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

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

Preprint 156
The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.
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

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

(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 \( a: \mathbb{R}^d \to \mathbb{R}^d \) and \( b: \mathbb{R}^d \to \mathbb{R}^{d \times d} \), 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 $\mu$ 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

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

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

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

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 $\hat{S}$ for a particular SDE is defined in terms of a metric on $M(\mathbb{R}^d)$, which is induced by a class $F$ of test functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$. Basically, $F$ 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 $\hat{S}$ 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

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

We also impose smoothness assumptions on $a$ and $b$. Basically, $C_1$ and $C_2$ 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, $C_0$ is a bounded subset of $\mathbb{R}^d$. This leads to the definition

$$e(\hat{S}) = \sup_{(x_0, a, b) \in C} \rho_F(S(x_0, a, b), \hat{S}(x_0, a, b))$$

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

The total cost, $\text{cost}^{\text{total}}(\hat{S}, (x_0, a, b))$, for computing $\hat{S}(x_0, a, b)$ 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 $\hat{S}$ for the input $(x_0, a, b)$ plus

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

The worst case cost of \( \hat{S} \) is defined by
\[
\text{cost}^\text{total}(\hat{S}) = \sup_{(x_0,a,b) \in \mathcal{C}} \text{cost}^\text{total}(\hat{S}, (x_0, a, b)).
\]

As the key quantity we study the \( N \)-th minimal error
\[
e_N^\text{total}(\mathcal{C}, \mathcal{F}) = \inf \{ e(\hat{S}) \mid \text{cost}^\text{total}(\hat{S}) \leq N \}, \quad N \in \mathbb{N},
\]
which depends on the smoothness parameters \( r, s, s_1, s_2 \) via \( \mathcal{F} \) and \( \mathcal{C} \) as well as on the dimension \( d \). In particular, we show that
\[
N^{-\eta^\text{total}} \lesssim e_N^\text{total}(\mathcal{C}, \mathcal{F}) \lesssim N^{-\eta^\text{total}+\varepsilon}
\]
for every \( \varepsilon > 0 \) with
\[
\eta^\text{total} = \min(r, s_1, s_2)/d.
\]
Thus, up to \( \varepsilon \), 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 cost\(^\ast(\hat{S})\) with \( \ast \in \{\inf, \supp\} \). The \( N \)-th minimal errors
\[
e_N^\ast(\mathcal{C}, \mathcal{F}) = \inf \{ e(\hat{S}) \mid \text{cost}^\ast(\hat{S}) \leq N \}, \quad N \in \mathbb{N},
\]
satisfy
\[
N^{-\eta^\ast} \lesssim e_N^\ast(\mathcal{C}, \mathcal{F}) \lesssim N^{-\eta^\ast+\varepsilon}
\]
for every \( \varepsilon > 0 \), where
\[
\eta^\ast = \begin{cases} 
\min(s_1, s_2)/d, & \text{if } \ast = \inf, \\
r/d, & \text{if } \ast = \supp.
\end{cases}
\]

We sketch how to derive the upper bounds for the minimal errors, and we start with the analysis of \( e_N^\ast(\mathcal{C}, \mathcal{F}) \) with \( \ast \in \{\inf, \supp\} \). 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 \( \ast = \inf \) and \( \ast = \supp \) 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 \( \tilde{a} \) and \( \tilde{b} \) to \( a \) and \( b \) such that the solutions \( \tilde{X} \) and \( X \) of the corresponding SDEs satisfy
\[
(\mathbb{E}[\|X(1) - \tilde{X}(1)\|^2])^{1/2} \lesssim n^{-s}
\]
with \( s = \min(s_1, s_2) \). In the next step a quantization procedure is applied to \( \tilde{X}(1) \). 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

\[
(\mathbb{E} \| \tilde{X}(1) - T(\tilde{X}(1)) \|^2)^{1/2} \precsim n^{-s}.
\]

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

\[
e(\hat{S}) \precsim n^{-s}
\]

for smoothness \( r = 1 \) of the test functions \( f \in \mathcal{F} \). To properly handle the remaining case \(* = \text{supp} \) and \( r > 1 \), a support reduction algorithm \( \mathcal{R} \) due to [2] is applied to \( \hat{S}(x_0, a, b) \), and this yields a measure \( \mathcal{R}(\hat{S}(x_0, a, b)) \in M_0(\mathbb{R}^d) \) 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

\[
e(\mathcal{R} \circ \hat{S}) \precsim n^{-s}.
\]

Unfortunately, \( \hat{S} \) relies on the probability values of a large number of rectangles in \( \mathbb{R}^d \) w.r.t. the distributions \( S(x_0, \tilde{a}, \tilde{b}) \) for all \( (x_0, a, b) \in \mathcal{C} \), i.e., for infinitely many SDEs. To address this problem we construct a modification of \( \hat{S} \), which is implementable in the real number model and achieves, up to \( \varepsilon \), the optimal order \( \eta^* \) of the error in terms of \( \text{cost}^* \) for \(* \in \{\text{supp}, \text{inf}, \text{total}\} \). We basically proceed as follows. By suitably rounding the function values of \( a \) and \( b \) that enter the computation of the approximations \( \tilde{a} \) and \( \tilde{b} \), 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 \( \hat{S}^{\text{mod}}(x_0, a, b) \in M_0(\mathbb{R}^d) \) so that \( \hat{S}^{\text{mod}} \) enjoys all the relevant properties of \( \hat{S} \). Finally, we consider the algorithm \( \mathcal{R} \circ \hat{S}^{\text{mod}} \). In addition to the error bound of order \( s \), and the bounds of orders \( d(1 + \varepsilon) \cdot s/r \) for the support size, \( \mathcal{R} \circ \hat{S}^{\text{mod}} \) 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 \( \mathcal{R} \circ \hat{S}^{\text{mod}} \) is hardly of a practical relevance, as \( \mathcal{R} \circ \hat{S}^{\text{mod}} \) still relies on heavy precomputation. The main purpose of \( \mathcal{R} \circ \hat{S}^{\text{mod}} \) 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 \(* \in \{\text{inf}, \text{supp}\} \). 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 \(* = \text{supp} \) 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 \(* = \text{inf} \) 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 \( c, c_1, \ldots \) and \( c(\cdot), c_1(\cdot), \ldots \) 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 \( \lambda_{\text{min}}(A) \) and \( \lambda_{\text{max}}(A) \) 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 \( \text{diag}(x) \) to denote the diagonal matrix \( (a_{i,j})_{1 \leq i,j \leq d} \in \mathbb{R}^{d \times d} \) with \( a_{i,i} = x_i \) for \( i = 1, \ldots, d \). For a function \( h : \mathbb{R}^d \to \mathbb{R}^d \) we use \( \text{diag}(h) \) to denote the mapping \( x \mapsto \text{diag}(h(x)) \). 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 \( \| f \| = \sup_{x \in \mathbb{R}^d} \| f(x) \| \), \( \| f \|_I = \sup_{x \in I} \| f(x) \| \) for a function \( f : \mathbb{R}^d \to \mathbb{R}^{m \times k} \) 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 \( \kappa \geq 0 \) w.r.t. the maximum norm on \( \mathbb{R}^d \), respectively, and we put

\[
\mathcal{L}^0(d, \kappa) = \{ f \in \mathcal{L}(d, \kappa) \mid |f(0)| \leq \kappa \}.
\]

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

\[
|f^{(\alpha)}(x)| \leq K \cdot (1 + \|x\|^{\beta})
\]

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

\[
\mathcal{F}^0(d, r, K) = \{ f \in \mathcal{F}(d, r, K, 0) \mid |f(0)| \leq K \}.
\]

For later purposes we note that

\[
|f(x) - f(\tilde{x})| \leq c(d, K) \cdot (1 + \|x\|^\beta + \|\tilde{x}\|^\beta) \cdot \|x - \tilde{x}\|
\]

and, consequently,

\[
|f(x)| \leq c(d, K, f) \cdot (1 + \|x\|^{1+\beta})
\]

(2.1)
for every $f \in \mathcal{F}(d, 1, K, \beta)$ and all $x, \bar{x} \in \mathbb{R}^d$.

Finally, for sequences $u_n$ and $v_n$ of non-negative reals the notion $u_n \preceq v_n$ 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 $\psi_i : \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d$, $i = 1, \ldots, N$.

Define

$$N_{\psi_1, \ldots, \psi_N} : \mathcal{C} \to \mathbb{R}^{d+N(d+d^2)}$$

by

$$N_{\psi_1, \ldots, \psi_N}(x_0, a, b) = (x_0, y_1, \ldots, y_N)$$

with

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

and

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

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

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

Then

$$(3.1) \quad \psi \circ N_{\psi_1, \ldots, \psi_N} : \mathcal{C} \to M_0(\mathbb{R}^d)$$

defines an algorithm, which sequentially evaluates the coefficients $a$ and $b$ from an equation $(x_0, a, b) \in \mathcal{C}$ at the points

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

and finally applies the mapping $\psi$ to the data $N_{\psi_1, \ldots, \psi_N}(x_0, a, b)$ to obtain a probability measure with finite support.

We use

$$\mathcal{S}_N(\mathcal{C}) = \{ \psi \circ N_{\psi_1, \ldots, \psi_N} \mid \psi : \mathbb{R}^{d+N(d+d^2)} \to M_0(\mathbb{R}^d),$$

$$\psi_i : \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d, i = 1, \ldots, N \}$$

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

$$\mathcal{S}(\mathcal{C}) = \bigcup_{n \in \mathbb{N}} \mathcal{S}_N(\mathcal{C}).$$
We use different notions of cost for the comparison of algorithms \( \hat{S} \in \mathcal{S}(C) \), in particular, the information cost
\[
\text{cost}^\text{inf}(\hat{S}) = \min \{ N \in \mathbb{N} \mid \hat{S} \in \mathcal{S}_N(C) \},
\]
i.e., the maximum number of evaluations of the coefficients \( a, b \) that are used by \( \hat{S} \), and
\[
\text{cost}^\text{supp}(\hat{S}) = \sup \{ |\text{supp}(\hat{S}(x_0, a, b))| \mid (x_0, a, b) \in C \},
\]
i.e., the maximum support size of the probability measures computed by \( \hat{S} \). 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 \( \text{cost}^\text{total}(\hat{S}) \) as the sum of \( \text{cost}^\text{inf}(\hat{S}) + \text{cost}^\text{supp}(\hat{S}) \)
and the worst case count of all basic computational operations that are carried out by an algorithm \( \hat{S} \) for input equations \((x_0, a, b)\) from the class \( C \).

