{1/2}\xi + y\|_2^\beta) \le c(y_0)(1+\|\xi\|_2^\beta).
\end{aligned}$$

Hence

$$\begin{aligned}
&\int_{(0,1)\times\mathbb{R}^d} \sup_{\|y-y_0\|_2\le\|y_0\|_2/2} \left|\frac{\partial}{\partial\xi_1} f((1-t)^{1/2}\xi + y) \cdot Z(0,0,t,y) \cdot Z(0,0,1,\xi)\right| d(t,\xi) \\
&\qquad \le c(y_0) \int_{(0,1)\times\mathbb{R}^d} (1+\|\xi\|_2^\beta) \cdot \varphi_{K^2\cdot E_d}(\xi)\, d(t,\xi) < \infty.
\end{aligned}$$

Since the integrand in (6.13) is continuous in y we conclude by the dominated convergence theorem that w is continuous in y_0.

We proceed with showing that $w(y) \neq 0$ for some $y \neq 0$. To this end let $(k_1,\dots,k_d) \in \mathbb{N}_0^d$ and consider the function $g\colon \mathbb{R}^d \setminus \{0\} \times (0,1) \times \mathbb{R}^d \to \mathbb{R}$ given by

$$g(y,t,\xi) = \Big(\prod_{j=1}^d y_j^{k_j}\Big) \cdot \frac{\partial}{\partial\xi_1} f(\xi) \cdot Z(0,0,t,y) \cdot Z(t,y,1,\xi). \tag{6.14}$$

Put $k = k_1 + \dots + k_d$. Then

$$\begin{aligned}
&\int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} |g(y,t,\xi)|\, d(y,t,\xi) \\
&\quad \le K \int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} \|y\|_2^k \cdot (1+\|\xi\|_2^\beta) \cdot Z(0,0,t,y) \cdot Z(t,y,1,\xi)\, d(y,t,\xi) \\
&\quad = K \int_{\mathbb{R}^d\times(0,1)} \|y\|_2^k \cdot Z(0,0,t,y) \int_{\mathbb{R}^d} (1 + \|(1-t)^{1/2}\xi + y\|_2^\beta) \cdot Z(0,0,1,\xi)\, d\xi\, d(y,t) \\
&\quad \le c \int_{\mathbb{R}^d\times(0,1)} (1+\|y\|_2^{k+\beta}) \cdot Z(0,0,t,y)\, d(y,t) \le c(k).
\end{aligned}$$

Hence g is integrable and using

$$Z(0,0,t,y) \cdot Z(t,y,1,\xi) = Z(0,0,t(1-t), y - t\cdot\xi) \cdot Z(0,0,1,\xi)$$

we obtain

$$\int_{\mathbb{R}^d} \Big(\prod_{j=1}^{d} y_j^{k_j}\Big) \cdot w(y)\, dy = \int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0,0,1,\xi) \cdot \psi(\xi)\, d\xi, \tag{6.15}$$

where

$$\psi(\xi) = \int_{(0,1)\times\mathbb{R}^d} \Big(\prod_{j=1}^{d} y_j^{k_j}\Big) \cdot Z(0,0,t(1-t), y - t\xi)\, d(t,y).$$

Let q_ℓ denote the ℓ-th moment of a standard normally distributed random variable for $\ell \in \mathbb{N}_0$. We have

$$\begin{aligned}
\psi(\xi) &= \int_{(0,1)\times\mathbb{R}^d} \Big(\prod_{j=1}^{d} \big(K\sqrt{t(1-t)}y_j + t\xi_j\big)^{k_j}\Big) \cdot \varphi_{E_d}(y)\, d(t,y) \\
&= \int_{(0,1)} \prod_{j=1}^{d} \sum_{\ell=0}^{k_j} \binom{k_j}{\ell} (K\sqrt{t(1-t)})^{\ell} q_\ell (t\xi_j)^{k_j-\ell}\, dt \\
&= \sum_{\ell_1=0}^{k_1} \cdots \sum_{\ell_d=0}^{k_d} \Big(\prod_{j=1}^{d} \binom{k_j}{\ell_j} K^{l_j} q_{\ell_j} \xi_j^{k_j-\ell_j}\Big) \cdot \int_{(0,1)} \prod_{j=1}^{d} (1-t)^{\ell_j/2} t^{k_j-\ell_j/2}\, dt.
\end{aligned}$$

Hence ψ is a polynomial of degree $(k_1,\dots,k_d)$. Assume $w=0$. Then (6.15) implies that

$$\int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0,0,1,\xi) \cdot p(\xi)\, d\xi = 0$$

for every d-variate polynomial p, which in turn yields $\partial f/\partial \xi_1 = 0$ Lebesgue-almost everywhere. Since $\partial f/\partial \xi_1$ is continuous we get $\partial f/\partial \xi_1 = 0$, which is in contradiction to assumption (6.8).

Finally, let $N \in \mathbb{N}$ and $a_1 \in H_N$ and consider the mapping g given by (6.14) with $k_1 = \dots = k_d = 0$. Since a_1 is bounded and g is integrable we have

$$\begin{aligned}
\int_{\mathbb{R}^d} a_1(y) \cdot w(y)\, dy &= \int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} a_1(y) \cdot g(y,t,\xi)\, d(y,t,\xi) \\
&= \int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} a_1(y) \cdot \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0,0,t,y) \cdot Z(t,y,1,\xi)\, d(y,t,\xi).
\end{aligned}$$

By (2.2),

$$\lim_{\|\xi\|\to\infty} f(\xi) \cdot Z(t,y,1,\xi) = 0$$

for all $t \in (0,1)$, $y \in \mathbb{R}^d$ and $\xi_2,\dots,\xi_d \in \mathbb{R}$. Hence, by partial integration,

$$\int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(t,y,1,\xi)\, d\xi = \int_{\mathbb{R}^d} f(\xi) \cdot \frac{\partial}{\partial y_1} Z(t,y,1,\xi)\, d\xi$$

for all $t \in (0,1)$, $y \in \mathbb{R}^d$. It follows

$$\begin{aligned}\int_{\mathbb{R}^d} a_1(y) \cdot w(y)\, dy &= \int_{\mathbb{R}^d} f(\xi) \int_{(0,1)\times\mathbb{R}^d} a_1(y) \cdot Z(0,0,t,y) \cdot \frac{\partial}{\partial y_1} Z(t,y,1,\xi)\, d(t,y)\, d\xi \\ &= \int_{\mathbb{R}^d} f(\xi) \cdot G_1(0,0,1,\xi)\, d\xi = J_1(f,a_1),\end{aligned}$$

which finishes the proof. □

By Lemma 6.3 there exists a rectangle $R \subset \mathbb{R}^d$ such that $\inf_{x\in R} |w(x)| > 0$. We turn to the proof of the corresponding claim (6.9). Let $N \in \mathbb{N}$ and consider an algorithm $\widehat{A} \in \mathfrak{A}$. Let $a_1 \in H_N$. By Lemma 6.1(ii), Lemma 6.2 and Lemma 6.3,

$$\begin{aligned}&\big|\widehat{A}(0,(a_1,0,\dots,0),K\cdot E_d) - A^f(0,(a_1,0,\dots,0),K\cdot E_d)\big| \\ &\quad= \Big|\widehat{A}(0,(a_1,0,\dots,0),K\cdot E_d) - \sum_{m=0}^{\infty} J_m(f,a_1)\Big| \\ &\quad\geq \Big|\widehat{A}(0,(a_1,0,\dots,0),K\cdot E_d) - \sum_{m=0}^{1} J_m(f,a_1)\Big| - c\cdot N^{-2s/d} \\ &\quad= \Big|\widehat{A}(0,(a_1,0,\dots,0),K\cdot E_d) - \int_{\mathbb{R}^d} f(\xi)\cdot Z(0,0,1,\xi)\, d\xi \\ &\qquad\qquad - \int_{\mathbb{R}^d} a_1(y)\cdot w(y)\, dy\Big| - c\cdot N^{-2s/d}.\end{aligned}$$

If $\widehat{A}$ satisfies $\mathrm{cost}(\widehat{A}) \leq N$ then

$$\widehat{I} \colon H_N \to \mathbb{R}, \quad h \mapsto \widehat{A}(0,(h,0,\dots,0),K\cdot E_d) - \int_{\mathbb{R}^d} f(\xi)\cdot Z(0,0,1,\xi)\, d\xi$$

defines an algorithm $\widehat{I} \in \mathcal{I}_N(H_N)$ and we may conclude that

$$\sup_{(x_0,a,b)\in\mathcal{C}_N} |\widehat{A}(x_0,a,b) - A^f(x_0,a,b)| \geq \sup_{h\in H_N} |\widehat{I}(h) - I^w(h)| - c\cdot N^{-2s/d}.$$

Consequently,

$$e_N^f(\mathcal{C}_N) \geq \overline{e}_N^{\,w}(H_N) - c\cdot N^{-2s/d}.$$

It remains to apply (6.4) to complete the proof of (6.9).

