LOWER BOUNDS AND NON-UNIFORM TIME DISCRETIZATION FOR APPROXIMATION OF STOCHASTIC HEAT EQUATIONS

THOMAS MÜLLER-GRONBACH AND KLAUS RITTER

ABSTRACT. We study algorithms for approximation of the mild solution of stochastic heat equations on the spatial domain $]0,1[^d$. The error of an algorithm is defined in L_2-sense. We derive lower bounds for the error of every algorithm that uses a total of N evaluations of one-dimensional components of the driving Wiener process W. For equations with additive noise we derive matching upper bounds and we construct asymptotically optimal algorithms. The error bounds depend on N and d, and on the decay of eigenvalues of the covariance of W in the case of nuclear noise. In the latter case the use of non-uniform time discretizations is crucial.

1. INTRODUCTION

We study stochastic heat equations

$$(1.1) \qquad \begin{aligned} dX(t) &= \Delta X(t)\, dt + B(t, X(t))\, dW(t), \\ X(0) &= \xi \end{aligned}$$

on the Hilbert space $L_2(]0,1[^d)$. Here Δ denotes the Laplace operator with Dirichlet boundary conditions and $W = (W(t))_{t \in [0,1]}$ is a (cylindrical) Brownian motion on $L_2(]0,1[^d)$. Under suitable assumptions on B a mild solution $X = (X(t))_{t \in [0,1]}$ of (1.1) exists and is uniquely determined as a continuous process with values in $L_2(]0,1[^d)$. See, e.g., Da Prato, Zabczyk (1992) and Kallianpur, Xiong (1995).

The construction and analysis of algorithms that approximately solve (1.1) or more general stochastic evolution equations on Hilbert spaces H started with the work by Grecksch, Kloeden (1996) and Gyöngy, Nualart (1997). A partial list of further contributions includes Allen, Novosel, Zhang (1998), Davie, Gaines (2001), Du, Zhang (2002), Gyöngy (1999), Hausenblas (2002, 2003), Kloeden, Shott (2001), Lord, Rougemont (2003), Shardlow (1999), and Yan (2003a, 2003b).

The algorithms that were analyzed so far have the following property in common. For a finite number of elements $g_1, \dots, g_m \in H$ the one-dimensional Brownian motions $\langle W, g_k \rangle$ are evaluated with the same constant step-size $1/\nu$. Thus $N = \nu \cdot m$ is the total number of evaluations of one-dimensional components of W. Based upon this discretization of W an approximation $\widehat{X}_N$ to X is constructed. Here finite difference or Galerkin methods are often used, and the finite-dimensional systems of stochastic differential equations that arise in this way are approximately solved by an implicit or explicit Euler method. Typically, the error of $\widehat{X}_N$ is studied at the discrete time instances ℓ/ν w.r.t. the norm $\| \cdot \|$ in the

Hilbert space H, and error bounds of the form[1]

$$(1.2) \qquad \sup_{\ell=1,\dots,\nu} \left(E\|X(\ell/\nu) - \widehat{X}_N(\ell/\nu)\|^2 \right)^{1/2} \preceq N^{-\alpha}$$

are established. See Gyöngy (1999) and Hausenblas (2002, 2003) for results involving further (semi-)norms in suitable (function) spaces. Gyöngy, Nualart (1997), Gyöngy (1998), Printems (2001), and Yoo (2000) focus on orders of convergence for semi-discretizations.

The discretizations of W that are analyzed in the literature so far may be called uniform, since a fixed step-size is used for all one-dimensional components of W that are evaluated. More generally, one may use a non-uniform discretization: the one-dimensional Brownian motions may be evaluated non-equidistantly, and the number of evaluations may be related to, e.g., the variance function of $\langle W, g_k \rangle / \|g_k\|$. Then, for a fair comparison of algorithms, the error criterion should not depend on the time discretization. In this paper we therefore define the error of an approximation $\widehat{X}_N$ by

$$(1.3) \qquad e(\widehat{X}_N) = \left(E \left(\int_0^1 \|X(t) - \widehat{X}_N(t)\|^2 \, dt \right) \right)^{1/2}.$$

Other (semi-)norms w.r.t the time variable t might be used as well.

It is natural to compare different approximations $\widehat{X}_N$ that use the same number N of evaluations of one-dimensional Brownian motions. The Nth minimal error

$$(1.4) \qquad e(N) = \inf_{\widehat{X}_N} e(\widehat{X}_N)$$

indicates how well the solution X can be approximated by any such algorithm $\widehat{X}_N$, and one is clearly interested in methods $\widehat{X}_N$ with error close to $e(N)$. The analysis of this problem requires an upper bound for the error of a suitable algorithm $\widehat{X}_N$ and a lower bound that holds for the error of every algorithm $\widehat{X}_N$. We mention that minimal errors are a key quantity in information-based complexity, see Traub, Wasilkowski, and Woźniakowski (1988), and, e.g., Novak (1988) and Ritter (2000).

In this paper we have $H = L_2(]0,1[^d)$, and we study equations (1.1) with B given by pointwise multiplication

$$(1.5) \qquad B(t,x)h = G(t,x) \cdot h$$

for $x, h \in H$ and $t \in [0,1]$, where $G : [0,1] \times H \to H$ satisfies mild regularity conditions. Operators B of that type are also considered in, e.g., Grecksch, Kloeden (1996), Gyöngy, Nualart (1997), Gyöngy (1998, 1999), Allen, Novosel, Zhang (1998), Shardlow (1999), Davie, Gaines (2001), Kloeden, Shott (2001), Du, Zhang (2002) and Yan (2003a).

Furthermore, we consider nuclear as well as space-time white noise, i.e., for the covariance $Q : H \to H$ of W we either suppose that Q is a trace class operator or that $Q = \mathrm{id}$. In the sequel these cases are called (TC) and (ID), respectively. For (ID), $d = 1$ is assumed in order to guarantee existence of the mild solution in H. Let

$$(1.6) \qquad h_{\boldsymbol{i}}(u) = 2^{d/2} \cdot \prod_{\ell=1}^{d} \sin(i_\ell \pi u_\ell)$$

---

[1]By definition, $a_N \preceq b_N$ means $\sup_{N \in \mathbb{N}} a_N / b_N < \infty$ for sequences of positive reals a_N and b_N.

with $\boldsymbol{i} = (i_1, \ldots, i_d) \in \mathbb{N}^d$. For (TC) we assume that the normalized eigenfunctions $h_{\boldsymbol{i}}$ of Δ are also eigenfunctions of Q with

$$Qh_{\boldsymbol{i}} = \lambda(|\boldsymbol{i}|_2) \cdot h_{\boldsymbol{i}},$$

where $\lambda : [1, \infty[\to]0, \infty[$ is a non-increasing, regular varying function of index $-\gamma$ with $\gamma \in [d, \infty[\setminus \{2d\}$. Hence the smoothness of the noise and the smoothness of the solution $X(t)$, too, is controlled by γ, with larger values of γ leading to higher smoothness.

A stochastic basis of the (cylindrical) Brownian motion W is given by the independent one-dimensional Brownian motions $\langle W, h_{\boldsymbol{i}} \rangle$ with $\boldsymbol{i} \in \mathbb{N}^d$. In the definition (1.4) of minimal errors $e(N)$ we permit arbitrary methods $\widehat{X}_N$ that use a total of N evaluations of these processes $\langle W, h_{\boldsymbol{i}} \rangle$ at arbitrarily chosen points.