For the definition of the error of an algorithm \( \hat{S} \in \mathcal{S}(C) \) we employ metrics on the space \( M(\mathbb{R}^d) \) given by
\[
\rho_F(\mu_1, \mu_2) = \sup_{f \in F} \left| \int_{\mathbb{R}^d} f \, d\mu_1 - \int_{\mathbb{R}^d} f \, d\mu_2 \right|,
\]
where \( F \) is a class of functions \( f : \mathbb{R}^d \to \mathbb{R} \) that satisfies
\[
\sup_{f \in F} \sup_{x \in \mathbb{R}^d} \frac{|f(x) - f(0)|}{1 + \|x\|^p} < \infty
\]
for some \( p \in [1, \infty) \). In particular, \( \rho_F \) is called a metric with \( \zeta \)-structure, if all functions \( f \in F \) are bounded, see [23, p. 72]. The worst case error of \( \hat{S} \) on \( C \) w.r.t. the metric \( \rho_F \) is defined by
\[
e(\hat{S}) = \sup_{(x_0, a, b) \in \mathcal{C}} \rho_F(S(x_0, a, b), \hat{S}(x_0, a, b)).
\]

The key quantities in our analysis are the \( N \)-th minimal errors
\[
e_N^*(C, F) = \inf \{ e(\hat{S}) \mid \text{cost}^*(\hat{S}) \leq N \}, \quad N \in \mathbb{N},
\]
where \( * \in \{ \text{inf}, \text{supp}, \text{total} \} \). We determine the decay of these quantities dependent on the smoothness properties of the classes \( C \) and \( F \).

Additionally, we study the related problem of computing the expected values
\[
A^f(x_0, a, b) = \int_{\mathbb{R}^d} f \, dS(x_0, a, b), \quad (x_0, a, b) \in C,
\]
for a fixed integrand \( f \in F \) 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
\[ \mathfrak{A}_N = \{ \psi \circ \mathcal{N}_{\psi_1, \ldots, \psi_N} \mid \psi : \mathbb{R}^{d+N(d+d^2)} \to \mathbb{R}, \psi_i : \mathbb{R}^{d+(i-1)(d+d^2)} \to \mathbb{R}^d, \ i = 1, \ldots, N \}, \]
and we define the cost and the error of \( \hat{A} \in \mathfrak{A} \) by
\[ \text{cost}(\hat{A}) = \min\{ N \in \mathbb{N} \mid \hat{A} \in \mathfrak{A}_N \} \]
and
\[ e^f(\hat{A}) = \sup_{(x_0, a, b) \in \mathcal{C}} \| A^f(x_0, a, b) - \hat{A}(x_0, a, b) \|. \]
Thus the corresponding \( N \)-th minimal errors are given by
\[ e^f_N(\mathcal{C}) = \inf \{ e^f(\hat{A}) \mid \text{cost}(\hat{A}) \leq N \}, \quad N \in \mathbb{N}. \]
The study of \( e^f_N(\mathcal{C}) \) is primarily motivated by the following fact. If \( \hat{S} \in \mathcal{G} \) then
\[ \hat{A}(x_0, a, b) = \int_{\mathbb{R}^d} f \ d\hat{S}(x_0, a, b), \quad (x_0, a, b) \in \mathcal{C}, \]
defines an algorithm \( \hat{A} \in \mathfrak{A} \) with \( \text{cost}(\hat{A}) \leq \text{cost}^{\inf}(\hat{S}) \), and therefore
\[ e_{N}^{\inf}(\mathcal{C}, \mathcal{F}) \geq e_{N}^{f}(\mathcal{C}) \]
for every \( N \in \mathbb{N} \) and \( f \in \mathcal{F} \).

4. Results

We consider the classes
\[ \mathcal{C} = \mathcal{C}(d, s_1, s_2, K) = [-K, K]^d \times (\mathcal{F}^0(d, s_1, K))^d \times (\mathcal{F}^0(d, s_2, K))^d \]
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 \( \beta \geq 0 \). Then
\[ N^{-\eta^*} \ll e_{N}^{\star}(\mathcal{C}, \mathcal{F}) \ll N^{-\eta^*+\varepsilon} \]
with
\[ \eta^* = \begin{cases} \min(s_1, s_2)/d, & \text{if } \star = \text{inf}, \\ r/d, & \text{if } \star = \text{supp}, \\ \min(r, s_1, s_2)/d, & \text{if } \star = \text{total}, \end{cases} \]
for every \( \varepsilon > 0 \) and \( \star \in \{ \text{inf}, \text{supp}, \text{total} \} \).
We conjecture that the upper bounds in Theorem 4.1 actually hold with \( \varepsilon = 0 \). Obviously, it suffices to establish the lower bounds for \( \ast \in \{ \inf, \supp \} \).

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 \( \tilde{S}_{n, \varepsilon} \) such that
\[
\lim_{n \to \infty} \text{cost}^*(\tilde{S}_{n, \varepsilon}) = \infty \quad \text{as well as} \quad e(\tilde{S}_{n, \varepsilon}) \lesssim \text{cost}^*(\tilde{S}_{n, \varepsilon}) - \eta^* + \varepsilon
\]
simultaneously for \( \ast = \inf \) and \( \ast = \supp \), see Section 5.1. In Section 5.2 we present a modified version \( \tilde{S}^{\text{mod}}_{n, \varepsilon} \) of \( \tilde{S}_{n, \varepsilon} \), which satisfies \( \lim_{n \to \infty} \text{cost}^*(\tilde{S}^{\text{mod}}_{n, \varepsilon}) = \infty \) and achieves (4.1) with \( \tilde{S}_{n, \varepsilon} \) replaced by \( \tilde{S}^{\text{mod}}_{n, \varepsilon} \) as well as
\[
e(\tilde{S}^{\text{mod}}_{n, \varepsilon}) \lesssim (\text{cost}^{\text{total}}(\tilde{S}^{\text{mod}}_{n, \varepsilon})) - \eta^{\text{total}} + \varepsilon.
\]
However, implementing \( \tilde{S}^{\text{mod}}_{n, \varepsilon} \) requires at least \( n^2d^2n^d \min(s_1, s_2) \) 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
\[
r = 1 \quad \text{and} \quad \beta = 0,
\]
i.e., we essentially take Lipschitz continuous integrands \( f \) with Lipschitz constant \( K \), and consider the subclass \( \tilde{C} \subset C(d, s_1, s_2, K) \) of equations \((x_0, a, b)\) such that
\[
\|a\|, \|b\| \leq K, \quad \inf_{x \in \mathbb{R}^d} \lambda_{\min}(bb^T(x)) \geq \delta
\]
with a fixed \( \delta > 0 \). Assume further that
\[
\min(s_1, s_2) \geq 6.
\]
In the forthcoming work \[16\] algorithms \( \tilde{S}_{n, \varepsilon} \) 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 \( \tilde{C} \) satisfy
\[
e(\tilde{S}_{n, \varepsilon}) \lesssim (\text{cost}^{\inf}(\tilde{S}_{n, \varepsilon}))^{-\min(s_1, s_2)/d + \varepsilon}
\]
as well as
\[
e(\tilde{S}_{n, \varepsilon}) \lesssim (\text{cost}^{\text{total}}(\tilde{S}_{n, \varepsilon}))^{-1/d + \varepsilon}.
\]
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 \( (\tilde{S}_{n, \varepsilon})_{n \in \mathbb{N}} \) performs asymptotically optimal, up to the arbitrarily small exponent \( \varepsilon \), with respect to \( \text{cost}^* \) for \( \ast \in \{ \supp, \inf, \text{total} \} \).
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 \(\beta\) given by (4.2) and the subclass \(\tilde{C} \subset \mathcal{C}(d, s_1, s_2, K)\) 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 \(\mathcal{S}_n\) that satisfy

\[
e(\mathcal{S}_n) \ll \ln(n) \cdot n^{-\min(s_1, s_2)/2}.
\]

as well as

\[
\text{cost}^{\text{total}}(\mathcal{S}_n) \ll n^{(\min(s_1, s_2) - 1)/2}.
\]

Thus

\[
e(\mathcal{S}_n) \ll \ln(\text{cost}^* (\mathcal{S}_n)) \cdot (\text{cost}^* (\mathcal{S}_n))^{-(\min(s_1, s_2) - 2)/((\min(s_1, s_2) - 1)/2)}
\]

for \(* \in \{\text{supp}, \text{total}\}\), 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

\[
e_N^f (\mathcal{C}(d, s_1, s_2, K)) \ll N^{-\min(s_1, s_2)/d + \varepsilon}
\]

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

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

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
\[
e_{N}^{f}(C(1)) \succ N^{-s/d}.
\]

(ii) Assume that \( f \in \mathcal{F}(d, 2, K, \beta) \) is not multilinear. Then
\[
e_{N}^{f}(C(2)) \succ N^{-s/d}.
\]

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 \( e_{N}^{f}(C(i)) = 0 \) 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 \( \varepsilon \), than \( M_{0}(\mathbb{R}^{d}) \) with respect to costinf.

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 \( a : \mathbb{R}^{d} \times [0, 1] \rightarrow \mathbb{R}^{d} \) 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 \( \ast \in \{\inf, \supp\} \) the proofs from Section 5.1 carry over almost literally. The upper bound for \( \ast = \text{total} \) is obtained by an obvious modification of the algorithm \( \hat{S}_{n, \varepsilon}^{\text{mod}} \) 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
\[ x_0 = 0, \quad a(x) = K \begin{pmatrix} 0 \\ x_1 \\ \vdots \\ x_{d-1} \end{pmatrix} \in \mathbb{R}^d, \quad b(x) = \begin{pmatrix} K & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{d \times m}, \]
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., \( \text{cost}^{\supp} \), is used as a measure of cost. In the case \( m < d \), if the information cost, i.e., \( \text{cost}^{\inf} \), 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 \( C = C(d, s, s, K) \) and \( F = F(d, r, \beta, K) \).

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

5.1. The case \(* \in \{\inf, \supp\}\). For every \( n \in \mathbb{N} \) and \( \varepsilon \in (0, 1] \) we present an algorithm
\[ \hat{S}_{n, \varepsilon} : C \rightarrow M_0(\mathbb{R}^d), \]
which consists of the following ingredients.

1) Evaluation of the coefficients \( a \) and \( b \) from any input \((x_0, a, b) \in C\) on the uniform grid
\[ \mathcal{X} = \left( 1/n \cdot \mathbb{Z} \cap \left[ -\left[ n^\varepsilon \right], \left[ n^\varepsilon \right] \right] \right)^d, \]
i.e., the mapping
\[ \mathcal{N} : \mathbb{R}^d \rightarrow \mathbb{R}^{|\mathcal{X}|}, \quad f \mapsto (f(x))_{x \in \mathcal{X}}, \]
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

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

where

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

2) A linear mapping

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

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) $\|f - \mathcal{A}(\mathcal{N}(f))\|_{[-n^d,n^d]^d} \leq c \cdot n^{-s}$ 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

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

where

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

3) A mapping

$$\mathcal{Q}: \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^{d^2} \to \mathcal{M}_0(\mathbb{R}^d),$$

which satisfies

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

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

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

$$\mathcal{Q}(\overline{\mathcal{A}}(\mathcal{N}(x_0, a, b))) \in \mathcal{M}_0(\mathbb{R}^d)$$

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

$$(5.1) \quad \hat{S}_{n,\varepsilon} = \mathcal{Q} \circ \overline{\mathcal{A}} \circ \mathcal{N}.$$ 

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

$$e(\hat{S}_{n,\varepsilon}) \leq c \cdot n^{-s}$$
and

\[
\text{cost}^\ast(\hat{S}_{n,\varepsilon}) \leq c \cdot \begin{cases} 
  n^{d(1+\varepsilon)}, & \text{if } \ast = \inf, \\
  n^{d(1+\varepsilon) s/r}, & \text{if } \ast = \supp.
\end{cases}
\]

Clearly, Theorem 5.1 implies the upper bounds in Theorem 4.1 except for \( \ast = \text{total} \).