6.5. **Proof of Theorem 4.2(ii).** Fix a function $f \in \mathcal{F}(d,2,K,\beta)$, which is not multilinear. For convenience we assume that

$$\frac{\partial^2}{\partial \xi_1^2} f \neq 0. \tag{6.16}$$

Put $q = s$, consider a non-trivial rectangle $R \subset \mathbb{R}^d$, recall the definition (6.3) of the corresponding function classes H_N and let

$$\widetilde{H}_N = \{K/2 + h/4 \mid h \in H_N\} \subset \mathcal{F}^0(d, s, K).$$

We consider the classes of SDEs

$$\mathcal{C}_N = \big\{(0, 0, \operatorname{diag}(b_1, K/2, \dots, K/2)) \mid b_1 \in \widetilde{H}_N\big\}, \qquad N \in \mathbb{N}.$$

Clearly,

$$e_N^f(\mathcal{C}^{(2)}) \geq e_N^f(\mathcal{C}_N).$$

It thus suffices to show that

$$e_N^f(\mathcal{C}_N) \geq c \cdot N^{-s/d} \tag{6.17}$$

for an appropriately chosen rectangle R.

Note that for $(x_0, a, b) \in \mathcal{C}_N$ the conditions (A1)-(A3) in Section 6.3 are satisfied with

$$B = K,\ \Lambda = \min(1, K^2/16).$$

In particular, we have

$$\|b'\| = \|b_1'\| \leq K \cdot N^{-(s-1)/d}, \quad \|b\|_* \leq K \cdot N^{-s/d},$$

see (6.2). Moreover, the functions Z and Φ_1 are given by

$$\begin{aligned} Z(t, x, \tau, \xi) = & \frac{1}{(2\pi(\tau - t))^{d/2} (K/2)^{d-1} \cdot b_1(\xi)} \\ & \cdot \exp\Big(-\frac{1}{2(\tau - t)}\Big(\frac{(x_1 - \xi_1)^2}{b_1^2(\xi)} + \sum_{i=2}^d \frac{(x_i - \xi_i)^2}{K^2/4}\Big)\Big) \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} \Phi_1(t, x, \tau, \xi) &= \frac{b_1^2(x) - b_1^2(\xi)}{2} \cdot \frac{\partial^2}{\partial x_1^2} Z(t, x, \tau, \xi) \\ &= \frac{b_1^2(x) - b_1^2(\xi)}{2(\tau - t) \cdot b_1^2(\xi)} \cdot \Big(\frac{(\xi_1 - x_1)^2}{(\tau - t) \cdot b_1^2(\xi)} - 1\Big) \cdot Z(t, x, \tau, \xi) \end{aligned} \tag{6.19}$$

for all $(t, x, \tau, \xi) \in V$, respectively, and therefore the integrals $J_m(f)$ depend on the diffusion coefficient b only via the component $b_1 \in \widetilde{H}_N$. We write $J_m(f, b_1)$ in place of $J_m(f)$ in the sequel.

Similarly to Lemma 6.2 we show that all integrals $J_m(f, b_1)$ with $m \geq 2$ are neglegible, asymptotically.

Lemma 6.4. *We have*

$$\forall\, \gamma \in (0, 1]\ \forall\, N \in \mathbb{N}\ \forall\, b_1 \in \widetilde{H}_N \colon \sum_{m=2}^{\infty} |J_m(f, b_1)| \leq c(\gamma) \cdot N^{-2(s-\gamma)/d}.$$

Proof. Let $\gamma \in (0,1]$, $N \in \mathbb{N}$ and $b_1 \in \widetilde{H}_N$. Apply Lemma 6.1 to get

$$|J_m(f,b_1)| \le (c(\gamma) \cdot N^{-(s-\gamma)/d} m^{-\gamma/4})^m$$

for every $m \in \mathbb{N}_0$. Thus

$$\sum_{m=2}^{\infty} |J_m(f,b_1)| \le N^{-2(s-\gamma)/d} \cdot \sum_{m=2}^{\infty} (c(\gamma) \cdot m^{-\gamma/4})^m \le c(\gamma) \cdot N^{-2(s-\gamma)/d},$$

which completes the proof. □

We proceed with the analysis of $J_0(f,b_1)$. Define

$$T_0 \colon \mathbb{R}^d \times [K/4,K] \to \mathbb{R}, \quad (\xi,v) \mapsto \frac{1}{(2\pi)^{d/2}(K/2)^{d-1} \cdot v} \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi_1^2}{v^2} + \sum_{i=2}^{d} \frac{\xi_i^2}{K^2/4}\right)\right).$$

Thus $Z(0,0,1,\xi) = T_0(\xi, b_1(\xi))$ and we have

$$J_0(f,b_1) = \int_{\mathbb{R}^d} f(\xi) \cdot T_0(\xi, b_1(\xi))\, d\xi$$

for every $b_1 \in \widetilde{H}_N$. Put

$$J_{0,0}(f) = \int_{\mathbb{R}^d} f(\xi) \cdot T_0(\xi, K/2)\, d\xi,$$
$$J_{0,1}(f,b_1) = \int_{\mathbb{R}^d} f(\xi) \cdot \frac{\partial}{\partial v} T_0(\xi, K/2) \cdot (b_1(\xi) - K/2)\, d\xi$$

for $b_1 \in \widetilde{H}_N$ and note that due to (2.2) both integrals are well-defined.

Lemma 6.5. *We have*

$$\forall\, N \in \mathbb{N}\ \forall\, b_1 \in \widetilde{H}_N \colon\ |J_0(f,b_1) - J_{0,0}(f) - J_{0,1}(f,b_1)| \le c \cdot N^{-2s/d}.$$

Proof. Straightforward calculations yield

$$\frac{\partial}{\partial v} T_0(\xi,v) = \frac{1}{(2\pi)^{d/2}(K/2)^{d-1}} \cdot \left(\frac{\xi_1^2}{v^4} - \frac{1}{v^2}\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi_1^2}{v^2} + \sum_{i=2}^{d} \frac{\xi_i^2}{K^2/4}\right)\right),$$
$$\frac{\partial^2}{\partial v^2} T_0(\xi,v) = \frac{1}{(2\pi)^{d/2}(K/2)^{d-1}} \cdot \left(\frac{\xi_1^4}{v^7} - \frac{5\xi_1^2}{v^5} + \frac{2}{v^3}\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\xi_1^2}{v^2} + \sum_{i=2}^{d} \frac{\xi_i^2}{K^2/4}\right)\right).$$