To slightly simplify the presentation we assume that

$$\lambda(|\boldsymbol{i}|_2) = |\boldsymbol{i}|_2^{-\gamma}$$

with

$$\gamma \in]d, \infty[\setminus \{2d\}$$

in the case (TC) throughout the rest of the Introduction. We put $\gamma = 0$ in the case (ID).

In Theorem 1 we obtain a lower bound:

The Nth minimal errors satisfy

$$e(N) \succeq N^{-\alpha^*},$$

where

$$\alpha^* = 1/2 - \frac{d - \gamma/2}{d + 2}$$

for (TC) with $\gamma < 2d$ and (ID) with $d = 1$, and

$$\alpha^* = 1/2$$

for (TC) with $\gamma > 2d$.

Hence there exists a constant $c > 0$ such that $e(\widehat{X}_N) \geq c \cdot N^{-\alpha^*}$ for every algorithm that uses a total of N evaluations of one-dimensional components $\langle W, h_{\boldsymbol{i}} \rangle$ of W. The constant c only depends on the equation, i.e., on d, γ, G, and ξ.

A matching upper bound is provided in Theorem 2 for equations with additive noise:

Suppose that $B(t, x)h = G(t) \cdot h$ with $G : [0, 1] \to H$. Then

$$e(\widehat{X}_N^*) \preceq N^{-\alpha^*}$$

holds for suitable approximations $\widehat{X}_N^*$.

Hence the upper and lower bound are sharp, up to constants, and the approximations $\widehat{X}_N^*$ are asymptotically optimal for these equations, see Corollary 1.

In the case (TC) our method $\widehat{X}_N^*$ uses a non-uniform discretization, while uniform discretizations cannot lead to asymptotically optimal approximations in general, see Remark 6. By our non-uniform discretization, the one-dimensional Brownian motions $\langle W, h_{\boldsymbol{i}} \rangle$ with $\boldsymbol{i}$ in a ball of radius $N^{1/(d+2)}$ are evaluated with step-size proportional to $|\boldsymbol{i}|_2^{\gamma/2}$. These data are used to approximate the random Fourier coefficients $\langle X, h_{\boldsymbol{j}} \rangle$ with $\boldsymbol{j}$ in a suitable hyperbolic cross $\{\boldsymbol{j} \in \mathbb{N}^d : j_1 \cdots j_d \leq J_N\}$.

Lower bounds for stochastic evolution equations were unknown so far, except for the heat equation

$$(1.7) \qquad dX(t) = \frac{\partial^2}{\partial u^2} X(t)\, dt + X(t)\, dW(t)$$

on the spatial domain $]0, 1[$ in the case (ID), see Davie, Gaines (2001). Note that (1.7) corresponds to the particular case $G(t, x) = x$ in (1.5). We get $\alpha^* = 1/6$ from Theorem 1, since $d = 1$ and $\gamma = 0$, in coincidence with the lower bound for equation (1.7) from Davie, Gaines (2001, Thm. 3.1). See Problem 8 for further discussion.

We compare our results with known upper bounds from the literature, where Δ is the Laplace operator on a bounded domain $D \subseteq \mathbb{R}^d$, which sometimes is assumed to have a smooth boundary. In the sequel $\varepsilon > 0$ may be arbitrarily small.

Hausenblas (2003) studies algorithms for stochastic evolution equations with nuclear noise in a very general approach, see also Hausenblas (2002). Her results yield an upper bound (1.2) for stochastic heat equations in the case (TC) with

$$(1.8) \qquad \alpha = \frac{1}{2} \cdot \frac{(\gamma - d + 2) \cdot (\gamma - d)}{4 + (2 + d) \cdot (\gamma - d)} - \varepsilon$$

if $\gamma \leq d + 2$ and

$$(1.9) \qquad \alpha = \frac{1}{2} \cdot \left(1 - \frac{d}{\gamma + 2} \right) - \varepsilon$$

if $\gamma > d + 2$.

Yan (2003b) studies algorithms for stochastic heat equations equations (1.1) in the case (TC), and he obtains an upper bound (1.2) with

$$(1.10) \qquad \alpha = \frac{1}{d + 2} - \varepsilon,$$

see also Yan (2003a).

Now we turn to stochastic heat equations in the case (ID) with $d = 1$. Here an upper bound (1.2) with

$$\alpha = 1/6 - \varepsilon$$

is due to Shardlow (1999) for certain equations with additive noise and due to Yan (2003b) for equation (1.1) in general, see also Yan (2003a). Particular cases are studied by Allen, Novosel, Zhang (1998) and Du, Zhang (2002). Gyöngy (1999) obtains an even stronger result, namely

$$\sup_{t \in [0,1],\ u \in [0,1]} \left(E |X(t, u) - \widehat{X}_N(t, u)|^2 \right)^{1/2} \preceq N^{-\alpha}$$

with

$$\alpha = 1/6.$$

Recall that (1.2) is an error bound that holds at discrete time instances ℓ/ν, while the error $e(\widehat{X}_N)$ according to (1.3) is analyzed in this paper. Still, the algorithms $\widehat{X}_N$ studied by Hausenblas (2003), Yan (2003b), and Shardlow (1999) satisfy $e(\widehat{X}_N) \preceq N^{-\alpha}$, too, with the respective order α, which follows from the mean square-$L_2(D)$-smoothness of the solution X and from the step-sizes $1/\nu$ involved. The analogous statement is obviously true for Gyöngy's estimate, where we have $\alpha = \alpha^*$.

For those results mentioned before that deal with the case (TC) we always have $\alpha < \alpha^*$. In the limit for a low degree of smoothness and with ε tending to zero,

$$\lim_{\gamma \to d+} \alpha = 0$$

for α according to (1.8), while $\alpha^* > 1/(d+2)$ and

$$\lim_{\gamma \to d+} \alpha^* = \frac{1}{d+2}.$$

Conversely, for a high degree of smoothness,

$$\lim_{\gamma \to \infty} \alpha = 1/2$$

for α according to (1.9), while $\alpha^* = 1/2$ already holds if $\gamma > 2d$. Finally, α according to (1.10) does not take into account the smoothness; it coincides with α^* in the limit if γ tends to d and ε tends to zero.

For finite-dimensional systems of stochastic differential equations the minimal errors w.r.t. the L_2-norm are of the order $N^{-1/2}$, see Hofmann, Müller-Gronbach, Ritter (2001) and Müller-Gronbach (2002). Theorem 1 shows that solving a stochastic heat equation is computationally harder than solving a finite-dimensional system if $\gamma < 2d$. In this case the minimal errors even tend to zero arbitrarily slowly in the power scale if the smoothness parameter γ is small compared to the dimension d. On the other hand, if $\gamma > 2d$ then solving a heat equations with additive noise is computationally as hard as solving a scalar stochastic differential equation.

We briefly outline the content of the paper. Assumptions and basic facts about the heat equation (1.1) are stated in Section 2. In Section 3 we formally introduce the computational problem. The lower bound for minimal errors is presented in Section 4. The matching upper bound and an asymptotically optimal algorithm for equations with additive noise are given in Section 5. Some open problems are discussed in Section 6, and proofs are deferred to an Appendix.

2. Assumptions and Basic Facts

Let $H = L_2(D)$ with $D =]0,1[^d$, and let $\|\cdot\|$ denote the norm in H. Furthermore, let $|\boldsymbol{i}|_2$ denote the euclidean norm of $\boldsymbol{i} \in \mathbb{N}^d$.