**Proof of Theorem 5.1.** We have

\[
|X| \leq c \cdot n^{d(1+\varepsilon)},
\]

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

Let \((x_0, a, b) \in C\) and put \((x_0, \tilde{a}, \tilde{b}) = A(N(x_0, a, b))\).

Thus

\[
\tilde{a}_i = A(N(a_i)), \quad \tilde{b}_{i,j} = A(N(b_{i,j}))
\]

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

\[
(5.3) \quad \tilde{a}_i, \tilde{b}_{i,j} \in L^0(d, c)
\]

follows from property 2)(i) of \( A \). Furthermore, by property 2)(ii) of \( A \),

\[
(5.4) \quad \|a_i - \tilde{a}_i\|_{[-n^s, n^s]^d} + \|b_{i,j} - \tilde{b}_{i,j}\|_{[-n^s, n^s]^d} \leq c \cdot n^{-s}.
\]

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

\[
(5.5) \quad \mathbb{E} \left( \sup_{t \in [0,1]} (\|X(t)\|^p + \|\tilde{X}(t)\|^p) \right) \leq c(p) \cdot (1 + \|x_0\|^p)
\]

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

\[
(5.6) \quad (\mathbb{E}\|X(1) - \tilde{X}(1)\|^2)^{1/2} \leq c \cdot n^{-s}
\]

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),

\[
|\mathbb{E}(f(X(1))) - \mathbb{E}(f(\tilde{X}(1)))| \leq c \cdot (1 + (\mathbb{E}\|X(1)\|^{2\beta})^{1/2} + (\mathbb{E}\|\tilde{X}(1)\|^{2\beta})^{1/2}) \cdot (\mathbb{E}\|X(1) - \tilde{X}(1)\|^2)^{1/2} \leq c \cdot n^{-s}.
\]

It follows

\[
\rho_F(S(x_0, a, b), S(x_0, \tilde{a}, \tilde{b})) \leq c \cdot n^{-s}.
\]

By (5.3) and property 3)(ii) of the mapping \( Q \) we have

\[
\rho_F(S(x_0, \tilde{a}, \tilde{b}), Q(x_0, \tilde{a}, \tilde{b})) \leq c \cdot n^{-s},
\]

which completes the proof. \( \square \)
We construct \( Q \) in two steps, i.e.,
\[
Q = Q_2 \circ Q_1,
\]
where the mapping
\[
Q_1 : \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^d^2 \to M_0(\mathbb{R}^d)
\]
is obtained by applying a quantization procedure for random variables with finite moments of any order, see Section 5.1.1 and satisfies
\[
\max_{x \in \text{supp}(Q_1(x_0,a,b))} \|x\| \leq n^{\varepsilon s/(r+\beta)},
\]
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
\[
Q_2 : M_0(\mathbb{R}^d) \to M_0(\mathbb{R}^d)
\]
is obtained by applying a measure support reduction algorithm, see Section 5.1.2 and satisfies
\[
|\text{supp}(Q_2(\mu))| \leq c \cdot n^{d(1+\varepsilon) - s/r},
\]
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 \( Q_1 \). 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
\[
T_{\gamma,N} = \sum_{j=1}^{J} \sum_{k=1}^{N_j} (y_{j,k} \cdot 1_{[y_{j,k}, y_{j,k+1}]} - y_{j,k} \cdot 1_{(-y_{j,k+1}, -y_{j,k})})
\]
and consider the mapping
\[
T_{\gamma,N}^{(d)} : \mathbb{R}^d \to \mathbb{R}, \quad (x_1, \ldots, x_d) \mapsto (T_{\gamma,N}(x_1), \ldots, T_{\gamma,N}(x_d)).
\]
Thus
\[
T_{\gamma,N}^{(d)}(\mathbb{R}^d) \subset [-N^{1/\gamma}, N^{1/\gamma}]^d
\]
and we have
\begin{equation}
(2N - 1)^d \leq |T^{(d)}_{\gamma,N}(\mathbb{R}^d)| = \left(2 \sum_{j=1}^{J} N_j - 1\right)^d \leq (2N\gamma/(\gamma - 1))^d.
\end{equation}

Note further that
\begin{equation}
\|T^{(d)}_{\gamma,N}(x)\|_p \leq \|x\|_p
\end{equation}
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 $q \geq 1$, every $p > 1 + (1 + \gamma)q$ and every $\mathbb{R}^d$-valued random variable $Z$ we have
\[
\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q \leq c(q,p) \max\{1,\mathbb{E}\|Z\|^p\} N^{-q}.
\]

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
\[
\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q = \sum_{j=1}^{\infty} \mathbb{E}\left(1_{B_j}(\|Z\|) \cdot \|Z - T^{(d)}_{\gamma,N}(Z)\|^q\right)
\leq \sum_{j=1}^{J} N_j^{-q} \cdot \mathbb{P}(\|Z\| \in B_j) + \sum_{j>J} j^q \cdot \mathbb{P}(\|Z\| \in B_j).
\]

We have
\[
N_j^{-q} \cdot \mathbb{P}(\|Z\| \in B_j) \leq c(q) N^{-q} j^\gamma q \cdot \mathbb{P}(\|Z\| \geq j - 1) \leq c(q,p) N^{-q} j^\gamma q - p \cdot \mathbb{E}\|Z\|^p
\]
for $j = 2,\ldots,J$, and
\[
j^q \cdot \mathbb{P}(\|Z\| \in B_j) \leq (j^\gamma/N)^q j^q \cdot \mathbb{P}(\|Z\| \geq j - 1) \leq c(p) N^{-q} j^{(1+\gamma)q-p} \cdot \mathbb{E}\|Z\|^p
\]
for $j > J$. Hence
\[
\mathbb{E}\|Z - T^{(d)}_{\gamma,N}(Z)\|^q \leq c(q,p) N^{-q} \left(1 + \mathbb{E}\|Z\|^p \cdot \sum_{j \geq 2} j^{(1+\gamma)q-p}\right),
\]
which finishes the proof. \hfill \square

Take
\[
\gamma = (r + \beta) \cdot \varepsilon^{-1}, \quad N = n^s
\]
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 $\|x_0\| \leq \kappa$ and $a_i, b_{i,j} \in L^0(d, \kappa)$ for $i, j = 1, \ldots, d$. Then the measure $Q_1(x_0, a, b)$ satisfies (5.7).

Proof. By definition,

$$\text{supp}(Q_1(x_0, a, b)) \subset T_{(r+\beta), \varepsilon^{-1}, n^s}(\mathbb{R}^d).$$

Thus

$$\max_{x \in \text{supp}(Q_1(x_0, a, b))} \|x\| \leq n^{\varepsilon s/(r+\beta)}$$

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

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

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

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

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

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

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

$$|\mathbb{E}(f(X(1))) - \mathbb{E}(f(V))| \leq c \cdot \mathbb{E}((1 + \|X(1)\|^\beta + \|V\|^\beta) \cdot \|X(1) - V\|) \leq c(\kappa) \cdot n^{-s}$$

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

5.1.2. The mapping $Q_2$. 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 $\tau$. By a well-known sequential support point elimination procedure due to [2] we obtain an algorithm

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

which satisfies

(5.12) \quad \text{supp}(\mathcal{R}_\tau(\mu)) \subset \text{supp}(\mu)

as well as

(5.13) \quad |\text{supp}(\mathcal{R}_\tau(\mu))| \leq \dim(\mathcal{P}_\tau)

and

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

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

We define

$$Q_2 = \mathcal{R}_\tau$$
with
\[
\tau = \max(r - 1, \lceil n^{(1+\varepsilon)/s} \rceil).
\]

**Lemma 5.3.** The mapping \( Q_2 \) satisfies (5.8).

**Proof.** Let \( \mu \in M_0(\mathbb{R}^d) \). By (5.13) we have
\[
|\text{supp}(Q_2(\mu))| \leq \left( \tau + \frac{d}{d} \right) \leq 2^d \cdot \tau^d,
\]
which yields the first estimate in (5.8).

Put
\[
\eta = \max_{x \in \text{supp}(\mu)} \|x\|.
\]
Let \( f \in \mathcal{F}(d, r, K, \beta) \). By (5.12) and (5.14) we have
\[
\left| \int_{\mathbb{R}^d} f \, d\mu - \int_{\mathbb{R}^d} f \, dQ_2(\mu) \right| \leq \int_{\mathbb{R}^d} |f - p| \, d\mu + \int_{\mathbb{R}^d} |f - p| \, dQ_2(\mu)
\leq 2 \max_{z \in \text{supp}(\mu)} |f(z) - p(z)|
\leq 2 \|f - p\|_{[-\eta, \eta]^d}
\]
for every \( p \in \mathcal{P}_r \). Using a Jackson-type theorem, see, e.g., [24, Theorem 3.4], we get
\[
\inf_{p \in \mathcal{P}_r} \|f - p\|_{[-\eta, \eta]^d} \leq c \cdot \tau^{-r} \cdot \eta^r \cdot \max_{|\alpha|=r} \|f^{(\alpha)}\|_{[-\eta, \eta]^d}
\leq c \cdot n^{-(1+\varepsilon)/s} \cdot (1 + \eta^{r+\beta}),
\]
which completes the proof. \( \square \)