Hence

$$\sup_{v \in [K/4,K]} \left(\left|\frac{\partial}{\partial v} T_0(\xi,v)\right| + \left|\frac{\partial^2}{\partial v^2} T_0(\xi,v)\right|\right) \le c \cdot (1 + \|\xi\|_2^4) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2}\right) \tag{6.20}$$

for all $\xi \in \mathbb{R}^d$. Observing (2.2) as well as $4b_1 - 2K \in H_N$ it follows

$$
\begin{aligned}
&|J_0(f,b_1) - J_{0,0}(f) - J_{0,1}(f,b_1)| \\
&\qquad = \left| \int_{\mathbb{R}^d} f(\xi) \int_{K/2}^{b_1(\xi)} (b_1(\xi) - v) \cdot \frac{\partial^2}{\partial v^2} T_0(\xi, v)\, dv\, d\xi \right| \\
&\qquad \le c \cdot \int_{\mathbb{R}^d} |f(\xi)| \cdot (1 + \|\xi\|_2^4) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2} \right) (b_1(\xi) - K/2)^2\, d\xi \\
&\qquad \le c \cdot \|b_1 - K/2\|^2 \cdot \int_{\mathbb{R}^d} (1 + \|\xi\|_2^{5+\beta}) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2} \right) d\xi \\
&\qquad \le c \cdot N^{-2s/d},
\end{aligned}
$$

which completes the proof. □

Next we consider the term $J_1(f, b_1)$. Define

$$T_1\colon (0,1) \times \mathbb{R}^d \times \mathbb{R}^d \times [K/4, K] \times [K/4, K] \to \mathbb{R}$$

by

$$
\begin{aligned}
&T_1(t, y, \xi, u, v) \\
&\quad = \frac{1}{2(2\pi)^d (K/2)^{2(d-1)} \cdot (t(1-t))^{d/2}} \cdot \left(\frac{1}{u} + \frac{1}{v} \right) \cdot \left(\frac{(\xi_1 - y_1)^2}{(1-t)^2 \cdot v^4} - \frac{1}{(1-t) \cdot v^2} \right) \\
&\qquad \cdot \exp\left(-\frac{1}{2t} \left(\frac{y_1^2}{u^2} + \sum_{i=2}^d \frac{y_i^2}{K^2/4} \right) - \frac{1}{2(1-t)} \left(\frac{(\xi_1 - y_1)^2}{v^2} + \sum_{i=2}^d \frac{(\xi_i - y_i)^2}{K^2/4} \right) \right).
\end{aligned}
$$

Then

$$Z(0,0,t,y) \cdot \Phi_1(t, y, 1, \xi) = (b_1(y) - b_1(\xi)) \cdot T_1(t, y, \xi, b_1(y), b_1(\xi)),$$

see (6.18) and (6.19), and by (7.8) we have

$$J_1(f, b_1) = \int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} f(\xi) \cdot (b_1(y) - b_1(\xi)) \cdot T_1(t, y, \xi, b_1(y), b_1(\xi))\, d(\xi, t, y)$$

for all $b_1 \in \widetilde{H}_N$.

Lemma 6.6. *We have*

$$\int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} \big| f(\xi) | \cdot \|\xi - y\|_2 \cdot |T_1(t, y, \xi, K/2, K/2) \big|\, d(\xi, t, y) \le c.$$

Proof. By definition,

$$
\begin{aligned}
T_1(t, y, \xi, K/2, K/2) &= \frac{1}{(K/2) \cdot (1-t)} \left(\frac{(\xi_1 - y_1)^2}{(K/2)^2 \cdot (1-t)} - 1 \right) \\
&\qquad \cdot \varphi_{(K/2)^2(1-t) \cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y)
\end{aligned}
$$

and therefore

$$\big|T_1(t,y,\xi,K/2,K/2)\big| \le \frac{c}{t^{d/2}(1-t)^{d/2+1}} \cdot \left(\frac{\|\xi-y\|_2^2}{1-t}+1\right) \cdot \exp\left(-\frac{\|y\|_2^2}{K^2/2\cdot t} - \frac{\|\xi-y\|_2^2}{K^2/2\cdot(1-t)}\right) \tag{6.21}$$

for all $y,\xi \in \mathbb{R}^d$ and $t \in (0,1)$. Using [5, p.15, Lemma 3] we conclude

$$\begin{aligned}
&\int_{(0,1)\times\mathbb{R}^d} \|\xi-y\|_2 \cdot \big|T_1(t,y,\xi,K/2,K/2)\big|\, d(t,y) \\
&\qquad \le c\cdot \int_{(0,1)\times\mathbb{R}^d} \frac{1}{t^{d/2}(1-t)^{d/2+1/2}} \cdot \exp\left(-\frac{\|y\|_2^2}{K^2 t} - \frac{\|\xi-y\|_2^2}{K^2(1-t)}\right) d(t,y) \\
&\qquad = c\cdot \exp\left(-\frac{\|\xi\|_2^2}{K^2}\right)
\end{aligned}$$

for all $\xi \in \mathbb{R}^d$, and it remains to observe (2.2) to complete the proof. □

For $b_1 \in \widetilde{H}_N$ we have $|b_1(y)-b_1(\xi)| \le K\|\xi-y\|_2$. Hence, by Lemma 6.6, the integral

$$J_{1,0}(f,b_1) = \int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} f(\xi)\cdot(b_1(y)-b_1(\xi))\cdot T_1(t,y,\xi,K/2,K/2)\, d(\xi,t,y)$$

is well-defined.

Lemma 6.7. *We have*

$$\forall\, N \in \mathbb{N}\ \forall\, b_1 \in \widetilde{H}_N\colon\ \ |J_1(f,b_1) - J_{1,0}(f,b_1)| \le c\cdot N^{-(2s-1/2)/d}.$$

Proof. Straightforward calculations yield

$$\begin{aligned}
&\left|\frac{\partial}{\partial u} T_1(t,y,\xi,u,v)\right| \\
&\qquad \le \frac{c}{t^{d/2}(1-t)^{d/2+1}} \cdot \left(\frac{\|y\|_2^2}{t}+1\right) \cdot \left(\frac{\|\xi-y\|_2^2}{1-t}+1\right) \cdot \exp\left(-\frac{\|y\|_2^2}{2K^2 t} - \frac{\|\xi-y\|_2^2}{2K^2(1-t)}\right)
\end{aligned}$$

and

$$\left|\frac{\partial}{\partial v} T_1(t,y,\xi,u,v)\right| \le \frac{c}{t^{d/2}(1-t)^{d/2+1}} \cdot \left(\frac{\|\xi-y\|_2^4}{(1-t)^2}+1\right) \cdot \exp\left(-\frac{\|y\|_2^2}{2K^2 t} - \frac{\|\xi-y\|_2^2}{2K^2(1-t)}\right)$$

for all $y, \xi \in \mathbb{R}^d$, $t \in (0,1)$ and $u, v \in [K/4, K]$. Observing $4b_1 - 2K \in H_N$ we obtain

$$\begin{aligned}
&|b_1(y) - b_1(\xi)| \cdot |T_1(t, y, \xi, b_1(y), b_1(\xi)) - T_1(t, y, \xi, K/2, K/2)| \\
&\qquad \leq c \cdot N^{-(2s-1)/(2d)} \cdot \frac{\|\xi - y\|_2^{1/2}}{t^{d/2}(1-t)^{d/2+1}} \cdot \left(\frac{\|y\|_2^2}{t} + 1\right)\left(\frac{\|\xi - y\|_2^4}{(1-t)^2} + 1\right) \\
&\qquad\qquad \cdot \exp\left(-\frac{\|y\|_2^2}{2K^2 t} - \frac{\|\xi - y\|_2^2}{2K^2(1-t)}\right) \cdot \big(|b_1(\xi) - K/2| + |b_1(y) - K/2|\big) \\
&\qquad \leq c \cdot N^{-(2s-1/2)/d} \cdot \frac{1}{t^{d/2}(1-t)^{d/2+3/4}} \cdot \exp\left(-\frac{\|y\|_2^2}{4K^2 t} - \frac{\|\xi - y\|_2^2}{4K^2(1-t)}\right)
\end{aligned}$$

for all $y, \xi \in \mathbb{R}^d$ and $t \in (0,1)$. Using [5, p.15, Lemma 3] we conclude that

$$\begin{aligned}
&\int_{(0,1)\times\mathbb{R}^d} |b_1(y) - b_1(\xi)| \cdot |T_1(t, y, \xi, b_1(y), b_1(\xi)) - T_1(t, y, \xi, K/2, K/2)| \, d(t, y) \\
&\qquad \leq c \cdot N^{-(2s-1/2)/d} \cdot \exp\left(-\frac{\|\xi\|_2^2}{4K^2}\right)
\end{aligned}$$

for every $\xi \in \mathbb{R}^d$. Hence

$$|J_1(f, b_1) - J_{1,0}(f, b_1)| \leq c \cdot N^{-(2s-1/2)/d} \cdot \int_{\mathbb{R}^d} |f(\xi)| \cdot \exp\left(-\frac{\|\xi\|_2^2}{4K^2}\right) d\xi$$

and it remains to observe (2.2) to complete the proof. □

Define $w \colon \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ by

$$w(y) = \frac{K}{2} \int_{(0,1)\times\mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2(1-t)\cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) \, d(t, \xi). \tag{6.22}$$