Concerning the covariance $Q : H \to H$ of the (cylindrical) Brownian motion W we study two cases:

(TC) Q is a trace class operator with eigenfunctions $h_{\boldsymbol{i}}$ given by (1.6) and corresponding eigenvalues

$$\lambda_{\boldsymbol{i}} = \lambda(|\boldsymbol{i}|_2),$$

where

$$\lambda : [1, \infty[\to]0, \infty[$$

is a non-increasing and regularly varying function of index $-\gamma$ for

$$\gamma \in [d, \infty[\setminus \{2d\}.$$

Furthermore,

$$\int_1^\infty \lambda(r) \cdot r^{d-1} \, dr < \infty.$$

We add that the latter property always holds if $\gamma > d$.

(ID) $Q = \text{id}$ and $d = 1$. Here we put

$$\lambda_{\boldsymbol{i}} = 1.$$

We refer to Bingham, Goldie, Teugels (1987) for a comprehensive study of regular variation. The simplest example of a regularly varying function λ of index $-\gamma$ is given by $\lambda(r) = r^{-\gamma}$.

In this paper we consider stochastic heat equations (1.1) with mappings B of the specific form (1.5). We assume that G satisfies the Lipschitz condition

(2.1) $$\|G(t, x) - G(t, y)\| \leq c \cdot \|x - y\|$$

as well as

(2.2) $$\|G(s, x) - G(t, x)\| \leq c \cdot |s - t|^{\eta} \cdot (1 + \|x\|)$$

for all $s, t \in [0, 1]$ and $x \in H$ with constants $c \geq 0$ and

$$\eta = \begin{cases} 1/2 & \text{for (TC)} \\ 1/4 & \text{for (ID)}. \end{cases}$$

Thus G is continuous and satisfies the linear growth condition

(2.3) $$\|G(t, x)\| \leq c \cdot (1 + \|x\|)$$

for some constant $c > 0$. Finally, we suppose that

(2.4) $$\int_0^1 E\|G(t, X(t))\|^2 \, dt > 0,$$

which precisely excludes deterministic heat equations.

We briefly discuss properties of the operator B, which follow from (1.5), (2.1), and (2.2). Let $\langle \cdot, \cdot \rangle$ denote the inner product in H. In the case (TC) we introduce the Hilbert space

$$H_0 = Q^{1/2}(H),$$

equipped with the scalar product

$$\langle Q^{1/2} h_1, Q^{1/2} h_2 \rangle_0 = \langle h_1, h_2 \rangle$$

for $h_1, h_2 \in H$. Moreover, we let

$$\mathcal{L} = \mathcal{L}_2(H_0, H)$$

denote the class of Hilbert-Schmidt operators from H_0 into H, equipped with the Hilbert-Schmidt norm $\|\cdot\|_{\mathcal{L}}$. In the case (ID) we let

$$\mathcal{L} = \mathcal{L}(H, H)$$

denote the class of bounded linear operators from H into H, and we use $\|\cdot\|_{\mathcal{L}}$ to denote the operator norm. Moreover, we put $H_0 = H$.

Remark 1. In the case (TC) we have $H_0 \subseteq L_\infty(D)$, since

(2.5) $$\sup_{\boldsymbol{i} \in \mathbb{N}^d} \|h_{\boldsymbol{i}}\|_\infty < \infty,$$

see Manthey, Zausinger (1999, Lemma 2.2). Hence $B(t, x)h \in H$ for every $h \in H_0$. Moreover,

$$\sum_{\boldsymbol{i} \in \mathbb{N}^d}^{\infty} \lambda_{\boldsymbol{i}} \cdot \|B(t, x)h_{\boldsymbol{i}}\|^2 \leq \sup_{\boldsymbol{i} \in \mathbb{N}^d} \|h_{\boldsymbol{i}}\|_\infty^2 \cdot \sum_{\boldsymbol{i} \in \mathbb{N}^d}^{\infty} \lambda_{\boldsymbol{i}} \cdot \|G(t, x)\|^2,$$

and therefore

$$B : [0,1] \times H \to \mathcal{L}.$$

Note that the eigenvalues $\mu_{\boldsymbol{i}}$ of $-\Delta$ that correspond to the eigenfunctions $h_{\boldsymbol{i}}$ are given by

$$\mu_{\boldsymbol{i}} = \pi^2 \cdot \sum_{\ell=1}^{d} i_\ell^2$$

and the associated semigroup $(S(t))_{t \geq 0}$ in H satisfies

$$S(t)h_{\boldsymbol{i}} = \exp(-\mu_{\boldsymbol{i}} t) \cdot h_{\boldsymbol{i}}.$$

Remark 2. Clearly (1.5) defines a mapping

$$B : [0,1] \times H \to \mathcal{L}(H, L_1(D)).$$

Because of (2.5) we can extend $S(t)$ for $t > 0$ to a mapping $S(t) \in \mathcal{L}(L_1(D), H)$ by

$$S(t)h = \sum_{\boldsymbol{i} \in \mathbb{N}^d} \exp(-\mu_{\boldsymbol{i}} t) \cdot \int_D h_{\boldsymbol{i}}(u) \cdot h(u) \, du \cdot h_{\boldsymbol{i}}.$$

It is straightforward to verify that in the case (ID), since $d = 1$,

$$S(t) \circ B(t,x) \in \mathcal{L}_2(H, H)$$

for $t > 0$ and

$$\|S(t) \circ B(t,x)\|_{\mathcal{L}_2(H,H)} \leq c \cdot t^{-1/4} \cdot \|G(t,x)\|$$

for all $t > 0$ and $x \in H$ with a constant $c > 0$.

We turn to existence and properties of a mild solution of equation (1.1). Take any $\xi \in H$. Then, under the above assumptions, there exists a continuous process $(X(t))_{t \in [0,1]}$ with values in H, which is adapted to the underlying filtration, such that, for every $t \in [0,1]$,

$$X(t) = S(t)\xi + \int_0^t S(t-s)B(s, X(s)) \, dW(s)$$

holds a.s. Moreover, this process is uniquely determined a.s., and it satisfies

(2.6) $$\sup_{t \in [0,1]} E\|X(t)\|^p < \infty$$

for every $p \geq 1$.

Put

(2.7) $$\beta_{\boldsymbol{i}}(t) = \lambda_{\boldsymbol{i}}^{-1/2} \cdot \langle W(t), h_{\boldsymbol{i}} \rangle$$

for $\boldsymbol{i} \in \mathbb{N}^d$ and $t \geq 0$. Then $(\beta_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{N}^d}$ is an independent family of standard one-dimensional Brownian motions. Let

(2.8) $$Y_{\boldsymbol{j}}(t) = \exp(-\mu_{\boldsymbol{j}} t) \cdot \langle \xi, h_{\boldsymbol{j}} \rangle + \sum_{\boldsymbol{i} \in \mathbb{N}^d} \lambda_{\boldsymbol{i}}^{1/2} \cdot Z_{\boldsymbol{i},\boldsymbol{j}}(t),$$

where

(2.9) $$Z_{\boldsymbol{i},\boldsymbol{j}}(t) = \int_0^t \exp(-\mu_{\boldsymbol{j}}(t-s)) \cdot \langle B(s, X(s))h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \, d\beta_{\boldsymbol{i}}(s).$$

Then

$$X(t) = \sum_{\boldsymbol{j} \in \mathbb{N}^d} Y_{\boldsymbol{j}}(t) \cdot h_{\boldsymbol{j}}.$$

holds a.s. and in mean-square sense.