5.2. The case \( * = \text{total} \). The algorithm \( \hat{S}_{n, \varepsilon} = Q_2 \circ Q_1 \circ \overline{\mathcal{A}} \circ \mathcal{N} \) is hardly implementable since the computation of the discrete measures \( Q_1(\overline{\mathcal{A}}(\mathcal{N}(x_0, a, b))) \) requires the knowledge of the marginal probabilities (5.15) for certain rectangles \( R \subset \mathbb{R}^d \) and all of the equations \( (x_0, \tilde{a}, \tilde{b}) = \overline{\mathcal{A}}(\mathcal{N}(x_0, a, b)) \) with \((x_0, a, b) \in C \), see Section 5.1.1. We present a modified version \( \hat{S}^\text{mod}_{n, \varepsilon} \) of \( \hat{S}_{n, \varepsilon} \), which is implementable in the real number model and achieves, up to the arbitrarily small exponent \( \varepsilon \), the optimal order \( \eta^* \) in terms of cost \( * \in \{\text{supp, inf, total}\} \), see Theorem 4.1. The idea is, essentially, to replace the mapping \( Q_1 \) 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 \( C \) 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 \( Z_1, Z_2, \ldots \) denote an independent sequence of \( d \)-dimensional standard normally distributed random variables. Fix \( x_0 \in \mathbb{R}^d \) as well as \( a: \mathbb{R}^d \to \mathbb{R}^d \)
and $b: \mathbb{R}^d \to \mathbb{R}^{d \times d}$. Let $m \in \mathbb{N}$ and consider the corresponding weak Euler scheme with step-size $1/m$ given by
\[
U_0 = x_0, \\
U_\ell = U_{\ell-1} + a(U_{\ell-1}) \cdot m^{-1} + b(U_{\ell-1}) \cdot m^{-1/2} \cdot Z_\ell
\]
for $\ell = 1, \ldots, m$. We obtain a quantization $(\tilde{U}_\ell)_{\ell=0,\ldots,m}$ of this scheme by iteratively applying the mapping $T_{\gamma,N}^{(d)}$ to the consecutive single Euler steps, i.e., we put
\[
\tilde{U}_0 = U_0 = x_0
\]
and we define recursively
\[
U_\ell = \tilde{U}_{\ell-1} + a(\tilde{U}_{\ell-1}) \cdot m^{-1} + b(\tilde{U}_{\ell-1}) \cdot m^{-1/2} \cdot Z_\ell, \\
\tilde{U}_\ell = T_{\gamma,N}^{(d)}(\tilde{U}_\ell)
\]
for $\ell = 1, \ldots, m$.
We take
\[
\gamma = 3(r + \beta) \cdot \varepsilon^{-1}, \quad N = n^{3s}, \quad m = n^{2s}
\]
and we define a mapping
\[
Q^\text{mod}_1: \mathbb{R}^d \times (\mathcal{L}(d))^d \times (\mathcal{L}(d))^{d^2} \to M_0(\mathbb{R}^d)
\]
by
\[
Q^\text{mod}_1(x_0, a, b) = P_{\tilde{U}_m}.
\]
In order to replace $\mathcal{C}$ by a finite set of equations we employ a projection
\[
\mathcal{P}: \mathbb{R} \to \mathcal{Y}
\]
into the uniform grid
\[
\mathcal{Y} = 1/n^s \cdot \mathbb{Z} \cap [-2K \cdot \lfloor n^\varepsilon \rfloor, 2K \cdot \lceil n^\varepsilon \rceil]
\]
such that
\[
|x - \mathcal{P}(x)| = \min_{y \in \mathcal{Y}} |x - y|
\]
for all $x \in \mathbb{R}$. Hereby the available data $\mathcal{N}(x_0, a, b)$ from an input $(x_0, a, b) \in \mathcal{C}$ is approximated by the vector
\[
\mathcal{P}(\mathcal{N}(x_0, a, b)) \in \mathcal{Y}^{d+|X| \cdot (d+d^2)},
\]
where
\[
\mathcal{P}: \mathbb{R}^{d+|X| \cdot (d+d^2)} \to \mathcal{Y}^{d+|X| \cdot (d+d^2)}, \quad y \mapsto (\mathcal{P}(y_i))_{i=1,\ldots,d+|X| \cdot (d+d^2)}.
\]
Finally we require that the linear mapping $\mathcal{A}$ satisfies
\[
\forall x \in \mathbb{R}^{|X|}: \|\mathcal{A}(x)\|_{[-n^\varepsilon, n^\varepsilon]^d} \leq c \cdot \|x\|,
\]
additionally to the conditions 2)(i) and 2)(ii), and we define
\[ \hat{S}^{\text{mod}}_{n,\varepsilon} = Q_2 \circ Q_1^{\text{mod}} \circ \bar{A} \circ \bar{P} \circ \bar{N}. \]

Using Lemma 5.1 and standard error estimates for the Euler scheme one can show that Lemma 5.2 still holds true with \( Q_1 \) replaced by \( Q^{\text{mod}}_1 \), and therefore the mapping \( Q_2 \circ Q_1^{\text{mod}} \) enjoys the same properties 3)(i) and 3)(ii) as the mapping \( Q = Q_2 \circ Q_1 \). Put \((\hat{x}_0, \hat{a}, \hat{b}) = \bar{A}(\bar{P}(\bar{N}(x_0, a, b))) \) for \((x_0, a, b) \in C \). Using properties 2)(i), 2)(ii) and (5.16) of \( A \) one can show that (5.3) as well as (5.4) hold true with \( a_i \) and \( b_{i,j} \) in place of \( \hat{a}_i \) and \( \hat{b}_{i,j} \), respectively. We conclude that Theorem 5.1 holds true with \( \hat{S}^{\text{mod}}_{n,\varepsilon} \) in place of \( \hat{S}^{\text{mod}}_{n,\varepsilon} \).

Next, we consider the total cost of \( \hat{S}^{\text{mod}}_{n,\varepsilon} \). Put
\[ V = Y^d + |X| \cdot (d + d^2). \]
In the real number model we may assume that all of the discrete measures
\[ Q_2(Q_1^{\text{mod}}(\bar{A}(\nu))), \quad \nu \in V, \]
which constitute the possible outcomes of the algorithm \( \hat{S}^{\text{mod}}_{n,\varepsilon} \), are available via precomputing in the form of \(|V|\) stored lists of support points and corresponding probability weights. As a consequence the mapping \( Q_2 \circ Q_1^{\text{mod}} \circ \bar{A} \) can be implemented in such a way that the number of basic computational operations needed to compute \( Q_2(Q_1^{\text{mod}}(\bar{A}(\nu))) \) at \( \nu \in V \) is bounded by \( c \cdot (|X| + |\text{supp}(Q_2(Q_1^{\text{mod}}(\bar{A}(\nu))))|) \). Furthermore, the projection \( \bar{P} : \mathbb{R} \to Y \) can be implemented in such a way that the number of basic computational operations needed for a single evaluation of \( \bar{P} \circ \bar{N} \) is bounded by \( c \cdot (d + |X| + d^2) \). In total we obtain the upper bound \( c \cdot (|X| + |\text{supp}(Q_2(Q_1^{\text{mod}}(\bar{A}(\nu))))|) \) for the number of basic computational operations needed for the evaluation of the algorithm \( \hat{S}^{\text{mod}}_{n,\varepsilon} \) at the input \((x_0, a, b)\). Using Property 3)(i) of \( Q_2 \circ Q_1^{\text{mod}} \) and (5.2) we conclude that
\[ \text{cost}^{\text{total}}(\hat{S}^{\text{mod}}_{n,\varepsilon}) \leq c \cdot n^{d+1}(1+\varepsilon) \max(1,s/r). \]

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 \(|V| \geq n d^2 s m^d \) 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 \( * = \text{supp} \).

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 \( * = \text{supp} \) 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 \( B, \Lambda \), 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 \( f : \mathbb{R}^d \to \mathbb{R} \) to be specified. Moreover, \( q \in \{ r, s \} \).

6.1. **Weighted integration on \( \mathbb{R}^d \).** Let \( w : \mathbb{R}^d \to \mathbb{R} \) 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 \( 0 \not\in R \) such that

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

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

\[
I_w(h) = \int_{\mathbb{R}^d} h(x) \cdot w(x) \, dx
\]

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

For \( N \in \mathbb{N} \) and functions \( \eta_i : \mathbb{R}^{i-1} \to \mathbb{R}^d \) for \( i = 1, \ldots, N \) define \( R_{\eta_1, \ldots, \eta_N} : H \to \mathbb{R}^N \) by

\[
R_{\eta_1, \ldots, \eta_N}(h) = (y_1, \ldots, y_N)
\]

with \( y_1 = h(\eta_1) \) and

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

for \( \ell = 2, \ldots, N \). Put

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

Then

\[
\tau_N^w(H) = \inf_{\tilde{I} \in \mathcal{I}_N(H)} \sup_{h \in H} |I_w(h) - \tilde{I}(h)|
\]

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 [11, 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 $h_1, \ldots, h_{2N} : \mathbb{R}^d \to [0, +\infty)$ such that

(a) the sets $\{h_i \neq 0\}$ are pairwise disjoint and contained in $R$,
(b) for all $\sigma_1, \ldots, \sigma_{2N} \in \{-1, 1\}$
\[
\sum_{i=1}^{2N} \sigma_i \cdot h_i \in H,
\]
(c) for every $i = 1, \ldots, 2N$
\[
\int_R h_i(x) \, dx \geq \varepsilon.
\]

Then
\[
\tau_N^w(H) \geq c \cdot N \cdot \varepsilon.
\]

For every $N \in \mathbb{N}$ we take functions $h_1, \ldots, h_{2N} \in F_0(d, q, K)$ 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

\[
\|h_i^{(\alpha)}\| \leq \frac{K}{N(q-\|\alpha\|_1)/d}
\]

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

Let
\[
H_N = \left\{ \sum_{i=1}^{2N} \sigma_i \cdot h_i \mid \sigma_1, \ldots, \sigma_{2N} \in \{-1, 1\} \right\},
\]
and note that by (6.2),
\[
H_N \subset F_0(d, q, K).
\]

Theorem 6.1 yields
\[
\tau_N^w(H_N) \geq c \cdot N^{-q/d}.
\]

6.2. Proof of the lower bound in Theorem 4.1 for $* = \text{supp}$. 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 $(0, 0, K \cdot E_d) \in C(d, s_1, s_2, K)$ and $H_N \subset F(d, r, K, \beta)$. Hence
\[
c_N^{\text{supp}}(C(d, s_1, s_2, K), F(d, r, K, \beta)) \geq c_N^{\text{supp}}(\{(0, 0, K \cdot E_d)\}, H_N).
\]

Consider an algorithm $\hat{S} : C(d, s_1, s_2, K) \to M_0(\mathbb{R}^d)$ that satisfies $\text{cost}^\text{supp}(\hat{S}) \leq N$. Then
\[
\hat{I} : H_N \to \mathbb{R}, \quad h \mapsto \int_{\mathbb{R}^d} h \, d\hat{S}(0, 0, K \cdot E_d),
\]
defines an algorithm $\tilde{T} \in I_N(H_N)$ and we have
\[ \rho_{H_N}(S(0,0,K \cdot E_d), \tilde{S}(0,0,K \cdot E_d)) = \sup_{h \in H_N} \left| \int_{\mathbb{R}^d} h(x) \cdot w(x) \, dx - \tilde{T}(h) \right|. \]
Thus
\[ e_N^{\text{supp}}(\{(0,0,K \cdot E_d)\}, H_N) \geq e_N^w(H_N) \geq c \cdot N^{-r/d} \]
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
\[ x_0 \in \mathbb{R}^d, a: \mathbb{R}^d \rightarrow \mathbb{R}^d, b: \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}, \]
and we put
\[ \sigma = bb^T. \]
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
  \[ \inf_{x \in \mathbb{R}^d} \lambda_{\text{min}}(\sigma(x)) \geq \Lambda. \]
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
\[ (6.5) \quad \frac{1}{2} \sum_{i,j=1}^d \sigma_{i,j}(x) \cdot \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^d a_i(x) \cdot \frac{\partial u}{\partial x_i}(t,x) + \frac{\partial u}{\partial t}(t,x) = 0 \]
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
\[ G: V \rightarrow [0,\infty) \]
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 \( \varphi_\Sigma \) denote the Lebesgue density of the centered \( d \)-dimensional normal distribution with positive definite covariance matrix \( \Sigma \in \mathbb{R}^{d \times d} \). Define
\[ Z : V \to \mathbb{R}, \quad (t, x, \tau, \xi) \mapsto \varphi_{\tau-t}\sigma(\xi)(x - \xi), \]
as well as a sequence of functions \( \Phi_m : V \to \mathbb{R} \) 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 \( G_m : V \to \mathbb{R} \) 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 \( \Phi_m \) 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
\[ (6.6) \quad G = \sum_{m=0}^{\infty} G_m. \]