Lemma 6.8. *The mapping w is well-defined, continuous and non-zero and satisfies*

$$J_{0,1}(f, b_1) + J_{1,0}(f, b_1) = \int_{\mathbb{R}^d} (b_1(y) - K/2) \cdot w(y) \, dy \tag{6.23}$$

for all $b_1 \in \widetilde{H}_N$ and $N \in \mathbb{N}$.

Proof. We have

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left|\frac{\partial^2}{\partial \xi_1^2} f(\xi)\right| \cdot \varphi_{(K/2)^2(1-t)\cdot E_d}(\xi - y) \, d\xi \\
&\qquad \leq K \int_{\mathbb{R}^d} (1 + \|(K/2)(1-t)^{1/2}\xi + y\|_2^\beta) \cdot \varphi_{E_d}(\xi) \, d\xi \leq c \cdot (1 + \|y\|_2^\beta)
\end{aligned} \tag{6.24}$$

for all $t \in (0,1)$ and $y \in \mathbb{R}^d$. Similar to (6.12),

$$\varphi_{(K/2)^2 t \cdot E_d}(y) \leq \frac{c}{\|y\|_2^d} \tag{6.25}$$

for all $t \in (0,1)$ and $y \in \mathbb{R}^d \setminus \{0\}$. Hence w is well-defined.

For the proof of the continuity of w we use the representation

$$w(y) = \frac{K}{2} \int_{(0,1)\times\mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(K/2(1-t)^{1/2}\xi + y) \cdot \varphi_{E_d}(\xi) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y)\, d(t,\xi). \tag{6.26}$$

Let $y_0 \in \mathbb{R}^d \setminus \{0\}$. By (6.25),

$$\begin{aligned}
\sup_{\|y-y_0\|_2 \le \|y_0\|_2/2} & \left| \frac{\partial^2}{\partial \xi_1^2} f(K/2(1-t)^{1/2}\xi + y) \right| \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) \\
&\le c \cdot \sup_{\|y-y_0\|_2 \le \|y_0\|_2/2} (1 + \|(K/2)(1-t)^{1/2}\xi + y\|_2^\beta) \cdot \|y\|_2^{-d} \\
&\le c \cdot (1 + \|y_0\|_2^\beta + \|\xi\|_2^\beta) \cdot \|y_0\|_2^{-d}
\end{aligned}$$

for all $t \in (0,1)$ and $\xi \in \mathbb{R}^d$ and, consequently,

$$\begin{aligned}
\int_{(0,1)\times\mathbb{R}^d} \sup_{\|y-y_0\|_2 \le \|y_0\|_2/2} & \left| \frac{\partial^2}{\partial \xi_1^2} f(K/2(1-t)^{1/2}\xi + y) \right| \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) \cdot \varphi_{E_d}(\xi)\, d(t,\xi) \\
&\le c(y_0) \int_{(0,1)\times\mathbb{R}^d} (1 + \|\xi\|_2^\beta) \cdot \varphi_{E_d}(\xi)\, d(t,\xi) \\
&\le c(y_0).
\end{aligned}$$

Since the integrand in (6.26) is continuous in y we conclude by the dominated convergence theorem that w is continuous in y_0.

Let $(k_1, \dots, k_d) \in \mathbb{N}_0^d$ and define $g\colon \mathbb{R}^d \setminus \{0\} \times (0,1) \times \mathbb{R}^d \to \mathbb{R}$ by

$$g(y,t,\xi) = \Big(\prod_{j=1}^d y_j^{k_j}\Big) \cdot \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2(1-t)\cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y). \tag{6.27}$$

Put $k = k_1 + \dots + k_d$. Observe (6.24) to get

$$\begin{aligned}
\int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} & |g(y,t,\xi)|\, d(y,t,\xi) \\
&\le c \int_{\mathbb{R}^d\times(0,1)} \|y\|_2^k \cdot (1 + \|y\|_2^\beta) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y)\, d(y,t) \\
&\le c \int_{\mathbb{R}^d\times(0,1)} (1 + \|(K/2)t^{1/2} y\|_2^{\beta+k}) \cdot \varphi_{E_d}(y)\, d(y,t) \\
&\le c(k).
\end{aligned}$$

Hence g is integrable. Using

$$\varphi_{(K/2)^2(1-t)\cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) = \varphi_{(K/2)^2 t(1-t)\cdot E_d}(y - t\xi) \cdot \varphi_{(K/2)^2 \cdot E_d}(\xi) \tag{6.28}$$

we obtain

$$\int_{\mathbb{R}^d} \Big(\prod_{j=1}^{d} y_j^{k_j}\Big) \cdot w(y)\, dy = \frac{K}{2} \int_{\mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2 \cdot E_d}(\xi) \cdot \psi(\xi)\, d\xi,$$

where

$$\psi(\xi) = \int_{(0,1)\times\mathbb{R}^d} \Big(\prod_{j=1}^{d} y_j^{k_j}\Big)\, \varphi_{(K/2)^2 t(1-t)\cdot E_d}(y - t\xi)\, d(t,y).$$

The mapping ψ is a polynomial of degree $(k_1,\dots,k_d)$, as shown in the proof of Lemma 6.3. Hence $w = 0$ implies that

$$\int_{\mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2 \cdot E_d}(\xi) \cdot p(\xi)\, d\xi = 0$$

for every d-variate polynomial p, which in turn yields $\partial^2 f/\partial \xi_1^2 = 0$, in contradiction to assumption (6.16).

It remains to establish (6.23). Let $b_1 \in \widetilde{H}_N$ and consider the mapping g given by (6.27) with $k_1 = \cdots = k_d = 0$. Since b_1 is bounded and g is integrable we have

$$\begin{aligned}\int_{\mathbb{R}^d} (b_1(y) - K/2) \cdot w(y)\, dy &= \int_{\mathbb{R}^d\times(0,1)\times\mathbb{R}^d} (b_1(y) - K/2)\cdot(K/2)\cdot g(y,t,\xi)\, d(y,t,\xi)\\ &= \lim_{\eta\to 1} \int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(y) - K/2)\cdot(K/2)\cdot g(y,t,\xi)\, d(y,t,\xi).\end{aligned}$$

Let $\eta \in (0,1)$. Since $f \in \mathcal{F}(d,2,K,\beta)$ we have

$$\lim_{\|\xi\|\to\infty} (|f(\xi)| + |\partial f(\xi)/\partial \xi_1|)\cdot \|\xi - y\|^m \cdot \exp\Big(-\frac{\|\xi - y\|_2^2}{2(K/2)^2\cdot(1-t)}\Big) = 0$$

for every $y \in \mathbb{R}^d$, every $t \in (0,\eta)$ and every $m \in \mathbb{N}$. Note that the mapping

$$(y,t,\xi) \to f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)$$

is integrable on $\mathbb{R}^d \times (0,\eta) \times \mathbb{R}^d$. Hence, by partial integration,

$$\begin{aligned}&\int_{\mathbb{R}^d} (K/2)\cdot g(y,t,\xi)\, d\xi\\ &\quad= \int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot \frac{\xi_1 - y_1}{(K/2)(1-t)} \cdot \varphi_{(K/2)^2(1-t)\cdot E_d}(\xi - y)\cdot \varphi_{(K/2)^2 t\cdot E_d}(y)\, d\xi\\ &\quad= \int_{\mathbb{R}^d} f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)\, d\xi\end{aligned}$$

for every $y \in \mathbb{R}^d \setminus \{0\}$ and every $t \in (0,\eta)$ and we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(y)-K/2)\cdot(K/2)\cdot g(y,t,\xi)\, d(y,t,\xi)\\
&\qquad = \int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(y)-K/2)\cdot f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)\ d(y,t,\xi)\\
&\qquad = \int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(y)-b_1(\xi))\cdot f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)\, d(y,t,\xi)\\
&\qquad\qquad + \int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(\xi)-K/2)\cdot f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)\, d(y,t,\xi).
\end{aligned}$$