For (TC) we refer to Da Prato, Zabczyk (1992, Sec. 7.1), where a general existence and uniqueness theorem as well as the bound (2.6) for the mild solution are given. These results are applicable to our equations due to Remark 1. In the case (ID), a slight modification of Theorem 3.2.1.(i) from Manthey, Zausinger (1999) together with Remark 2 yields existence and uniqueness of the mild solution as well as (2.6).

3. The Computational Problem

We approximate the mild solution X of the stochastic heat equation (1.1) for every $t \in [0,1]$. We study methods that evaluate a finite number of Brownian motions $\beta_{\boldsymbol{i}}$ at a finite number of points and then produce a curve in H that is close to the corresponding realization of X. The selection and evaluation of the scalar Brownian motions $\beta_{\boldsymbol{i}}$, i.e., the discretization of the (cylindrical) Brownian motion W, is specified by a finite set

$$\mathcal{I} \subseteq \mathbb{N}^d$$

and knots

$$0 < t_{1,\boldsymbol{i}} < \cdots < t_{n_{\boldsymbol{i}},\boldsymbol{i}} \leq 1$$

for $\boldsymbol{i} \in \mathcal{I}$ and $n_{\boldsymbol{i}} \in \mathbb{N}$. Every Brownian motion $\beta_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathcal{I}$ is evaluated at the corresponding knots $t_{\ell,\boldsymbol{i}}$. Letting $n_{\boldsymbol{i}} = 0$ for $\boldsymbol{i} \notin \mathcal{I}$, we have $n \in \mathbb{N}_0^{\mathbb{N}^d}$ and the total number of evaluations is given by

$$|n|_1 = \sum_{\boldsymbol{i} \in \mathbb{N}^d} n_{\boldsymbol{i}}.$$

Formally, an approximation $\widehat{X}$ to X is given by

$$\widehat{X}(t) = \phi\big(t, \beta_{\boldsymbol{i}_1}(t_{1,\boldsymbol{i}_1}), \ldots, \beta_{\boldsymbol{i}_1}(t_{n_{\boldsymbol{i}_1},\boldsymbol{i}_1}), \ldots, \beta_{\boldsymbol{i}_k}(t_{1,\boldsymbol{i}_k}), \ldots, \beta_{\boldsymbol{i}_k}(t_{n_{\boldsymbol{i}_k},\boldsymbol{i}_k})\big),$$

where

$$\phi : [0,1] \times \mathbb{R}^{|n|_1} \to H$$

is any measurable mapping and $\mathcal{I} = \{\boldsymbol{i}_1, \ldots, \boldsymbol{i}_k\}$. The error of $\widehat{X}$ is defined by

$$e(\widehat{X}) = \left(E\left(\int_0^1 \|X(t) - \widehat{X}(t)\|^2 \, dt \right) \right)^{1/2}.$$

We wish to minimize the error among all methods that use a total of at most N evaluations of the scalar Brownian motions. To this end we define

$$e(n) = \inf\{e(\widehat{X}) : \ \widehat{X} \text{ uses } n_{\boldsymbol{i}} \text{ evaluations of } \beta_{\boldsymbol{i}} \text{ for every } \boldsymbol{i} \in \mathbb{N}^d\}$$

for $n \in \mathbb{N}_0^{\mathbb{N}^d}$ with $|n|_1 < \infty$ and the Nth minimal error

$$e(N) = \inf_{|n|_1 \leq N} e(n)$$

for $N \in \mathbb{N}$.

4. The Lower Bound

We establish a lower bound for the Nth minimal error $e(N)$, which holds for every stochastic heat equation (1.1) that satisfies the assumptions from Section 2. Henceforth constants that are hidden in notations like[2] $\asymp$ or $\preceq$ may only depend on d, $(\lambda_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{N}^d}$, G, and ξ.

By assumption (TC),

$$\lambda(r) = r^{-\gamma} \cdot L(r) \tag{4.1}$$

with a slowly varying function $L : [1, \infty[\, \to \,]0, \infty[$. We define

$$e_*(N) = N^{-1/2 + \frac{d - \gamma/2}{d+2}} \cdot \left(L(N^{1/(d+2)}) \right)^{1/2}$$

for (TC) with $\gamma \in [d, 2d[$,

$$e_*(N) = N^{-1/2},$$

for (TC) with $\gamma \in \,]2d, \infty[$, and

$$e_*(N) = N^{-1/6}$$

for (ID). The quantity $e_*(N)$ provides a lower bound for $e(N)$, which depends on N, d, γ, and, for slowly decaying eigenvalues, on L as well. We refer to Appendix A for the proof.

Theorem 1. *The Nth minimal error satisfies*

$$e(N) \succeq e_*(N)$$

for (TC) and any $\xi \in H$ as well as for (ID) and any $\xi \in C^1([0, 1])$.

Remark 3. Regularly varying functions of index $-2d$ are excluded by assumption (TC). Our analysis can be generalized to partially cover this case, too. Suppose, for instance, that $\lambda(r) = r^{-2d}$. Then one can show that

$$e(N) \succeq N^{-1/2} \cdot \ln N.$$

5. The Upper Bound and an Asymptotically Optimal Algorithm for Equations with Additive Noise

In this section we consider stochastic heat equations with additive noise,

$$\begin{aligned} dX(t) &= \Delta X(t)\, dt + B(t)\, dW(t), \\ X(0) &= \xi, \end{aligned} \tag{5.1}$$

so that (1.5) now reads

$$B(t)h = G(t) \cdot h$$

for $h \in H$ and $t \in [0, 1]$, where $G : [0, 1] \to H$. We write $\mathcal{G}(t, u) = G(t)(u)$ for simplicity, where $t \in [0, 1]$ and $u \in \overline{D}$, and we suppose that

$$\mathcal{G} \in C^{(1, 1, \ldots, 1)}([0, 1] \times \overline{D}). \tag{5.2}$$

In the case of additive noise the processes $\beta_{\boldsymbol{i}}$, $Y_{\boldsymbol{j}}$, and $Z_{\boldsymbol{i}, \boldsymbol{j}}$, form a Gaussian system, see (2.7), (2.8), and (2.9). Hence the conditional expectations $\widehat{Y}_{\boldsymbol{j}}(t)$ and $\widehat{Z}_{\boldsymbol{i}, \boldsymbol{j}}(t)$ of $Y_{\boldsymbol{j}}(t)$ and

---

[2]By definition, $a_N \asymp b_N$ means $a_N \preceq b_N$ and $b_N \preceq a_N$ for sequences of positive reals a_N and b_N.

$Z_{\boldsymbol{i},\boldsymbol{j}}(t)$, respectively, given the data $\beta_{\boldsymbol{i}}(t_{\ell,\boldsymbol{i}})$ with $\boldsymbol{i} \in \mathcal{I}$ and $\ell = 1, \dots, n_{\boldsymbol{i}}$, can be computed explicitly. We have

$$\widehat{Y}_{\boldsymbol{j}}(t) = \exp(-\mu_{\boldsymbol{j}} t) \cdot \langle \xi, h_{\boldsymbol{j}} \rangle + \sum_{\boldsymbol{i} \in \mathcal{I}} \lambda_{\boldsymbol{i}}^{1/2} \cdot \widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}(t),$$

and $\widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}$ depends linearly on the data. For approximation of X it suffices to study methods of the form

$$(5.3) \qquad\qquad \widehat{X}(t) = \sum_{\boldsymbol{j} \in \mathcal{J}} \widehat{Y}_{\boldsymbol{j}}(t) \cdot h_{\boldsymbol{j}}$$

with a finite set

$$\mathcal{J} \subseteq \mathbb{N}^d.$$

Let $N \in \mathbb{N}$ be given as an upper bound for the total number of evaluations of the scalar Brownian motions. We determine sets $\mathcal{I} = \mathcal{I}_N$ and $\mathcal{J} = \mathcal{J}_N$, numbers $n_{\boldsymbol{i}} = n_{\boldsymbol{i},N} \in \mathbb{N}$ with $\sum_{\boldsymbol{i} \in \mathcal{I}} n_{\boldsymbol{i}} \preceq N$, and knots $t_{\ell,\boldsymbol{i}} = t_{\ell,\boldsymbol{i},N}$ such that the corresponding method $\widehat{X} = \widehat{X}_N^*$ is asymptotically optimal, i.e., $e(\widehat{X}_N^*) \asymp e(N)$. In fact, we show that the lower bound for the minimal error $e(N)$ from Theorem 1 is an upper bound (up to constants) for the error of our method $\widehat{X}_N^*$, see Theorem 2 below.