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

We use (6.6) to obtain a series representation of
\[ (6.7) \quad \int_{\mathbb{R}^d} f \, dS(x_0, a, b) = \int_{\mathbb{R}^d} f(\xi) \cdot G(0, x_0, 1, \xi) \, d\xi \]
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

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

(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

\[(6.8) \quad \frac{\partial}{\partial \xi_1} f \neq 0.\]

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,

$$e^f_N(\mathcal{C}^{(1)}) \geq e^f_N(\mathcal{C}_N).$$

It thus suffices to show that

\[(6.9) \quad e^f_N(\mathcal{C}_N) \geq c \cdot N^{-s/d}\]

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

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

see (6.2), and

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

Moreover, the functions $Z$ and $\Phi_1$ are given by

\[(6.10) \quad Z(t, x, \tau, \xi) = \varphi_{K^2(\tau-t)}E_d(x - \xi)\]

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

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

**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. \(\square\)

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

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

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

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

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

\[
Z(0, 0, t, y) \leq \frac{1}{\|y\|_2^d} \cdot \frac{\|y\|_2^2 \cdot \exp\left(-\frac{\|y\|_2^2}{2K^2 t}\right)}{(2K^2 t)^{d/2}} \leq c \cdot \frac{1}{\|y\|_2^2}.
\]

Hence

\[
\int_{(0,1) \times \mathbb{R}^d} \left| \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0, 0, t, y) \cdot Z(t, y, 1, \xi) \right| \, d(t, \xi)
\]

\[
\leq \frac{c}{\|y\|_2^2} \int_{(0,1) \times \mathbb{R}^d} \left(1 + \|\xi\|^2 \cdot \varphi_{K^2(1-t)}(y - \xi) \right) \, d(t, \xi)
\]

\[
= \frac{c}{\|y\|_2^2} \int_{(0,1) \times \mathbb{R}^d} \left(1 + \|K(1-t)^{1/2} \xi + y\|^2 \cdot \varphi_{E_d}(\xi) \right) \, d(t, \xi)
\]

\[
\leq c \cdot \frac{1 + \|y\|_2^2}{\|y\|_2^2},
\]

which shows that \(w\) is well-defined.
In order to prove continuity of $w$ note that

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

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

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

see (6.12). Moreover,

$$\sup_{t \in (0,1)} \sup_{\|y - y_0\|_2 \leq \|y_0\|_2/2} \left| \frac{\partial}{\partial \xi_1} f((1-t)^{1/2} \xi + y) \right| \leq K (1 + \|(1-t)^{1/2} \xi + y\|^\beta_2) \leq c(y_0) (1 + \|\xi\|^\beta_2).$$

Hence

$$\int_{(0,1) \times \mathbb{R}^d} \sup_{\|y - y_0\|_2 \leq \|y_0\|_2/2} \left| \frac{\partial}{\partial \xi_1} f((1-t)^{1/2} \xi + y) \right| Z(0,0,t,y) \cdot Z(0,0,1,\xi) \, d(t,\xi)$$

$$\leq c(y_0) \int_{(0,1) \times \mathbb{R}^d} (1 + \|\xi\|^\beta_2) \cdot \varphi_{K^2} (\xi) \, d(t,\xi) < \infty.$$

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 $(k_1, \ldots, k_d) \in \mathbb{N}_0^d$ and consider the function $g: \mathbb{R}^d \setminus \{0\} \times (0,1) \times \mathbb{R}^d \to \mathbb{R}$ given by

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

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

$$\int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} |g(y,t,\xi)| \, d(y,t,\xi)$$

$$\leq K \int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} \|y\|_2^k \cdot (1 + \|\xi\|^\beta_2) \cdot Z(0,0,t,y) \cdot Z(t,y,1,\xi) \, d(y,t,\xi)$$

$$= K \int_{\mathbb{R}^d \times (0,1)} \|y\|_2^k \cdot Z(0,0,t,y) \int_{\mathbb{R}^d} (1 + \|(1-t)^{1/2} \xi + y\|^\beta_2) \cdot Z(0,0,1,\xi) \, d\xi \, d(y,t)$$

$$\leq c \int_{\mathbb{R}^d \times (0,1)} (1 + \|y\|_2^{k+\beta_2}) \cdot Z(0,0,t,y) \, d(y,t) \leq c(k).$$

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

\[
\int_{\mathbb{R}^d} \left( \prod_{j=1}^d y_j^{k_j} \right) \cdot w(y) \, dy = \int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0, 0, 1, \xi) \cdot \psi(\xi) \, d\xi,
\]

where

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

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

\[
\psi(\xi) = \int_{(0,1) \times \mathbb{R}^d} \left( \prod_{j=1}^d \left( K \sqrt{t(1-t)} y_j + t\xi_j \right)^{k_j} \right) \cdot \varphi_{E_d}(y) \, d(t,y)
\]

\[
= \int_{(0,1)} \prod_{j=1}^d \sum_{\ell_j=0}^{k_j} \binom{k_j}{\ell_j} \left( K \sqrt{t(1-t)} \right)^{\ell_j} q_{\ell_j}(t\xi_j)^{k_j-\ell_j} \, dt
\]

\[
= \sum_{\ell_1=0}^{k_1} \cdots \sum_{\ell_d=0}^{k_d} \left( \prod_{j=1}^d \binom{k_j}{\ell_j} K^{\ell_j} q_{\ell_j}(\xi_j)^{k_j-\ell_j} \right) \cdot \int_{(0,1)} \prod_{j=1}^d (1-t)^{\ell_j/2} t^{k_j-\ell_j/2} \, dt.
\]

Hence \( \psi \) is a polynomial of degree \((k_1, \ldots, k_d)\). Assume \( w = 0 \). Then (6.15) implies

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

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 \( k_1 = \cdots = k_d = 0 \). Since \( a_1 \) is bounded and \( g \) is integrable we have

\[
\int_{\mathbb{R}^d} a_1(y) \cdot w(y) \, dy = \int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} a_1(y) \cdot g(y, t, \xi) \, d(y, t, \xi)
\]

\[
= \int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} a_1(y) \cdot \frac{\partial}{\partial \xi_1} f(\xi) \cdot Z(0, 0, t, y) \cdot Z(t, y, 1, \xi) \, d(y, t, \xi).
\]

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 \( \xi_2, \ldots, \xi_d \in \mathbb{R} \). 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
\[
\int_{\mathbb{R}^d} a_1(y) \cdot w(y) \, dy = \int_{\mathbb{R}^d} f(\xi) \int_{(0,1) \times \mathbb{R}^d} a_1(y) \cdot Z(0, 0, t, y) \cdot \frac{\partial}{\partial y_1} Z(t, y, 1, \xi) \, dt \, d\xi
\]
\[
= \int_{\mathbb{R}^d} f(\xi) \cdot G_1(0, 0, 1, \xi) \, d\xi = J_1(f, a_1),
\]
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 $\hat{A} \in \mathfrak{A}$. Let $a_1 \in H_N$. By Lemma 6.1(ii), Lemma 6.2 and Lemma 6.3,
\[
|\hat{A}(0, (a_1, 0, \ldots, 0), K \cdot E_d) - A^f(0, (a_1, 0, \ldots, 0), K \cdot E_d)|
\]
\[
= |\hat{A}(0, (a_1, 0, \ldots, 0), K \cdot E_d) - \sum_{m=0}^{\infty} J_m(f, a_1)|
\]
\[
\geq |\hat{A}(0, (a_1, 0, \ldots, 0), K \cdot E_d) - \sum_{m=0}^{1} J_m(f, a_1)| - c \cdot N^{-2s/d}
\]
\[
= |\hat{A}(0, (a_1, 0, \ldots, 0), K \cdot E_d) - \int_{\mathbb{R}^d} f(\xi) \cdot Z(0, 0, 1, \xi) \, d\xi
\]
\[
- \int_{\mathbb{R}^d} a_1(y) \cdot w(y) \, dy| - c \cdot N^{-2s/d}.
\]
If $\hat{A}$ satisfies $\text{cost}(\hat{A}) \leq N$ then
\[
\hat{I}: H_N \rightarrow \mathbb{R}, \quad h \mapsto \hat{A}(0, (h, 0, \ldots, 0), K \cdot E_d) - \int_{\mathbb{R}^d} f(\xi) \cdot Z(0, 0, 1, \xi) \, d\xi
\]
defines an algorithm $\hat{I} \in \mathcal{I}_N(H_N)$ and we may conclude that
\[
\sup_{(x_0, a, b) \in C_N} |\hat{A}(x_0, a, b) - A^f(x_0, a, b)| \geq \sup_{h \in H_N} |\hat{I}(h) - I^w(h)| - c \cdot N^{-2s/d}.
\]
Consequently,
\[
e^f_N(C_N) \geq \tau^w_N(H_N) - c \cdot N^{-2s/d}.
\]
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
\[
\frac{\partial^2}{\partial \xi_1^2} f \neq 0.
\]
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

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

We consider the classes of SDEs

$$C_N = \{(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) \mid b_1 \in \tilde{H}_N\}, \quad N \in \mathbb{N}.$$ 

Clearly,

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

It thus suffices to show that

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

for an appropriately chosen rectangle $R$.

Note that for $(x_0, a, b) \in C_N$ 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 $\Phi_1$ are given by

(6.18) $$Z(t, x, \tau, \xi) = \frac{1}{(2\pi(\tau-t))^{d/2}(K/2)^{d-1} \cdot b_1(\xi)} 
\cdot \exp\left(-\frac{1}{2(\tau-t)}\left(\frac{(x_1-\xi_1)^2}{b_2^2(\xi)} + \sum_{i=2}^{d} \frac{(x_i-\xi_i)^2}{K^2/4}\right)\right)$$

and

(6.19) $$\Phi_1(t, x, \tau, \xi) = \frac{b_2^2(x) - b_2^2(\xi)}{2} \cdot \frac{\partial^2}{\partial x_1^2} Z(t, x, \tau, \xi) \cdot \frac{\partial^2}{\partial x_1^2} Z(t, x, \tau, \xi)$$

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 $b_1 \in \tilde{H}_N$. 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 negligible, asymptotically.