By Lemma 6.6,

$$\lim_{\eta\to 1}\int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} (b_1(y)-b_1(\xi))\cdot f(\xi)\cdot T_1(t,y,\xi,K/2,K/2)\, d(y,t,\xi) = J_{1,0}(f,b_1).$$

It therefore remains to show that

$$\lim_{\eta\to 1}\int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} h_\eta(y,t,\xi)\, d(y,t,\xi) = J_{0,1}(f,b_1), \tag{6.29}$$

where $h_\eta\colon \mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d\to\mathbb{R}$ is given by

$$h_\eta(y,t,\xi) = (b_1(\xi)-K/2)\cdot f(\xi)\cdot T_1(t,y,\xi,K/2,K/2).$$

By (2.2),

$$\begin{aligned}
|h_\eta(y,t,\xi)| \le c(\eta)\cdot(1+\|y\|_2^{1+\beta}+\|\xi-y\|_2^{3+\beta})&\\
\cdot\, \varphi_{(K/2)^2(1-t)\cdot E_d}(\xi-y)\cdot\varphi_{(K/2)^2 t\cdot E_d}(y)&
\end{aligned}$$

for all $y,\xi\in\mathbb{R}^d$ and $t\in(0,\eta)$. Hence h_η is integrable. Observing (6.28) we conclude that

$$\begin{aligned}
&\int_{\mathbb{R}^d\times(0,\eta)\times\mathbb{R}^d} h_\eta(y,t,\xi)\, d(y,t,\xi)\\
&\quad = \int_{\mathbb{R}^d}(b_1(\xi)-K/2)\cdot f(\xi)\cdot\varphi_{(K/2)^2\cdot E_d}(\xi)\int_{(0,\eta)}\frac{1}{(K/2)(1-t)}\\
&\qquad\qquad \cdot\int_{\mathbb{R}^d}\left(\frac{(\xi_1-y_1)^2}{(K/2)^2\cdot(1-t)}-1\right)\varphi_{(K/2)^2 t(1-t)\cdot E_d}(y-t\xi)\, dy\, dt\, d\xi\\
&\quad = \int_{\mathbb{R}^d}(b_1(\xi)-K/2)\cdot f(\xi)\cdot\varphi_{(K/2)^2\cdot E_d}(\xi)\int_{(0,\eta)}\frac{1}{K/2}\cdot\left(\frac{\xi_1^2}{(K/2)^2}-1\right)dt\, d\xi\\
&\quad = \eta\int_{\mathbb{R}^d}(b_1(\xi)-K/2)\cdot f(\xi)\cdot\frac{\partial}{\partial v}T_0(\xi,K/2)\, d\xi = \eta\cdot J_{0,1}(f,b_1),
\end{aligned}$$

which yields (6.29) and hereby completes the proof. $\square$

By Lemma 6.8 there exists a rectangle $R \subset \mathbb{R}^d$ such that $\inf_{x\in R} |w(x)| > 0$. We turn to the proof of the corresponding claim (6.17). Let $N \in \mathbb{N}$ and consider an algorithm $\widehat{A} \in \mathfrak{A}$. Let $b_1 \in \widetilde{H}_N$. By Lemma 6.1(ii), Lemma 6.4 with $\gamma = 1/4$, Lemmas 6.5, 6.7 and by (6.23) we get

$$
\begin{aligned}
&\big|\widehat{A}(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2)) - A^f(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2))\big|\\
&\quad = \Big|\widehat{A}(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2)) - \sum_{m=0}^{\infty} J_m(f,b_1)\Big|\\
&\quad \geq \Big|\widehat{A}(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2)) - \sum_{m=0}^{1} J_m(f,b_1)\Big| - \frac{c}{N^{3s/(2d)}}\\
&\quad \geq \Big|\widehat{A}(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2)) - J_{0,0}(f) - J_{0,1}(f,b_1) - J_{1,0}(f,b_1)\Big|\\
&\qquad\qquad - \frac{c}{N^{-3s/(2d)}}\\
&\quad = \Big|\widehat{A}(0,0,\mathrm{diag}(b_1,K/2,\dots,K/2)) - J_{0,0}(f) - \int_{\mathbb{R}^d} (b_1(y)-K/2)\cdot w(y)\,dy\Big|\\
&\qquad\qquad - \frac{c}{N^{-3s/(2d)}}.
\end{aligned}
$$

If $\widehat{A}$ satisfies $\mathrm{cost}(\widehat{A}) \leq N$ then

$$
\widehat{I}\colon H_N \to \mathbb{R}, \quad h \mapsto \widehat{A}(0,0,\mathrm{diag}(K/2+h/4,K/2,\dots,K/2)) - J_{0,0}(f)
$$

defines an algorithm $\widehat{I} \in \mathcal{I}_N(H_N)$ and we obtain

$$
\sup_{(x_0,a,b)\in\mathcal{C}_N} |\widehat{A}(x_0,a,b) - A^f(x_0,a,b)| \geq \sup_{h\in H_N} |\widehat{I}(h) - I^{w/4}(h)| - c\cdot N^{-3s/(2d)}.
$$

Consequently,

$$
e_N^f(\mathcal{C}_N) \geq \overline{e}_N^{\,w/4}(H_N) - c\cdot N^{-3s/(2d)},
$$

and (6.17) follows by (6.4).

7. Appendix

7.1. **A comparison result for SDEs.** We study the dependence in p-th mean of the solution of an SDE on the initial value and the coefficients of the equation.

Lemma 7.1. *Let $\kappa > 0$ and*

$$
(x_0,a,b), (\widetilde{x}_0,\widetilde{a},\widetilde{b}) \in [-\kappa,\kappa]^d \times (\mathcal{L}^0(d,\kappa))^d \times (\mathcal{L}^0(d,\kappa))^{d^2},
$$

and let X and $\widetilde{X}$ denote the strong solutions of the respective SDEs driven by W. Then

$$\mathbb{E}\big(\sup_{t\in[0,1]} \|X(t)-\widetilde{X}(t)\|^p\big)^{1/p} \le c(d,p,q,\kappa)\cdot(1+\|x_0\|^{q+1}+\|\widetilde{x}_0\|^{q+1})\cdot(\|x_0-\widetilde{x}_0\| + \|a-\widetilde{a}\|_{[-u,u]^d} + \|b-\widetilde{b}\|_{[-u,u]^d} + u^{-q})$$

for all $p,q\ge 1$ and all $u>0$.

We believe that Lemma 7.1 or a similar result is already known but could not find it in the literature. For the convenience of the reader we therefore provide a proof of Lemma 7.1.