We take

$$\mathcal{I}_N = \left\{ \boldsymbol{i} \in \mathbb{N}^d : |\boldsymbol{i}|_2 \leq N^{1/(d+2)} \right\},$$

and we put $n_{\boldsymbol{i},N} = 0$ if $\boldsymbol{i} \notin \mathcal{I}_N$. Otherwise $n_{\boldsymbol{i},N}$ is defined by

$$n_{\boldsymbol{i},N} = \left\lceil \lambda_{\boldsymbol{i}}^{1/2} \cdot N^{1 - \frac{d - \gamma/2}{d+2}} \cdot \left(L(N^{1/(d+2)}) \right)^{-1/2} \right\rceil$$

for (TC) with $\gamma \in [d, 2d[$ and L denoting the slowly varying function in (4.1),

$$n_{\boldsymbol{i},N} = \left\lceil \lambda_{\boldsymbol{i}}^{1/2} \cdot N \right\rceil$$

for (TC) with $\gamma \in]2d, \infty[$, and

$$n_{\boldsymbol{i},N} = \left\lceil N^{2/3} \right\rceil$$

for (ID). We obtain

$$|n_N|_1 = \sum_{\boldsymbol{i} \in \mathcal{I}_N} n_{\boldsymbol{i},N} \preceq N,$$

which obviously holds in the case (ID), and which follows from $\sum_{\boldsymbol{i} \in \mathbb{N}^d} \lambda_{\boldsymbol{i}}^{1/2} < \infty$ in the case (TC) with $\gamma > 2d$. In the case (TC) with $\gamma < 2d$ this bound is a consequence of (A.15) in Appendix A.2. The scalar Brownian motions $\beta_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathcal{I}_N$ are evaluated at the points

$$t_{\ell,\boldsymbol{i},N} = \ell/n_{\boldsymbol{i},N}, \qquad \ell = 1, \dots, n_{\boldsymbol{i},N}.$$

It remains to specify the set $\mathcal{J}_N$. In the case (ID) we take

$$\mathcal{J}_N = \mathcal{I}_N.$$

In the case (TC) we use a so-called hyperbolic cross

$$\mathcal{J}_N = \left\{ \boldsymbol{j} \in \mathbb{N}^d : \prod_{\ell=1}^{d} j_\ell \leq J_N \right\}$$

with

$$J_N = \begin{cases} N^{d/(d+2)} \cdot (\ln N)^{d(d-1)/(\gamma-d+2)} \cdot \left(1 + \left(L(N^{1/(d+2)})\right)^{-d/(\gamma-d+2)}\right) & \text{if } \gamma \in [d, 2d[\\ N^{d/(d+2)} \cdot (\ln N)^{d(d-1)/(d+2)} & \text{if } \gamma \in]2d, \infty[. \end{cases}$$

In this way we have constructed a method $\widehat{X} = \widehat{X}_N^*$, which yields a process with values in span$\{h_{\boldsymbol{j}} : \boldsymbol{j} \in \mathcal{J}_N\}$. See, e.g., Temlyakov (1994) and Ritter (2000) for results and references concerning approximation based on hyperbolic crosses, either for deterministic functions or for random fields.

Recall the definition of $e_*(N)$ from Section 4. We refer to Appendix B for the proof of the following result.

Theorem 2. *The error of the algorithm $\widehat{X}_N$ satisfies*

$$e(\widehat{X}_N^*) \preceq e_*(N)$$

for (TC) and any $\xi \in C^{(1,\dots,1)}(\overline{D})$ as well as for (ID) and any $\xi \in H$.

Remark 4. As in Theorem 1 regular varying functions λ of index $\gamma = -2d$ are not covered by Theorem 2. However, by a slight extension of our methods of proof we are able to treat this case if, e.g., $\lambda(r) = r^{-2d}$.

Use the same set $\mathcal{I}_N$ as above and define

$$n_{\boldsymbol{i},N} = \left\lceil \lambda_{\boldsymbol{i}}^{1/2} \cdot N/\ln N \right\rceil$$

for $\boldsymbol{i} \in \mathcal{I}_N$. Furthermore, take $\mathcal{J}_N$ to be the hyperbolic cross determined by

$$J_N = N^{d/(d+2)} \cdot (\ln N)^{\frac{d(d-3)}{d+2}}.$$

Then $\widehat{X}_N^*$ satisfies

$$e(\widehat{X}_N^*) \preceq N^{-1/2} \cdot \ln N.$$

Remark 5. We discuss the size of the hyperbolic cross $\mathcal{J}_N$, i.e., the number of random Fourier coefficients that are approximated by $\widehat{X}_N^*$. Clearly, in the case (ID)

$$\#\mathcal{J}_N = \#\mathcal{I}_N \asymp N^{1/3}.$$

For the trace class case we use

$$\#\mathcal{J}_N \asymp J_N \cdot (\ln J_N)^{d-1},$$

see Papageorgiou, Wasilkowski (1990, Lemma A.1). Consequently,

$$\#\mathcal{J}_N \asymp N^{d/(d+2)} \cdot (\ln N)^{\frac{2(d^2-1)}{d+2}}$$

for (TC) with $\gamma > 2d$, and

$$\#\mathcal{J}_N \asymp N^{d/(d+2)} \cdot (\ln N)^{\frac{(\gamma+2)(d-1)}{\gamma-d+2}} \cdot \left(1 + \left(L(N^{1/(d+2)})\right)^{-d/(\gamma-d+2)}\right)$$

in the case (TC) with $\gamma < 2d$. Summarizing, for (ID) as well as for (TC) the size of the set $\mathcal{J}_N$ is essentially given by the size of $\mathcal{I}_N$, since

$$\lim_{N \to \infty} N^{-\varepsilon} \cdot \frac{\#\mathcal{J}_N}{N^{d/(d+2)}} = 0$$

for every $\varepsilon > 0$. We add that the latter property together with an upper bound $e_*(N)$ cannot be achieved in general, if a ball is used instead of a hyperbolic cross. The same conclusions hold true in the case $\lambda(r) = r^{-2d}$, see Remark 4, where

$$\#\mathcal{J}_N \asymp N^{d/(d+2)} \cdot (\ln N)^{\frac{2(d^2-d-1)}{d+2}}.$$

The lower and upper bound from Theorems 1 and 2, respectively, coincide, up to constants. Furthermore, the lower bound from Remark 3 matches the upper bound from Remark 4. Therefore the algorithm $\widehat{X}_N^*$ is asymptotically optimal for equations with additive noise under the appropriate assumptions.