Lemma 6.4. We have

$$\forall \gamma \in (0, 1] \forall N \in \mathbb{N} \forall b_1 \in \tilde{H}_N: \sum_{m=2}^{\infty} |J_m(f, b_1)| \leq c(\gamma) \cdot N^{-2(\gamma-s)/d}.$$
Proof. Let $\gamma \in (0, 1]$, $N \in \mathbb{N}$ and $b_1 \in \tilde{H}_N$. Apply Lemma 6.1 to get

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

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

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

which completes the proof. □

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

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

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

**Lemma 6.5.** We have

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

**Proof.** Straightforward calculations yield

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

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

Hence

$$(6.20) \sup_{v \in [K/4, K]} \left(\left|\frac{\partial}{\partial v} T_0(\xi, v)\right| + \left|\frac{\partial^2}{\partial v^2} T_0(\xi, v)\right|\right) \leq c \cdot (1 + \|\xi\|_2^4) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2}\right)$$
for all $\xi \in \mathbb{R}^d$. Observing (2.18) as well as $4b_1 - 2K \in H_N$ it follows

\[ |J_0(f, b_1) - J_{0,0}(f) - J_{0,1}(f, b_1)| \]

\[ = \left| \int_{\mathbb{R}^d} f(\xi) \int_{K/2}^{b_1(\xi)} (b_1(\xi) - \nu) \cdot \frac{\partial^2}{\partial\nu^2} T_0(\xi, \nu) \, d\nu \, d\xi \right| \]

\[ \leq c \cdot \int_{\mathbb{R}^d} |f(\xi)| \cdot (1 + \|\xi\|_2^4) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2}\right) (b_1(\xi) - K/2)^2 \, d\xi \]

\[ \leq c \cdot \|b_1 - K/2\|^2 \cdot \int_{\mathbb{R}^d} (1 + \|\xi\|_2^{5+\beta}) \cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2}\right) \, d\xi \]

\[ \leq c \cdot N^{-2s/d}, \]

which completes the proof. \qed

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

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

by

\[ T_1(t, y, \xi, u, v) = \frac{1}{2(2\pi)^d(K/2)^{2(d-1)} \cdot (t(1-t))^d/2} \cdot \left( \frac{1}{u} + \frac{1}{v} \right) \cdot \left( \frac{(\xi_1 - y_1)^2}{(1-t)^2 \cdot v^2} - \frac{1}{(1-t) \cdot v^2} \right) \cdot \exp\left(-\frac{1}{2t} \left( \frac{y_1^2}{u^2} + \sum_{i=2}^{d} \frac{y_i^2}{K^2/4} \right) - \frac{1}{2(1-t)} \left( \frac{(\xi_1 - y_1)^2}{v^2} + \sum_{i=2}^{d} \frac{(\xi_i - y_i)^2}{K^2/4} \right) \right). \]

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 $b_1 \in \tilde{H}_N$.

**Lemma 6.6.** We have

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

**Proof.** By definition,

\[ T_1(t, y, \xi, K/2, K/2) = \frac{1}{(K/2)^2 \cdot (1-t) \cdot (1-t) - 1} \cdot \varphi(K/2)^2(1-t) \cdot E_d(\xi - y) \cdot \varphi(K/2)^2(1-t) \cdot E_d(y) \]
and therefore

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

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

\[ \int_{(0,1) \times \mathbb{R}^d} \|\xi - y\|_2 \cdot |T_1(t, y, \xi, K/2, K/2)| \, d(t, y) \]

\[ \leq c \cdot \int_{(0,1) \times \mathbb{R}^d} \frac{1}{td/(1-t)^{d/2+1}} \cdot \exp\left( -\frac{\|y\|_2^2}{K^2t} - \frac{\|\xi - y\|_2^2}{K^2(1-t)} \right) \, d(t, y) \]

\[ = c \cdot \exp\left( -\frac{\|\xi\|_2^2}{K^2} \right) \]

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

For $b_1 \in \tilde{H}_N$ we have $|b_1(y) - b_1(\xi)| \leq K\|\xi - y\|_2$. 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

\[ \forall N \in \mathbb{N} \forall b_1 \in \tilde{H}_N : \ |J_1(f, b_1) - J_{1,0}(f, b_1)| \leq c \cdot N^{-(2s-1/2)/d}. \]

**Proof.** Straightforward calculations yield

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

and

\[ \left| \frac{\partial}{\partial v} T_1(t, y, \xi, u, v) \right| \leq \frac{c}{td/(1-t)^{d/2+1}} \cdot \left( \frac{\|\xi - y\|_2^3}{(1-t)^2} + 1 \right) \cdot \exp\left( -\frac{\|y\|_2^2}{2K^2t} - \frac{\|\xi - y\|_2^2}{2K^2(1-t)} \right) \]
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
\[
|b_1(y) - b_1(\xi)| \cdot |T_1(t, y, \xi, b_1(y), b_1(\xi)) - T_1(t, y, \xi, K/2, K/2)|
\leq c \cdot N^{-(2s-1)/(2d)} \cdot \frac{\|\xi - y\|^2/2}{t^{d/2}(1 - t)^{d/2+1}} \cdot \left(\frac{\|\xi - y\|_2^3}{t} + 1\right) \left(\frac{\|\xi - y\|_2^4}{(1 - t)^2} + 1\right)
\cdot \exp\left(-\frac{\|\xi\|_2^2}{2K^2t} - \frac{\|\xi - y\|_2^2}{2K^2(1 - t)}\right) \cdot (|b_1(\xi) - K/2| + |b_1(y) - K/2|)
\leq c \cdot N^{-(2s-1)/(2d)} \cdot \frac{1}{t^{d/2}(1 - t)^{d/2+3/4}} \cdot \exp\left(-\frac{\|\xi\|_2^2}{4K^2} - \frac{\|\xi - y\|_2^2}{4K^2(1 - t)}\right)
\] for all \(y, \xi \in \mathbb{R}^d\) and \(t \in (0, 1)\). Using \[5, \text{p.15, Lemma 3}\] we conclude that
\[
\int_{(0,1) \times \mathbb{R}^d} |b_1(y) - b_1(\xi)| \cdot |T_1(t, y, \xi, b_1(y), b_1(\xi)) - T_1(t, y, \xi, K/2, K/2)| d(t, y)
\leq c \cdot N^{-(2s-1)/(2d)} \cdot \exp\left(-\frac{\|\xi\|_2^2}{4K^2}\right)
\] for every \(\xi \in \mathbb{R}^d\). Hence
\[
|J_1(f, b_1) - J_{1,0}(f, b_1)| \leq c \cdot N^{-(2s-1)/(2d)} \cdot \int_{\mathbb{R}^d} |f(\xi)| \cdot \exp\left(-\frac{\|\xi\|_2^2}{4K^2}\right) d\xi
\] and it remains to observe (2.2) to complete the proof. \(\square\)

Define \(w: \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}\) by
\[
w(y) = \frac{K}{2} \int_{(0,1) \times \mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi(K/2)^2(1-t) \cdot E_d(\xi - y) \cdot \varphi(K/2)^2t \cdot E_d(y) d(t, \xi).
\]

**Lemma 6.8.** The mapping \(w\) is well-defined, continuous and non-zero and satisfies
\[
J_{0,1}(f, b_1) + J_{1,0}(f, b_1) = \int_{\mathbb{R}^d} (b_1(y) - K/2) \cdot w(y) dy
\] for all \(b_1 \in \tilde{H}_N\) and \(N \in \mathbb{N}\).

**Proof.** We have
\[
\int_{\mathbb{R}^d} \left|\frac{\partial^2}{\partial \xi_1^2} f(\xi)\right| \cdot \varphi(K/2)^2(1-t) \cdot E_d(\xi - y) d\xi
\leq K \int_{\mathbb{R}^d} (1 + \|(K/2)(1-t)^{1/2} \cdot \xi + y\|_2^\beta) \cdot \varphi E_d(\xi) d\xi \leq c \cdot (1 + \|y\|_2^\beta)
\] for all \(t \in (0, 1)\) and \(y \in \mathbb{R}^d\). Similar to (6.12),
\[
\varphi(K/2)^2t \cdot E_d(y) \leq \frac{c}{\|y\|_2^2}
\] 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

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

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

\[
\sup_{\|y-y_0\|_2 \leq \|y_0\|_2 / 2} \left| \frac{\partial^2}{\partial \xi_1^2} f(K/2(1-t)^{1/2} \xi + y) \right| \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) \\
\leq c \cdot \sup_{\|y-y_0\|_2 \leq \|y_0\|_2 / 2} (1 + \| (K/2) (1-t)^{1/2} \xi + y \|_2^2) \cdot \| y \|_2^d \\
\leq c \cdot (1 + \| y_0 \|_2^2 + \| \xi \|_2^2) \cdot \| y_0 \|_2^{-d}
\]

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

\[
\int_{(0,1) \times \mathbb{R}^d} \sup_{\|y-y_0\|_2 \leq \|y_0\|_2 / 2} \left| \frac{\partial^2}{\partial \xi_1^2} f(K/2(1-t)^{1/2} \xi + y) \right| \cdot \varphi_{(K/2)^2 t \cdot E_d}(\xi) \cdot \varphi_{E_d}(y) \, d(t, \xi) \\
\leq c(y_0) \int_{(0,1) \times \mathbb{R}^d} (1 + \| \xi \|_2^2) \cdot \varphi_{E_d}(\xi) \, d(t, \xi) \\
\leq c(y_0).
\]

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

\[
(6.27) \quad g(y, t, \xi) = \left( \prod_{j=1}^d y_j^{k_j} \right) \cdot \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2(1-t) \cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y).
\]

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

\[
\int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} |g(y, t, \xi)| \, d(y, t, \xi) \\
\leq c \int_{\mathbb{R}^d \times (0,1)} \| y \|_2^k \cdot (1 + \| y \|_2^2) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) \, d(y, t) \\
\leq c \int_{\mathbb{R}^d \times (0,1)} (1 + \| (K/2) t^1/2 \|_2^{k+1}) \cdot \varphi_{E_d}(y) \, d(y, t) \\
\leq c(k).
\]

Hence \( g \) is integrable. Using

\[
(6.28) \quad \varphi_{(K/2)^2(1-t) \cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 t \cdot E_d}(y) = \varphi_{(K/2)^2(1-t) \cdot E_d}(y - t\xi) \cdot \varphi_{(K/2)^2 \cdot E_d}(\xi)
\]
we obtain
\[
\int_{\mathbb{R}^d} \left( \prod_{j=1}^d y_j^{k_j} \right) \cdot w(y) \, dy = \frac{K}{2} \int_{\mathbb{R}^d} \frac{\partial^2}{\partial \xi_1^2} f(\xi) \cdot \varphi_{(K/2)^2 \cdot E_d}(\xi) \cdot \psi(\xi) \, d\xi,
\]
where
\[
\psi(\xi) = \int_{(0,1) \times \mathbb{R}^d} \left( \prod_{j=1}^d y_j^{k_j} \right) \varphi(\xi/2) \cdot (1-t) \cdot E_d(y-t\xi) \, d(t, y).
\]