Proof. Fix $\kappa>0$, $p\ge 2$ and $q\ge 1$. Unspecified positive constants c and $c(\cdot)$ may only depend on the parameters eventually specified in brackets, on the dimension d and on the parameters κ, p and q. Let $(x_0,a,b),(\widetilde{x}_0,\widetilde{a},\widetilde{b})\in[-\kappa,\kappa]^d\times(\mathcal{L}^0(d,\kappa))^d\times(\mathcal{L}^0(d,\kappa))^{d^2}$. Put

$$Z^*(t)=\sup_{0\le v\le t}\|Z(v)\|$$

for $t\in[0,1]$ and any $\mathbb{R}^d$-valued stochastic process $Z=(Z(v))_{v\in[0,1]}$ with continuous paths. Fix $u>0$ and let $\widetilde{q}\ge 1$. By the properties of a, $\widetilde{a}$, b, $\widetilde{b}$,

$$\mathbb{E}(X^*(1))^{\widetilde{q}}+\mathbb{E}(\widetilde{X}^*(1))^{\widetilde{q}}\le c(\widetilde{q})\cdot(1+\|x_0\|^{\widetilde{q}}+\|\widetilde{x}_0\|^{\widetilde{q}}). \tag{7.1}$$

Consider the stopping time

$$\tau=\inf\{t\in[0,1]\mid \max(\|X(t)\|,\|\widetilde{X}(t)\|)\ge u\}.$$

By the Markov inequality and (7.1) with $\widetilde{q}=2pq$ we have

$$\mathbb{P}(\tau\le 1)\le\mathbb{P}(X^*(1)\ge u)+\mathbb{P}(\widetilde{X}^*(1)\ge u)\le c\cdot(1+\|x_0\|^{2pq}+\|\widetilde{x}_0\|^{2pq})\cdot u^{-2pq}.$$

Put $Y=X-\widetilde{X}$ and use (7.1) with $\widetilde{q}=2p$ to obtain

$$\begin{aligned}\big(\mathbb{E}\big((Y^*(1))^p\cdot 1_{\{\tau\le 1\}}\big)\big)^{1/p} &\le \big(\mathbb{E}(Y^*(1))^{2p}\big)^{1/(2p)}\cdot\big(\mathbb{P}(\tau\le 1)\big)^{1/(2p)}\\ &\le c\cdot(1+\|x_0\|^{q+1}+\|\widetilde{x}_0\|^{q+1})\cdot u^{-q}.\end{aligned}$$

Note that

$$Y^*(1)\cdot 1_{\{\tau>1\}}\le Y^*(\tau\wedge 1).$$

Therefore the proof of the lemma is complete once we have shown that

$$\mathbb{E}(Y^*(\tau\wedge 1))^p\le c\cdot(\|x_0-\widetilde{x}\|^p+\|a-\widetilde{a}\|^p_{[-u,u]^d}+\|b-\widetilde{b}\|^p_{[-u,u]^d}). \tag{7.2}$$

Clearly,

$$\begin{aligned}Y(\tau\wedge t)=x_0-\widetilde{x}_0&+\int_0^{\tau\wedge t}(a(X(v))-\widetilde{a}(\widetilde{X}(v)))\,dv\\ &+\int_0^{\tau\wedge t}(b(X(v))-\widetilde{b}(\widetilde{X}(v)))\,dW(v)\end{aligned}$$

for every $t \in [0,1]$. By the Hölder inequality and the Lipschitz continuity of a and $\widetilde{a}$ we obtain

$$\begin{aligned}&\Big\|\int_0^{\tau\wedge t}(a(X(v))-\widetilde{a}(\widetilde{X}(v)))\,dv\Big\|^p\\&\qquad\le(\tau\wedge t)^{p-1}\cdot\int_0^{\tau\wedge t}\|a(X(v))-\widetilde{a}(\widetilde{X}(v))\|^p\,dv\\&\qquad\le c\cdot\Big(\|a-\widetilde{a}\|^p_{[-u,u]^d}+\kappa^p\cdot\int_0^{\tau\wedge t}\|X(v)-\widetilde{X}(v)\|^p\,dv\Big),\end{aligned}$$

which implies

$$\begin{aligned}&\mathbb{E}\Big(\sup_{0\le t\le T}\Big\|\int_0^{\tau\wedge t}(a(X(v))-\widetilde{a}(\widetilde{X}(v)))\,dv\Big\|^p\Big)\\&\qquad\le c\cdot\Big(\|a-\widetilde{a}\|^p_{[-u,u]^d}+\mathbb{E}\Big(\int_0^{\tau\wedge T}\|X(v)-\widetilde{X}(v)\|^p\,dv\Big)\Big)\\&\qquad\le c\cdot\Big(\|a-\widetilde{a}\|^p_{[-u,u]^d}+\int_0^{T}\mathbb{E}\big(Y^*(\tau\wedge v)\big)^p\,dv\Big)\end{aligned}$$

for every $T \in [0,1]$. Similarly, using the Burkholder-Davis-Gundy inequality and the Lipschitz continuity of b and $\widetilde{b}$ we get

$$\begin{aligned}&\mathbb{E}\Big(\sup_{0\le t\le T}\Big\|\int_0^{\tau\wedge t}(b(X(v))-\widetilde{b}(\widetilde{X}(v)))\,dW(v)\Big\|^p\Big)\\&\qquad\le c\cdot\mathbb{E}\Big(\int_0^{\tau\wedge T}\|b(X(v))-\widetilde{b}(\widetilde{X}(v)))\|^2\,dv\Big)^{p/2}\\&\qquad\le c\cdot\Big(\|b-\widetilde{b}\|^p_{[-u,u]^d}+\int_0^{T}\mathbb{E}\big(Y^*(\tau\wedge v)\big)^p\,dv\Big)\end{aligned}$$

for every $T \in [0,1]$. Hence

$$\begin{aligned}\mathbb{E}(Y^*(\tau\wedge T))^p\le c\cdot\Big(&\|x_0-\widetilde{x}_0\|^p+\|a-\widetilde{a}\|^p_{[-u,u]^d}\\&+\|b-\widetilde{b}\|^p_{[-u,u]^d}+\int_0^{T}\mathbb{E}\big(Y^*(\tau\wedge v)\big)^p\,dv\Big)\end{aligned}$$

for every $T \in [0,1]$, which implies (7.2) by the Gronwall inequality. □

7.2. **Proof of Lemma 6.1.** We first provide estimates of the functions Φ_m. Similar estimates are established in [7, 3] and [21] but without tracing their particular dependence on all of the quantities $\|a\|$, $\|b'\|$ and $\|b\|_*$.

Lemma 7.2. *The functions Φ_m satisfy*

$$|\Phi_m(t,x,\tau,\xi)| \le (c(\gamma))^m (4\pi d^2 B^2)^{(m-1)d/2} \frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)} \cdot \frac{(\|a\| + \|b'\|^\gamma \|b\|_*^{1-\gamma})^m}{(\tau - t)^{1+d/2-m\gamma/2}} \cdot \exp\Big(-\frac{\|x-\xi\|_2^2}{4d^2B^2(\tau-t)}\Big) \tag{7.3}$$

for all $(t,x,\tau,\xi) \in V$, $\gamma \in (0,1]$ and $m \in \mathbb{N}$.

Proof. It is straightforward to see that the function Z satisfies

$$Z(t,x,\tau,\xi) \le \frac{1}{\Lambda^{d/2}\ (\tau-t)^{d/2}} \cdot \exp\Big(-\frac{\|x-\xi\|_2^2}{2d^2B^2(\tau-t)}\Big), \tag{7.4}$$

$$\Big|\frac{\partial}{\partial x_i} Z(t,x,\tau,\xi)\Big| \le \frac{\|x-\xi\|_2}{\Lambda(\tau-t)} \cdot Z(t,x,\tau,\xi), \tag{7.5}$$

$$\Big|\frac{\partial^2}{\partial x_i\,\partial x_j} Z(t,x,\tau,\xi)\Big| \le \Big(\frac{\|x-\xi\|_2^2}{\Lambda^2(\tau-t)^2} + \frac{1}{\Lambda(\tau-t)}\Big) \cdot Z(t,x,\tau,\xi) \tag{7.6}$$

for all $(t,x,\tau,\xi) \in V$ and $i,j = 1,\ldots,d$.