Corollary 1. *Consider the stochastic heat equation* (5.1) *with additive noise, and suppose that $\xi \in C^{(1,\dots,1)}(\overline{D})$ and $\mathcal{G} \in C^{(1,1,\dots,1)}([0,1] \times \overline{D})$. Then, for (TC) and (ID),*

$$e(\widehat{X}_N^*) \asymp e(N) \asymp e_*(N).$$

Furthermore, if $\lambda_{\boldsymbol{i}} = |\boldsymbol{i}|_2^{-2d}$ then

$$e(\widehat{X}_N^*) \asymp e(N) \asymp N^{-1/2} \cdot \ln N.$$

Remark 6. In the case (TC) the lower bound from Theorem 1 can be improved for the class of algorithms that use the same number $\nu \in \mathbb{N}$ of evaluations for every Brownian motion $\beta_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathcal{I}$, i.e., $n \in \{0, \nu\}^{\mathbb{N}^d}$. We illustrate this fact for the simple equation

$$dX(t) = \Delta X(t)\,dt + dW(t),$$
$$X(0) = 0,$$

i.e., for $\mathcal{G} = 1$ or, equivalently, for $B(t,x) = \text{id}$. Recall the definition of $e(n)$ from Section 3. The mild solution is given by

$$X(t) = \sum_{\boldsymbol{i} \in \mathbb{N}^d} \lambda_{\boldsymbol{i}}^{1/2} \cdot \int_0^t \exp(-\mu_{\boldsymbol{i}}(t-s))\,d\beta_{\boldsymbol{i}}(s) \cdot h_{\boldsymbol{i}},$$

and Lemma 6 immediately yields

$$e^2(n) \succeq \sum_{\boldsymbol{i} \in \mathbb{N}^d} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}).$$

Assume that $\lambda_{\boldsymbol{i}} = |\boldsymbol{i}|_2^{-\gamma}$ with $\gamma \in {]d, \infty[} \setminus \{2d\}$. Then

$$e(n) \succeq |n|_1^{-1/2 + \frac{d}{2(\gamma+2)}}$$

follows in a straightforward manner.

To achieve an error $e(n) \preceq \varepsilon$ we therefore need at total of

$$N(\varepsilon) \geq \varepsilon^{-\frac{2(\gamma+2)}{\gamma-d+2}}$$

evaluation, if we discretize the (cylindrical) Brownian motion W in this way. On the other hand,

$$N^*(\varepsilon) = \varepsilon^{-\frac{2(d+2)}{\gamma-d+2}}$$

if $\gamma < 2d$ and

$$N^*(\varepsilon) = \varepsilon^{-2}$$

if $\gamma > 2d$ suffices to achieve the accuracy ε with the algorithm $\widehat{X}_N^*$ and $N = N^*(\varepsilon)$. We always have

$$\lim_{\varepsilon \to 0} \frac{N(\varepsilon)}{N^*(\varepsilon)} = \infty.$$

Moreover, $N(\varepsilon)/N^*(\varepsilon)$ increases if γ approaches $2d$, and this ratio tends to one if γ tends to d or ∞. Formally, for $\gamma = 2d$

$$\frac{N(\varepsilon)}{N^*(\varepsilon)} \geq \varepsilon^{-\frac{2d}{d+2}}.$$

For the moderate accuracy $\varepsilon = 10^{-3}$ the right-hand side takes the values 100, 3981, and 10^5 for $d = 1, 3$, and 10, respectively.

Thus one always benefits from adjusting the number of evaluations of the scalar Brownian motions $\beta_{\boldsymbol{i}}$ to the eigenvalues $\lambda_{\boldsymbol{i}}$. Sometimes the benefit is extremely large.

6. Outlook and Discussion

Remark 7. We have presented an asymptotically optimal algorithm only in the case of stochastic heat equations with additive noise. For general equations with multiplication operators, see (1.5), upper bounds are available in the literature, as discussed in Section 1. However, these upper bounds and the lower bound from Theorem 1 do not coincide in the case (TC), and it remains to determine the order of the minimal error and to construct asymptotically optimal algorithms for general equations with multiplication operators.

We conjecture that the lower bound from Theorem 1 is sharp, i.e.,

$$e(N) \asymp e_*(N)$$

also holds in this general case, cf. Corollary 1. Furthermore, we conjecture that an asymptotically optimal algorithm may be based on the discretization of W that is used by the algorithm $\widehat{X}_N^*$ for equations with additive noise. Of course, conditional expectations are infeasible in the general case.

Remark 8. We have analyzed arbitrary approximations $\widehat{X}$ to X of the form

$$\widehat{X}(t) = \phi(t, \chi_1, \ldots, \chi_N),$$

where $\phi : [0,1] \times \mathbb{R}^N \to H$ is any measurable mapping and $\chi_1, \ldots, \chi_N$ with

$$\chi_i \in \mathfrak{W}_0 = \{W(t, h_{\boldsymbol{i}}) : t \in [0,1], \ \boldsymbol{i} \in \mathbb{N}^d\}$$

can be selected in any way. Hence the essential restriction is: only those one-dimensional components of the (cylindrical) Brownian motion W that correspond to an eigenfunction of the Laplacian L may be evaluated. More generally, algorithms $\widehat{X}$ with

$$\chi_i \in \mathfrak{W}_1 = \{W(t, g) : t \in [0,1], \ g \in H\}$$

are studied by, e.g., Allen, Novosel, Zhang (1998), Davie, Gaines (2001), Du, Zhang (2002), Gyöngy (1999), Hausenblas (2002, 2003), and Yan (2003a, 2003b). Here any one-dimensional component of W may be evaluated. For instance, in the case (ID), the random field $(W(t, g_x)_{t, x \in [0,1]})$ with $g_x = 1_{[0,x]}$ defines a Brownian sheet, and evaluation of this random field on a grid is used in several papers. Finally, algorithms $\widehat{X}$ with

$$\chi_i \in \mathfrak{W}_2 = \overline{\operatorname{span}} \, \mathfrak{W}_1$$

are studied by Davie, Gaines (2001). The closure is considered in the space of square-integrable random variables, and $\mathfrak{W}_2$ is called the Hilbert space generated by W.

Associated with every class $\mathfrak{W}_i$ we have the minimal errors $e_i(N)$, and obviously

$$e_2(N) \leq e_1(N) \leq e_0(N) = e(N).$$

We conjecture that these minimal errors coincide, up to constants, i.e.,

$$e(N) \preceq e_2(N),$$

so that $\mathfrak{W}_0$ is essentially as powerful as $\mathfrak{W}_2$.

Due to the upper bound from Gyöngy (1999) and the lower bound from Davie, Gaines (2001) at least $e_1(N) \preceq e_2(N)$ holds for the particular equation

$$dX(t) = \frac{\partial^2}{\partial u^2} X(t)\, dt + X(t)\, dW(t)$$

in the case (ID).

Remark 9. We say that the cylindrical Brownian motion W is uniformly discretized, if $N = \nu \cdot m$ and $\widehat{X}$ is based on

$$\{\chi_1, \ldots, \chi_N\} = \{W(\ell/\nu, g_k) : \ell = 1, \ldots, \nu, \ k = 1, \ldots, m\}$$

for any choice of $g_1, \ldots, g_m \in H$, i.e., a finite number of one-dimensional components of W are evaluated with constant step-size $1/\nu$. It seems that all algorithms that were considered in the literature use a uniform discretization. According to Remark 6 a uniform discretization cannot be asymptotically optimal for that particular equation, as long as $g_1, \ldots, g_m$ are actually eigenfunctions of L. It would be interesting to know, whether this extends to stochastic heat equations in general for the case (TC) and arbitrary $g_1, \ldots, g_m \in H$.