The mapping \( \psi \) is a polynomial of degree \((k_1, \ldots, k_d)\), 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 \( b_1 \in \tilde{H}_N \) 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
\[
\int_{\mathbb{R}^d} (b_1(y) - K/2) \cdot w(y) \, dy = \int_{\mathbb{R}^d \times (0,1) \times \mathbb{R}^d} (b_1(y) - K/2) \cdot g(y,t,\xi) \, d(y,t,\xi)
\]
\[
= \lim_{\eta \to 1} \int_{\mathbb{R}^d \times (0,\eta) \times \mathbb{R}^d} (b_1(y) - K/2) \cdot g(y,t,\xi) \, d(y,t,\xi).
\]

Let \( \eta \in (0,1) \). Since \( f \in \mathcal{F}(d, 2, K, \beta) \) we have
\[
\lim_{\|\xi\| \to \infty} (|f(\xi)| + |\partial f(\xi)/\partial \xi_1|) \cdot \|\xi - y\|^{m} \cdot \exp\left(-\frac{\|\xi - y\|^2}{2(K/2)^2 \cdot (1-t)}\right) = 0
\]
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,
\[
\int_{\mathbb{R}^d} (K/2) \cdot g(y,t,\xi) \, d\xi
\]
\[
= \int_{\mathbb{R}^d} \frac{\partial}{\partial \xi_1} f(\xi) \cdot \frac{\xi_1 - y_1}{(K/2)(1-t)} \cdot \varphi_{(K/2)^2 \cdot (1-t) \cdot E_d}(\xi - y) \cdot \varphi_{(K/2)^2 \cdot E_d}(y) \, d\xi
\]
\[
= \int_{\mathbb{R}^d} f(\xi) \cdot T_1(t,y,\xi,K/2,K/2) \, d\xi
\]
for every $y \in \mathbb{R}^d \setminus \{0\}$ and every $t \in (0, \eta)$ and we obtain
\[
\int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} (b_1(y) - K/2) \cdot (K/2) \cdot g(y, t, \xi) \, d(y, t, \xi)
= \int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} (b_1(y) - K/2) \cdot f(\xi) \cdot T_1(t, y, \xi, K/2, K/2) \, d(y, t, \xi)
= \int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} (b_1(y) - b_1(\xi)) \cdot f(\xi) \cdot T_1(t, y, \xi, K/2, K/2) \, d(y, t, \xi)
+ \int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} (b_1(\xi) - K/2) \cdot f(\xi) \cdot T_1(t, y, \xi, K/2, K/2) \, d(y, t, \xi).
\]
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
\[
(6.29) \quad \lim_{\eta \to 1} \int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} h_\eta(y, t, \xi) \, d(y, t, \xi) = J_{0,1}(f, b_1),
\]
where $h_\eta : \mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d \to \mathbb{R}$ 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),
\[
|h_\eta(y, t, \xi)| \leq c(\eta) \cdot (1 + \|y\|_2^{1+\beta} + \|\xi - y\|_2^{3+\beta}) \cdot \varphi(K/2)^2(1-t) \cdot E_d(\xi - y) \cdot \varphi(K/2)^2t \cdot E_d(y)
\]
for all $y, \xi \in \mathbb{R}^d$ and $t \in (0, \eta)$. Hence $h_\eta$ is integrable. Observing (6.28) we conclude that
\[
\int_{\mathbb{R}^d \times (0, \eta) \times \mathbb{R}^d} h_\eta(y, t, \xi) \, d(y, t, \xi)
= \int_{\mathbb{R}^d} (b_1(\xi) - K/2) \cdot f(\xi) \cdot \varphi(K/2)^2 \cdot E_d(\xi) \int_{(0, \eta)} 1 \, (K/2)(1-t)
\cdot \int_{\mathbb{R}^d} \left( \frac{(\xi_1 - y_1)^2}{(K/2)^2} - 1 \right) \varphi(K/2)^2t(1-t) \cdot E_d(y - t\xi) \, dy \, dt \, d\xi
= \int_{\mathbb{R}^d} (b_1(\xi) - K/2) \cdot f(\xi) \cdot \varphi(K/2)^2 \cdot E_d(\xi) \int_{(0, \eta)} \frac{1}{K/2} \cdot \left( \frac{\xi_1^2}{(K/2)^2} - 1 \right) \, dt \, d\xi
= \eta \int_{\mathbb{R}^d} (b_1(\xi) - K/2) \cdot f(\xi) \cdot \frac{\partial}{\partial \xi} T_0(\xi, K/2) \, d\xi = \eta \cdot J_{0,1}(f, b_1),
\]
which yields (6.29) and hereby completes the proof.\qed
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 $\hat{A} \in \mathcal{A}$. Let $b_1 \in \tilde{H}_N$. By Lemma 6.1(ii), Lemma 6.4 with $\gamma = 1/4$, Lemmas 6.5, 6.7 and by (6.23) we get

$$\left| \hat{A}(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) - A^f(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) \right|
\geq \left| \hat{A}(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) - \sum_{m=0}^{\infty} J_m(f, b_1) \right|
\geq \left| \hat{A}(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) - \sum_{m=0}^{1} J_m(f, b_1) \right| - \frac{c}{N^{3s/(2d)}}
\geq \left| \hat{A}(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) - J_{0,0}(f) - J_{0,1}(f, b_1) - J_{1,0}(f, b_1) \right|
- \frac{c}{N^{3s/(2d)}}
\geq \left| \hat{A}(0, 0, \text{diag}(b_1, K/2, \ldots, K/2)) - J_{0,0}(f) - \int_{\mathbb{R}^d} (b_1(y) - K/2) \cdot w(y) \, dy \right|
- \frac{c}{N^{3s/(2d)}}.$$

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

$$\hat{I} : H_N \rightarrow \mathbb{R}, \quad h \mapsto \hat{A}(0, 0, \text{diag}(K/2 + h/4, K/2, \ldots, K/2)) - J_{0,0}(f)$$

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

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

Consequently,

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

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

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

$$e_N(\mathcal{C}_N) \geq \varepsilon^{w/4}_N(H_N) - c \cdot N^{-3s/(2d)},$$
and let $X$ and $\tilde{X}$ denote the strong solutions of the respective SDEs driven by $W$. Then

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

for all $p, q \geq 1$ 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 $\kappa > 0$, $p \geq 2$ and $q \geq 1$. 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 $\kappa, p$ and $q$. Let $(x_0, a, b, (\tilde{x}_0, \tilde{a}, \tilde{b}) \in [-\kappa, \kappa]^d \times (\mathcal{L}^d(d, \kappa))^d \times (\mathcal{L}^d(d, \kappa))^d$.

Put

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

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 $\tilde{q} \geq 2$. By the properties of $a, \tilde{a}, b, \tilde{b}$, for $t \in [0, 1]$ and any $\mathbb{R}^d$-valued stochastic process $Z = (Z(v))_{v \in [0,1]}$ with continuous paths, we have

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

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

$$\left(\mathbb{E}\left((Y^*(1))^{2p} \cdot \mathbf{1}_{\{\tau \leq 1}\}\right)\right)^{1/p} \leq \left(\mathbb{E}(Y^*(1))^{2p}\right)^{1/(2p)} \cdot \left(\mathbb{P}(\tau \leq 1)^{1/(2p)}\right) \leq c \cdot \left(1 + \|x_0\|^{q+1} + \|\tilde{x}_0\|^{q+1}\right) \cdot u^{-q}.$$

Note that

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

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

$$\mathbb{E}(Y^*(\tau \wedge 1))^{p} \leq c \cdot (\|x_0 - \tilde{x}\|^{p} + \|a - \tilde{a}\|_{[-u,u]^d}^{p} + \|b - \tilde{b}\|_{[-u,u]^d}^{p}).$$

Clearly,

$$Y(\tau \wedge t) = x_0 - \tilde{x}_0 + \int_{0}^{\tau \wedge t} (a(X(v)) - \tilde{a}(\tilde{X}(v))) \, dv
+ \int_{0}^{\tau \wedge t} (b(X(v)) - \tilde{b}(\tilde{X}(v))) \, dW(v)$$
for every $t \in [0, 1]$. By the Hölder inequality and the Lipschitz continuity of $a$ and $\tilde{a}$ we obtain
\[
\left\| \int_0^{\tau \wedge T} (a(X(v)) - \tilde{a}(X(v))) \, dv \right\|^p
\leq (\tau \wedge t)^{p-1} \cdot \int_0^{\tau \wedge T} \|a(X(v)) - \tilde{a}(X(v))\|^p \, dv
\leq c \cdot \left( \|a - \tilde{a}\|_{[-u,u]^d}^p + \kappa \cdot \int_0^{\tau \wedge T} \|X(v) - \tilde{X}(v)\|^p \, dv \right),
\]
which implies
\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} \left\| \int_0^{\tau \wedge T} (a(X(v)) - \tilde{a}(X(v))) \, dv \right\|^p \right)
\leq c \cdot \left( \|a - \tilde{a}\|_{[-u,u]^d}^p + \mathbb{E} \left( \int_0^{\tau \wedge T} \|X(v) - \tilde{X}(v)\|^p \, dv \right) \right)
\leq c \cdot \left( \|a - \tilde{a}\|_{[-u,u]^d}^p + \int_0^T \mathbb{E} \left( Y^*(\tau \wedge v) \right)^p \, dv \right)
\]
for every $T \in [0, 1]$. Similarly, using the Burkholder-Davis-Gundy inequality and the Lipschitz continuity of $b$ and $\tilde{b}$ we get
\[
\mathbb{E} \left( \sup_{0 \leq t \leq T} \left\| \int_0^{\tau \wedge T} (b(X(v)) - \tilde{b}(X(v))) \, dW(v) \right\|^p \right)
\leq c \cdot \mathbb{E} \left( \int_0^{\tau \wedge T} \|b(X(v)) - \tilde{b}(X(v))\|^2 \, dv \right)^{p/2}
\leq c \cdot \left( \|b - \tilde{b}\|_{[-u,u]^d}^p + \int_0^T \mathbb{E} \left( Y^*(\tau \wedge v) \right)^p \, dv \right)
\]
for every $T \in [0, 1]$. Hence
\[
\mathbb{E} \left( Y^*(\tau \wedge T) \right)^p \leq c \cdot \left( \|x_0 - \tilde{x}_0\|^p + \|a - \tilde{a}\|_{[-u,u]^d}^p + \|b - \tilde{b}\|_{[-u,u]^d}^p + \int_0^T \mathbb{E} \left( Y^*(\tau \wedge v) \right)^p \, dv \right)
\]
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 $\Phi_m$. Similar estimates are established in [7, 3] and [21], but without tracing their particular dependence on all of the quantities $\|a\|$, $\|b\|$ and $\|b\|_\star$. 