We prove the statement of the lemma by induction on m. Fix $\gamma \in (0,1]$, let

$$\kappa(\alpha) = \sup_{y \in \mathbb{R}} |y|^\alpha \exp(-y^2/4)$$

for $\alpha > 0$ and put

$$c^* = \frac{d^{13/2} B^4 (\kappa(\gamma) + \kappa(1) + \kappa(2+\gamma))}{\Lambda^{2+d/2}}.$$

We have

$$\begin{aligned}
|\sigma_{i,j}(x) - \sigma_{i,j}(\xi)| &= \Big|\sum_{k=1}^d \big(b_{i,k}(x)(b_{j,k}(x) - b_{j,k}(\xi)) + b_{j,k}(\xi)(b_{i,k}(x) - b_{i,k}(\xi))\big)\Big| \\
&\le \sum_{k=1}^d 2\|b\| \big(\|b'\| d^{1/2} \|x-\xi\|_2\big)^\gamma \|b\|_*^{1-\gamma} \le 2d^{3/2} B \|b'\|^\gamma \|x-\xi\|_2^\gamma \|b\|_*^{1-\gamma}.
\end{aligned}$$

Hence, by (7.5) and (7.6),

$$\begin{aligned}
&|\Phi_1(t,x,\tau,\xi)| \\
&\quad \le \bigg(d^{7/2} B \|b'\|^\gamma \|b\|_*^{1-\gamma} \Big(\frac{\|x-\xi\|_2^{2+\gamma}}{\Lambda^2(\tau-t)^2} + \frac{\|x-\xi\|_2^\gamma}{\Lambda(\tau-t)}\Big) + d\|a\| \frac{\|x-\xi\|_2}{\Lambda(\tau-t)}\bigg) Z(t,x,\tau,\xi) \\
&\quad \le \big(\|a\| + \|b'\|^\gamma \|b\|_*^{1-\gamma}\big) \frac{d^{7/2} B}{\Lambda^2} \bigg(\frac{\|x-\xi\|_2^{2+\gamma}}{(\tau-t)^2} + \frac{\|x-\xi\|_2^\gamma}{\tau-t} + \frac{\|x-\xi\|_2}{\tau-t}\bigg) Z(t,x,\tau,\xi).
\end{aligned}$$

Put

$$u = \frac{\|x-\xi\|_2}{dB(\tau-t)^{1/2}}.$$

Using (7.4) we obtain

$$\begin{aligned}
\Big(\frac{\|x-\xi\|_2^{2+\gamma}}{(\tau-t)^2} + \frac{\|x-\xi\|_2^{\gamma}}{\tau-t} + \frac{\|x-\xi\|_2}{\tau-t}\Big) \cdot Z(t,x,\tau,\xi) \\
&= \Big(\frac{(dBu)^{2+\gamma}+(dBu)^{\gamma}}{(\tau-t)^{1-\gamma/2}} + \frac{dBu}{(\tau-t)^{1/2}}\Big) \cdot Z(t,x,\tau,\xi) \\
&\le \frac{(dB)^{2+\gamma}}{(\tau-t)^{1-\gamma/2}} \cdot \frac{u^{2+\gamma}+u^{\gamma}+u}{\Lambda^{d/2}(\tau-t)^{d/2}} \cdot \exp(-u^2/2) \\
&\le \frac{(dB)^3}{\Lambda^{d/2}} \cdot \frac{\kappa(\gamma)+\kappa(1)+\kappa(2+\gamma)}{(\tau-t)^{1-\gamma/2+d/2}} \cdot \exp(-u^2/4),
\end{aligned}$$

whence (7.3) holds with $c(\gamma) = c^*$ in the case $m = 1$.

Assume that (7.3) holds with $c(\gamma) = c^*$ for some $m \in \mathbb{N}$. Then

$$\begin{aligned}
&\int_{(t,\tau)\times\mathbb{R}^d} \big|\Phi_m(t,x,t',y)\cdot\Phi_1(t',y,\tau,\xi)\big|\, d(t',y) \\
&\quad \le (c^*)^{m+1}(4\pi d^2B^2)^{(m-1)d/2}\frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)}\big(\|a\|+\|b'\|^{\gamma}\|b\|_*^{1-\gamma}\big)^{m+1} \\
&\qquad \cdot \int_{(t,\tau)\times\mathbb{R}^d} \frac{\exp\big(-\|x-y\|_2^2/(4d^2B^2(t'-t)) - \|\xi-y\|_2^2/(4d^2B^2(\tau-t'))\big)}{(t'-t)^{1+d/2-m\gamma/2}(\tau-t')^{1+d/2-\gamma/2}}\, d(t',y) \\
&\quad = (c^*)^{m+1}(4\pi d^2B^2)^{(m-1)d/2}\frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)}\big(\|a\|+\|b'\|^{\gamma}\|b\|_*^{1-\gamma}\big)^{m+1} \\
&\qquad \cdot (4\pi d^2B^2)^{d/2}\frac{\Gamma(m\gamma/2)\Gamma(\gamma/2)}{\Gamma((m+1)\gamma/2)}\frac{\exp\big(-\|x-\xi\|_2^2/(4d^2B^2(\tau-t))\big)}{(\tau-t)^{1+d/2-(m+1)\gamma/2}},
\end{aligned}$$

where the latter equality follows from [5, p.15, Lemma 3]. Thus

$$\begin{aligned}
|\Phi_{m+1}(t,x,\tau,\xi)| &\le (c^*)^{m+1}(4\pi d^2B^2)^{md/2}\frac{(\Gamma(\gamma/2))^{m+1}}{\Gamma((m+1)\gamma/2)} \\
&\qquad \cdot \frac{(\|a\|+\|b'\|^{\gamma}\|b\|_*^{1-\gamma})^{m+1}}{(\tau-t)^{1+d/2-(m+1)\gamma/2}}\exp\Big(-\frac{\|x-\xi\|_2^2}{4d^2B^2(\tau-t)}\Big),
\end{aligned}$$

which completes the proof. □

We turn to estimates of the functions G_m.

Lemma 7.3. *The functions G_m satisfy*

$$\begin{aligned}
|G_m(t,x,\tau,\xi)| &\le (c(\gamma))^m(\|a\|+\|b'\|^{\gamma}\|b\|_*^{1-\gamma})^m m^{-m\gamma/4} \\
&\qquad \cdot (\tau-t)^{(m\gamma-d)/2}\exp\Big(-\frac{\|x-\xi\|_2^2}{4d^2B^2(\tau-t)}\Big)
\end{aligned}$$

for all $(t,x,\tau,\xi) \in V$, $\gamma \in (0,1]$ and $m \in \mathbb{N}$.

Proof. Fix $\gamma \in (0,1]$ and $(t,x,\tau,\xi) \in V$. By (7.4), Lemma 7.2 and [5, p.15, Lemma 3] we have

$$\begin{aligned}
&\int_{(t,\tau)\times\mathbb{R}^d} |Z(t,x,t',y)\cdot\Phi_m(t',y,\tau,\xi)|\,d(t',y)\\
&\quad\le (c(\gamma))^m \frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)}\big(\|a\| + \|b'\|^\gamma\|b\|_*^{1-\gamma}\big)^m\\
&\qquad\cdot\int_{(t,\tau)\times\mathbb{R}^d} \frac{\exp\big(-\|x-y\|_2^2/(4d^2B^2(t'-t)) - \|\xi-y\|_2^2/(4d^2B^2(\tau-t'))\big)}{(t'-t)^{d/2}(\tau-t')^{1+d/2-m\gamma/2}}\,d(t',y)\\
&\quad= (c(\gamma))^m \frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)}\big(\|a\| + \|b'\|^\gamma\|b\|_*^{1-\gamma}\big)^m\\
&\qquad\cdot(4\pi d^2B^2)^{d/2}\frac{\Gamma(m\gamma/2)}{\Gamma(m\gamma/2+1)}\cdot\frac{\exp\big(-\|x-\xi\|_2^2/(4d^2B^2(\tau-t))\big)}{(\tau-t)^{d/2-m\gamma/2}}\\
&\quad\le (c(\gamma))^m \frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2+1)}\cdot\frac{\big(\|a\| + \|b'\|^\gamma\|b\|_*^{1-\gamma}\big)^m}{(\tau-t)^{d/2-m\gamma/2}}\exp\Big(-\frac{\|x-\xi\|_2^2}{4d^2B^2(\tau-t)}\Big).
\end{aligned}$$