Appendix A. Proof of Theorem 1

Put

$$\Lambda_1(R) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \asymp \int_R^\infty \lambda(r) \cdot r^{d-1}\, dr$$

for $R \geq 1$ in the trace class case.

Proposition 1. *For (TC) there exist constants $c \geq 0$ and $m_0, R_0 \geq 1$ such that*

$$e(N) \succeq \sup_{m \geq m_0} \sup_{R \geq R_0} \left(\inf_{|n|_1 \leq N} \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}, 1/m) - c \cdot \Lambda_1(R)/m^2 \right)^{1/2}.$$

For (ID) with $\xi \in C^1([0,1])$ there exist constants $c \geq 0$ and $m_0, R_0 \geq 1$ such that

$$e(N) \succeq \sup_{m \geq m_0} \left(\inf_{|n|_1 \leq N} \sum_{|\boldsymbol{i}|_2 \geq R_0} \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}, 1/m) - c \cdot 1/m \right)^{1/2}.$$

See Section A.1 for the proof of Proposition 1. A crude application of Proposition 1 immediately yields Theorem 1 in the trace class case with a high degree of smoothness.

Proof of Theorem 1 in the case (TC) with $\gamma > 2d$. Take $R = R_0$, $m = N^{2/3}$ and fix any $\boldsymbol{i}^* \in \mathbb{N}^d$ with $|\boldsymbol{i}^*|_2 \geq R_0$. Then

$$\sum_{|\boldsymbol{i}|_2 \geq R_0} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}, 1/m) \geq \lambda_{\boldsymbol{i}^*} \cdot \min(1/N, 1/\mu_{\boldsymbol{i}^*}, 1/m)$$

if $|n|_1 \leq N$, which yields

$$(A.1) \qquad e(N) \succeq \left(\lambda_{i^*}/N - c \cdot \Lambda_1(R_0)/N^{4/3}\right)^{1/2} \asymp N^{-1/2}.$$

$\square$

The proofs of Theorem 1 for the cases (ID) and (TC) with $\gamma < 2d$ are given in Section A.2. Clearly (A.1) is valid in the latter case, too, but a proper application of Proposition 1 yields a larger lower bound as stated in Theorem 1.

A.1. **Proof of Proposition 1.** First, we provide an upper bound for the mean-square smoothness of the mild solution X, cf. Da Prato, Zabczyk (1992, Thm. 9.1).

Lemma 1. *The mild solution is continuous in mean-square sense. For (TC)*

$$E\|X(s) - X(t)\|^2 \preceq |t - s| \cdot (1 + \psi(\min(s, t))),$$

where $\psi \in L_1([0, 1])$, and for (ID)

$$E\|X(s) - X(t)\|^2 \preceq |s - t|^{1/2}.$$

Proof. Clearly

$$E\|X(s) - X(t)\|^2 = \sum_{j \in \mathbb{N}^d} E(Y_j(s) - Y_j(t))^2.$$

Suppose that $s < t$. Then

$$(A.2) \qquad Y_j(t) - Y_j(s) = (\exp(-\mu_j(t - s)) - 1) \cdot Y_j(s)$$

$$+ \sum_{i \in \mathbb{N}^d} \lambda_i^{1/2} \cdot \int_s^t \exp(-\mu_j(t - r)) \cdot \langle B(r, X(r))h_i, h_j\rangle \, d\beta_i(r),$$

which implies

$$E(Y_j(s) - Y_j(t))^2 = (\exp(-\mu_j(t - s)) - 1)^2 \cdot E(Y_j^2(s))$$

$$+ \int_s^t \exp(-2\mu_j(t - r)) \cdot E\|B^*(r, X(r))h_j\|_{H_0}^2 \, dr.$$

Analogously,

$$E(Y_j(s))^2 = \exp(-2\mu_j s) \cdot \langle \xi, h_j\rangle^2 + \int_0^s \exp(-2\mu_j(s - r)) \cdot E\|B^*(r, X(r))h_j\|_{H_0}^2 \, dr.$$

Put

$$\Gamma_1 = \sum_{j \in \mathbb{N}^d} (\exp(-\mu_j(t - s)) - 1)^2 \cdot E(Y_j^2(s))$$

and

$$\Gamma_2 = \sum_{j \in \mathbb{N}^d} \int_s^t \exp(-2\mu_j(t - r)) \cdot E\|B^*(r, X(r))h_j\|_{H_0}^2 \, dr.$$

We use (2.6) and the linear growth condition (2.3) to obtain

$$\Gamma_2 \leq E\left(\int_s^t \sum_{j \in \mathbb{N}^d} \|B^*(r, X(r))h_j\|_{H_0}^2 \, dr\right) = E\left(\int_s^t \|B^*(r, X(r))\|_{\mathcal{L}_2(H, H_0)}^2 \, dr\right) \preceq t - s$$

for (TC) and

$$\Gamma_2 \leq \sum_{j \in \mathbb{N}^d} \int_s^t \exp(-2\mu_j(t - r)) \cdot E\|B^*(r, X(r))\|_{\mathcal{L}(H, H)}^2 \, dr \preceq \sum_{j \in \mathbb{N}^d} \min(t - s, 1/\mu_j)$$

for (ID). Note that $1 - \exp(-x) \asymp \min(1, x)$ for $x > 0$, and therefore

$$\Gamma_1 \preceq \sum_{\boldsymbol{j} \in \mathbb{N}^d} \min(1, \mu_{\boldsymbol{j}}(t - s)) \cdot E(Y_{\boldsymbol{j}}^2(s))$$

holds in both cases, (TC) and (ID).

Mean-square continuity for (TC) follows from $\sum_{\boldsymbol{j} \in \mathbb{N}^d} E(Y_{\boldsymbol{j}}^2(s)) = E\|X(s)\|^2 < \infty$. Obviously

$$\Gamma_1 \preceq (t - s) \cdot \psi(s)$$

for

$$\psi(s) = \sum_{\boldsymbol{j} \in \mathbb{N}^d} \mu_{\boldsymbol{j}} \cdot E(Y_{\boldsymbol{j}}^2(s)).$$

Since

$$\mu_{\boldsymbol{j}} \cdot \int_0^1 E(Y_{\boldsymbol{j}}^2(s)) \, ds \preceq \langle \xi, h_{\boldsymbol{j}} \rangle^2 + \int_0^1 E\|B^*(r, X(r))h_{\boldsymbol{j}}\|_{H_0}^2 \, dr,$$

we get $\psi \in L_1([0, 1])$. This completes the proof in the case (TC).

For (ID) we use $\xi \in C^1([0, 1])$ and again the linear growth condition and (2.6) to obtain

$$E(Y_{\boldsymbol{j}}^2(s)) \preceq 1/\mu_{\boldsymbol{j}}$$

and therefore

$$\Gamma_1 \preceq \sum_{\boldsymbol{j} \in \mathbb{N}^d} \min(t - s, 1/\mu_{\boldsymbol{j}}).$$

Finally, observe that

(A.3) $$\sum_{\boldsymbol{j} \in \mathbb{N}} \min(t - s, 1/\mu_{\boldsymbol{j}}) \preceq (t - s)^{1/2},$$

since $\mu_{\boldsymbol{j}} \asymp \boldsymbol{j}^2$ for $d = 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now we split the solution X into two parts $X^{(1)}$ and $X^{(2)}$, and we approximate both parts separately. The construction and the estimates are slightly different in the two cases, (TC) and (ID).