---

MÜLLER-GRONBACH, RITTER, AND YAROSLATSEVA
Lemma 7.2. The functions $\Phi_m$ satisfy

$$|\Phi_m(t, x, \tau, \xi)| \leq (c(\gamma))^m (4\pi d^2 B^2)^{(m-1)d/2} \frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2)} \cdot (\|a\| + \|b\| \|b\|_s^{1-\gamma})^{m} \cdot \exp\left(-\frac{\|x - \xi\|^2_2}{4d^2B^2(\tau - t)}\right)$$

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

$$Z(t, x, \tau, \xi) \leq \frac{1}{\Lambda^{d/2}(\tau - t)^{d/2}} \cdot \exp\left(-\frac{\|x - \xi\|^2_2}{2d^2B^2(\tau - t)}\right),$$

$$\left|\frac{\partial}{\partial x_i} Z(t, x, \tau, \xi)\right| \leq \frac{\|x - \xi\|^2_2}{\Lambda(\tau - t)} \cdot Z(t, x, \tau, \xi),$$

$$\left|\frac{\partial^2}{\partial x_i \partial x_j} Z(t, x, \tau, \xi)\right| \leq \left(\frac{\|x - \xi\|^2_2}{\Lambda^2(\tau - t)^2} + \frac{1}{\Lambda(\tau - t)}\right) \cdot Z(t, x, \tau, \xi)$$

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

$$|\sigma_{i,j}(x) - \sigma_{i,j}(\xi)| = \sum_{k=1}^{d} \left| (b_{i,k}(x)(b_{j,k}(x) - b_{j,k}(\xi)) + b_{j,k}(\xi)(b_{i,k}(x) - b_{i,k}(\xi))) \right|$$

$$\leq \sum_{k=1}^{d} 2\||b||d^{1/2}\|x - \xi\|_2\gamma||b||_s^{1-\gamma} \leq 2d^{3/2}B||b||^\gamma\|x - \xi\|_2^2||b||_s^{1-\gamma}.$$

Hence, by (7.5) and (7.6),

$$|\Phi_1(t, x, \tau, \xi)|$$

$$\leq \left(d^{7/2}B||b||^\gamma||b||_s^{1-\gamma} \left(\frac{\|x - \xi\|_2^{2+\gamma}}{\Lambda^2(\tau - t)^2} + \frac{\|x - \xi\|_2^2}{\Lambda(\tau - t)}\right) + d\|a\| \frac{\|x - \xi\|_2}{\Lambda(\tau - t)}\right) Z(t, x, \tau, \xi)$$

$$\leq \left(\|a\| + \|b\| \|b\|_s^{1-\gamma}\right) \frac{d^{7/2}B}{\Lambda^2} \left(\frac{\|x - \xi\|_2^{2+\gamma}}{\tau - t} + \frac{\|x - \xi\|_2^2}{\tau - t} + \frac{\|x - \xi\|_2}{\tau - t}\right) Z(t, x, \tau, \xi).$$

Put

$$u = \frac{\|x - \xi\|_2}{dB(\tau - t)^{1/2}}.$$
Lemma 7.3. The functions $G_m$ satisfy

$$|G_m(t,x,\tau,\xi)| \leq (c(\gamma))^m(\|a\| + \|b'\|\gamma\|b\|_{1-\gamma}^\gamma)^m m^{-m\gamma/4}$$

$$\cdot \frac{\Gamma((m + 1)\gamma/2)}{(\tau - t)^{1+d/2-(m+1)\gamma/2}} \exp\left(-\frac{\|x - \xi\|_2^2}{4d^2B^2(\tau - t)}\right)$$

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

We turn to estimates of the functions $G_m$. 

Using (7.4) we obtain

$$\left(\frac{\|x - \xi\|_2^{2+\gamma}}{(\tau - t)^2} + \frac{\|x - \xi\|_2^\gamma}{\tau - t} + \frac{\|x - \xi\|_2^\gamma}{\tau - t}\right) \cdot Z(t,x,\tau,\xi)$$

$$= \left(\frac{(dBu)^2 + (dBu)^\gamma}{(\tau - t)^{1-\gamma/2}} + \frac{dBu}{(\tau - t)^{1/2}}\right) \cdot Z(t,x,\tau,\xi)$$

$$\leq \frac{(dB)^2 + u^{2+\gamma} + u^\gamma + u}{\Lambda^{d/2}(\tau - t)^{d/2}} \cdot \exp(-u^2/2)$$

$$\leq \frac{(dB)^3 + \kappa(\gamma) + \kappa(1) + \kappa(2 + \gamma)}{(\tau - t)^{1-\gamma/2+d/2}} \cdot \exp(-u^2/4),$$

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

$$\int_{(t,\tau) \times \mathbb{R}^d} |\Phi_m(t,x,t',\xi) \cdot \Phi_1(t',y,\tau,\xi)| \, d(t',y)$$

$$\leq (c^*)^{m+1}(4\pi d^2B^2)^{(m-1)d/2} \frac{\Gamma((m+1)\gamma/2)}{\Gamma(m\gamma/2)} (\|a\| + \|b'\|\gamma\|b\|_{1-\gamma}^\gamma)^{m+1}$$

$$\cdot \int_{(t,\tau) \times \mathbb{R}^d} \frac{\exp\left(-\|x - y\|_2^2/(4d^2B^2(t' - t)) - \|\xi - y\|_2^2/(4d^2B^2(\tau - t'))\right)}{(t' - t)^{1+d/2-m\gamma/2}(\tau - t')^{1+d/2-\gamma/2}} \, d(t',y)$$

$$= (c^*)^{m+1}(4\pi d^2B^2)^{(m-1)d/2} \frac{\Gamma((m+1)\gamma/2)}{\Gamma(m\gamma/2)} (\|a\| + \|b'\|\gamma\|b\|_{1-\gamma}^\gamma)^{m+1}$$

$$\cdot \frac{(4\pi d^2B^2)^{d/2} \Gamma((m+1)\gamma/2) \Gamma(\gamma/2)}{\Gamma((m+1)\gamma/2) \Gamma(\gamma/2)} \exp\left(-\|x - \xi\|_2^2/(4d^2B^2(\tau - t))\right),$$

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

$$|\Phi_{m+1}(t,x,\tau,\xi)| \leq (c^*)^{m+1}(4\pi d^2B^2)^{md/2} \frac{\Gamma((m+1)\gamma/2)}{\Gamma((m+1)\gamma/2)} (\|a\| + \|b'\|\gamma\|b\|_{1-\gamma}^\gamma)^{m+1}$$

$$\cdot \frac{\exp\left(-\|x - \xi\|_2^2/(4d^2B^2(\tau - t))\right)}{(\tau - t)^{1+d/2-(m+1)\gamma/2}},$$

which completes the proof.
Proof. Fix $\gamma \in (0, 1]$ and $(t, x, \tau, \xi) \in V$. By \cite[Lemma 7.2 and p.15, Lemma 3]{[5]} we have

\[
\int_{(t, \tau) \times \mathbb{R}^d} |Z(t, x, t', y) \cdot \Phi_m(t', y, \tau, \xi)| \, d(t', y)
\]

\[
\leq \left( c(\gamma) \right)^m \frac{(\Gamma(\gamma/2))^m \left( \|a\| + \|b'\|^\gamma \|b\|_\gamma \right)^m}{\Gamma(m\gamma/2)} \cdot \int_{(t, \tau) \times \mathbb{R}^d} \frac{\exp\left( - \|x - y\|_2^2 / (4d^2B^2(t' - t)) - \|\xi - y\|_2^2 / (4d^2B^2(\tau - t')) \right)}{(t' - t)^{d/2} (\tau - t')^{1+d/2-m\gamma/2}} \, d(t', y)
\]

\[
= \left( c(\gamma) \right)^m \frac{(\Gamma(\gamma/2))^m \left( \|a\| + \|b'\|^\gamma \|b\|_\gamma \right)^m}{\Gamma(m\gamma/2)} \cdot (4\pi d^2B^2)^{d/2} \frac{\Gamma(m\gamma/2)}{\Gamma(m\gamma/2 + 1)} \cdot \frac{\exp\left( - \|x - \xi\|_2^2 / (4d^2B^2(\tau - t)) \right)}{(t - \tau)^{d/2-m\gamma/2}} \exp\left( - \frac{\|x - \xi\|_2^2}{4d^2B^2(\tau - t)} \right)
\]

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

\[
\Gamma(m\gamma/2 + 1) \geq \lfloor m\gamma/2 \rfloor! \geq c \cdot \left( \lfloor m\gamma/2 \rfloor / e \right)^{\lfloor m\gamma/2 \rfloor} \geq c \cdot \left( m\gamma / (4e) \right)^{m\gamma / (4e)} \geq (c(\gamma))^m m^{m\gamma/4}
\]

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

\[
\frac{(\Gamma(\gamma/2))^m}{\Gamma(m\gamma/2 + 1)} \leq (c(\gamma))^m m^{-m\gamma/4}
\]

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

\[
|G_m(t, x, \tau, \xi)| \leq \int_{(t, \tau) \times \mathbb{R}^d} |Z(t, x, t', y)\Phi_m(t', y, \tau, \xi)| \, d(t', y)
\]

\[
\leq (c(\gamma))^m \frac{(\|a\| + \|b'\|^\gamma \|b\|_\gamma)^m m^{-m\gamma/4}}{(\tau - t)^{(d-m\gamma)/2}} \cdot \exp\left( - \frac{\|x - \xi\|_2^2}{4d^2B^2(\tau - t)} \right)
\]

for all $m \in \mathbb{N}$, which finishes the proof. \qed
We proceed with the proof of Lemma 6.1. By (2.2) and (7.7) we obtain

\[
|J_m(f)| \leq \int_{\mathbb{R}^d} |f(\xi) \cdot G_m(0, x_0, 1, \xi)| \, d\xi \\
\leq \int_{\mathbb{R}^d} |f(\xi)| \int_{(0,1) \times \mathbb{R}^d} |Z(0, x_0, t', y) \cdot \Phi_m(t', y, 1, \xi)| \, d(t', y) \, d(\xi)
\]

(7.8)

\[
\leq c(f)(c(\gamma))m(\|a\| + \|b'\|^\gamma \|b\|^{1-\gamma})m_m^{-m\gamma/4} \\
\cdot \int_{\mathbb{R}^d} (1 + \|\xi\|^{1+\beta}) \exp\left(-\frac{\|x_0 - \xi\|_2^2}{4d^2B^2}\right) \, d\xi \\
\leq c(x_0, f)(c(\gamma))m(\|a\| + \|b'\|^\gamma \|b\|^{1-\gamma})m_m^{-m\gamma/4}
\]

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

\[
|J_0(f)| \leq \int_{\mathbb{R}^d} |f(\xi) \cdot G_0(0, x_0, 1, \xi)| \, d\xi \\
\leq c(f) A_d^{d/2} \int_{\mathbb{R}^d} (1 + \|\xi\|^{1+\beta}) \exp\left(-\frac{\|x_0 - \xi\|_2^2}{2d^2B^2}\right) \, d\xi \\
= c(f, x_0),
\]

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

By (7.8) and (7.9),

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

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


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