Note that Γ is increasing on $[2,\infty)$. Using the Stirling formula we thus obtain

$$\begin{aligned}
\Gamma(m\gamma/2+1) &\ge \lfloor m\gamma/2\rfloor! \ge c\cdot\big(\lfloor m\gamma/2\rfloor/e\big)^{\lfloor m\gamma/2\rfloor}\\
&\ge c\cdot\big(m\gamma/(4e)\big)^{m\gamma/(4e)} \ge (c(\gamma))^m m^{m\gamma/4}
\end{aligned}$$

for all $m \in \mathbb{N}$ with $m \ge 4/\gamma$. Hence

$$\frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2+1)} \le (c(\gamma))^m m^{-m\gamma/4}$$

for all $m \in \mathbb{N}$. We conclude that

$$\begin{aligned}
|G_m(t,x,\tau,\xi)| &\le \int_{(t,\tau)\times\mathbb{R}^d} |Z(t,x,t',y)\Phi_m(t',y,\tau,\xi)|\,d(t',y)\\
&\le (c(\gamma))^m\frac{(\|a\| + \|b'\|^\gamma\|b\|_*^{1-\gamma})^m m^{-m\gamma/4}}{(\tau-t)^{(d-m\gamma)/2}}\cdot\exp\Big(-\frac{\|x-\xi\|_2^2}{4d^2B^2(\tau-t)}\Big)
\end{aligned}\tag{7.7}$$

for all $m \in \mathbb{N}$, which finishes the proof. □

We proceed with the proof of Lemma 6.1. By (2.2) and (7.7) we obtain

$$
\begin{aligned}
|J_m(f)| &\le \int_{\mathbb{R}^d} |f(\xi) \cdot G_m(0,x_0,1,\xi)|\, d\xi \\
&\le \int_{\mathbb{R}^d} |f(\xi)| \int_{(0,1)\times\mathbb{R}^d} |Z(0,x_0,t',y)\cdot \Phi_m(t',y,1,\xi)|\, d(t',y)\, d(\xi) \\
&\le c(f)(c(\gamma))^m (\|a\| + \|b'\|^\gamma \|b\|_*^{1-\gamma})^m m^{-m\gamma/4} \\
&\qquad \cdot \int_{\mathbb{R}^d} (1+\|\xi\|_2^{1+\beta}) \exp\Big(-\frac{\|x_0-\xi\|_2^2}{4d^2B^2}\Big)\, d\xi \\
&\le c(x_0,f)(c(\gamma))^m (\|a\| + \|b'\|^\gamma \|b\|_*^{1-\gamma})^m m^{-m\gamma/4}
\end{aligned} \tag{7.8}
$$

for all $m \in \mathbb{N}$. Furthermore, by (7.4),

$$
\begin{aligned}
|J_0(f)| &\le \int_{\mathbb{R}^d} |f(\xi)\cdot G_0(0,x_0,1,\xi)|\, d\xi \\
&\le \frac{c(f)}{\Lambda^{d/2}} \int_{\mathbb{R}^d} (1+\|\xi\|_2^{1+\beta}) \exp\Big(-\frac{\|x_0-\xi\|_2^2}{2d^2B^2}\Big)\, d\xi \\
&= c(f,x_0),
\end{aligned} \tag{7.9}
$$

which finishes the proof of part (i) of the lemma.

By (7.8) and (7.9),

$$
\sum_{m=0}^\infty \int_{\mathbb{R}^d} |f(\xi)\cdot G_m(0,x_0,1,\xi)|\, d\xi \le \sum_{m=0}^\infty (c(x_0,a,b,\gamma))^{m+1} m^{-m\gamma/4} < \infty.
$$

Hence

$$
\int_{\mathbb{R}^d} f(\xi)\cdot G(0,x_0,1,\xi)\, d\xi = \sum_{m=0}^\infty J_m(f),
$$

which together with (6.7) yields part (ii) of the lemma and completes the proof of Lemma 6.1.

References

[1] Bakhvalov, N. S. (1959), On approximate computation of integrals, Vestnik MGU, Ser. Math. Mech. Astron. Phys. Chem. **4**, 3–18.
[2] Davis, P. (1967), A construction of nonnegative approximate quadratures, Math. Comp. **21**, 578–582.
[3] Deck, T. Kruse, S. (2002), Parabolic differential equations with unbounded coefficents – a generalization of the parametrix method, Acta Appl. Math. **74**, 71–91.
[4] Dereich, S., Vormoor, C. (2011), The high resolution vector quantization problem with Orlicz norm distortion, J. Theor. Probab. **24**, 517–544.
[5] Friedman, A. (1964), *Partial differential equations of parabolic type*, Prentice-Hall, Inc., Englewood Cliffs, N.J.
[6] Friedman, A. (1975), *Stochastic differential equatios and applications. Vol. 1.*, Academic Press.

[7] Kruse, S. (2001), *A generalized parametrix method, smoothness of random fields and application to parabolic stochastic partial differential equations*, Dissertation, Universität Mannheim.

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

[9] Kusuoka, S. (2004), Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus, Adv. Math. Econ. **6**, 69–83.

[10] Kwas, M., Li, Y. (2003), Worst case complexity of multivariate Feynman-Kac path integration, J. Complexity **19**, 730–743.

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

[12] Lyons, T., Gyurkó, L.G. (2011), Efficient and practical implementations of cubature on Wiener space, Stochastic Analysis **2010**, 73–111.

[13] Lyons, T., Victoir, N. (2004), Cubature on Wiener space, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **460**, 169–198.

[14] Müller-Gronbach, T., Ritter, K., Yaroslavtseva, L. (2012), Derandomization of the Euler scheme for scalar stochastic differential equations, J. Complexity **28**, 139–153.

[15] Müller-Gronbach, T., Ritter, K. (2013), A local refinement strategy for constructive quantization of scalar SDEs, Found. Comput. Math. **13**, 1005–1033.

[16] Müller-Gronbach, T., Yaroslavtseva, L. (2013), A deterministic quadrature formula for SDEs based on weak Itô-Taylor steps, in preparation.

[17] Novak, E. (1988), *Deterministic and stochastic error bounds in numerical analysis*, Lecture Notes in Mathematics **1349**, Springer, Berlin.

[18] Novak, E., Woźniakowski, H. (2008), *Tractability of multivariate problems. Volume I: Linear information*, EMS Tracts in Mathematics **6**, European Mathematical Society (EMS), Zürich.

[19] Novak, E., Woźniakowski, H. (2010), *Tractability of multivariate problems. Volume II: Standard information for functionals*, EMS Tracts in Mathematics **12**, European Mathematical Society (EMS), Zürich.

[20] Novak, E., Woźniakowski, H. (2012), *Tractability of multivariate problems. Volume III: Standard information for operators*, EMS Tracts in Mathematics **18**, European Mathematical Society (EMS), Zürich.

[21] Petras, K., Ritter, K. (2006), On the complexity of parabolic initial value problems with variable drift, J. Complexity **22**, 118–145.

[22] Plaskota, L., Wasilkowski, G. W., Woźniakowski, H. (2000), A new algorithm and worst case complexity for Feynman-Kac path integration, J. Comput. Phys. **164**, 335–353.

[23] Rachev, S. (1991). *Probability metrics and the stability of stochastic models*, Wiley, Chichester.

[24] Ragozin, D. (1971), Constructive polynomial approximation on spheres and projective spaces, Trans. Amer. Math. Soc. **162**, 157–170.

[25] Ritter, K. (2000), *Average-case analysis of numerical problems*, Lecture Notes in Mathematics **1733**, Springer, Berlin.

Fakultät für Informatik und Mathematik, Universität Passau, Innstr. 33, 94030 Passau, Germany
E-mail address: thomas.mueller-gronbach@uni-passau.de

Fachbereich Mathematik, Technische Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, Germany
E-mail address: ritter@mathematik.uni-kl.de

Fakultät für Informatik und Mathematik, Universität Passau, Innstr. 33, 94030 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.