In the trace class case we take $R \geq 1$ and $m \in \mathbb{N}$, and we choose points

$$\tau_\ell \in [(\ell - 1)/m, \ell/m]$$

with

$$1/m \cdot \psi(\tau_\ell) \leq \int_{(\ell-1)/m}^{\ell/m} \psi(s) \, ds$$

for $\ell = 1, \ldots, m$. Here ψ denotes the integrable function from Lemma 1. Put $\tau_0 = 0$ and $\tau_{m+1} = 1$. Define processes $Y_{\boldsymbol{j}}^{(1)}$ and $Y_{\boldsymbol{j}}^{(2)}$ by $Y_{\boldsymbol{j}}^{(1)}(0) = 0$ and $Y_{\boldsymbol{j}}^{(2)}(0) = \langle \xi, h_{\boldsymbol{j}} \rangle$ as well as

$$Y_{\boldsymbol{j}}^{(1)}(t) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \int_{\tau_\ell}^t \exp(-\mu_{\boldsymbol{j}}(t - s)) \cdot \langle B(s, X(s))h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \, d\beta_{\boldsymbol{i}}(s)$$

and

$$Y_{\boldsymbol{j}}^{(2)}(t) = \exp(-\mu_{\boldsymbol{j}}(t - \tau_\ell)) \cdot Y_{\boldsymbol{j}}(\tau_\ell) + \sum_{|\boldsymbol{i}|_2 < R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \int_{\tau_\ell}^t \exp(-\mu_{\boldsymbol{j}}(t - s)) \cdot \langle B(s, X(s))h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \, d\beta_{\boldsymbol{i}}(s)$$

for $t \in]\tau_\ell, \tau_{\ell+1}]$. Clearly, by (A.2),

$$X(t) = X^{(1)}(t) + X^{(2)}(t)$$

for
$$X^{(\nu)}(t) = \sum_{\boldsymbol{j} \in \mathbb{N}^d} Y_{\boldsymbol{j}}^{(\nu)}(t) \cdot h_{\boldsymbol{j}}.$$

A simple approximation of $X^{(1)}$ is given by
$$\overline{X}^{(1)}(t) = \sum_{\boldsymbol{j} \in \mathbb{N}^d} \overline{Y}_{\boldsymbol{j}}^{(1)}(t) \cdot h_{\boldsymbol{j}},$$

where $\overline{Y}_{\boldsymbol{j}}^{(1)}(0) = 0$ and
$$\overline{Y}_{\boldsymbol{j}}^{(1)}(t) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \langle B(\tau_\ell, X(\tau_\ell)) h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \cdot \int_{\tau_\ell}^t \exp(-\mu_{\boldsymbol{j}}(t-s)) \, d\beta_{\boldsymbol{i}}(s)$$

for $t \in \,]\tau_\ell, \tau_{\ell+1}]$. Thus, on each subinterval $[\tau_\ell, \tau_{\ell+1}[$, the process $\overline{X}^{(1)}$ follows a simplified evolution equation: the initial value is zero, $B(s, X(s))$ is replaced by $B(\tau_\ell, X(\tau_\ell))$, and only the Brownian motions $\beta_{\boldsymbol{i}}$ with $|\boldsymbol{i}|_2 \geq R$ are relevant.

Lemma 2. *For (TC)*
$$\int_0^1 E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \, dt \preceq \Lambda_1(R)/m^2$$

and
$$\int_{\tau_\ell}^{\tau_{\ell+1}} E\|X^{(1)}(t)\|^2 \, dt \preceq \Lambda_1(R)/m^2.$$

Proof. For $t \in \,]\tau_\ell, \tau_{\ell+1}]$ and $\ell = 0, \dots, m$ we have

(A.4) $\quad E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2$
$$= \sum_{\boldsymbol{j} \in \mathbb{N}^d} \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \int_{\tau_\ell}^t \exp(-2\mu_{\boldsymbol{j}}(t-s)) \cdot E\langle (B(s, X(s)) - B(\tau_\ell, X(\tau_\ell))) h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle^2 \, ds$$

Thus, since $\sup_{\boldsymbol{i} \in \mathbb{N}^d} \|h_{\boldsymbol{i}}\|_\infty < \infty$,
$$E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \preceq \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \int_{\tau_\ell}^t E\|G(s, X(s)) - G(\tau_\ell, X(\tau_\ell))\|^2 \, ds.$$

From (2.1), (2.2) with $\eta = 1/2$, (2.6), and Lemma 1 we get
$$E\|G(s, X(s)) - G(\tau_\ell, X(\tau_\ell))\|^2 \preceq (s - \tau_\ell) \cdot (1 + E\|X(s)\|^2) + E\|X(s) - X(\tau_\ell)\|^2$$
$$\preceq (s - \tau_\ell) \cdot (1 + \psi(\tau_\ell)).$$

Due to the properties of the points τ_ℓ we obtain
$$E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \preceq 1/m^2 \cdot (1 + \psi(\tau_\ell)) \cdot \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}}$$
$$\preceq \Lambda_1(R)/m \cdot \int_{(\ell-1)/m}^{\ell/m} (1 + \psi(s)) \, ds,$$

if $\ell \geq 1$, which implies
$$\int_{\tau_1}^1 E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \, dt \preceq \Lambda_1(R)/m^2.$$

For $\ell = 0$ we use

$$E\|G(s, X(s)) - G(0, \xi)\|^2 \preceq 1,$$

which follows from (2.3) and (2.6), to obtain

$$\int_0^{\tau_1} E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \, dt \preceq \Lambda_1(R)/m^2.$$

The second statement in the Lemma is verified analogously. $\square$

We proceed with an approximation of $X^{(2)}$ in the trace class case. Put

$$m' = \lceil m/\Lambda_1(R) \rceil$$

and

$$\Delta_\ell = (\tau_{\ell+1} - \tau_\ell)/m',$$

so that $\Delta_\ell \leq \Lambda_1(R)/m^2$. Choose points

$$\sigma_{\ell,k} \in [\tau_\ell + (k-1) \cdot \Delta_\ell, \tau_\ell + k \cdot \Delta_\ell]$$

with

$$\Delta_\ell \cdot \psi(\sigma_{\ell,k}) \leq \int_{\tau_\ell + (k-1)\cdot\Delta_\ell}^{\tau_\ell + k\cdot\Delta_\ell} \psi(s) \, ds$$

for $k = 1, \ldots, m'$ and $\ell = 0, \ldots, m$. Furthermore, put $\sigma_{\ell,0} = \tau_\ell$ and $\sigma_{\ell,m'+1} = \tau_{\ell+1}$. Define

$$\overline{X}^{(2)}(t) = \sum_{\boldsymbol{j} \in \mathbb{N}^d} \overline{Y}_{\boldsymbol{j}}^{(2)}(t) \cdot h_{\boldsymbol{j}},$$

where $\overline{Y}_{\boldsymbol{j}}^{(2)}(\tau_\ell) = Y_{\boldsymbol{j}}(\tau_\ell)$ for $\ell = 0, \ldots, m+1$ and

$$\overline{Y}_{\boldsymbol{j}}^{(2)}(t) = \exp(-\mu_{\boldsymbol{j}}(t - \sigma_{\ell,k})) \cdot \overline{Y}_{\boldsymbol{j}}^{(2)}(\sigma_{\ell,k})$$

$$+ \sum_{|\boldsymbol{i}|_2 < R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \langle B(\sigma_{\ell,k}, \overline{X}^{(2)}(\sigma_{\ell,k})) h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \cdot \int_{\sigma_{\ell,k}}^t \exp(-\mu_{\boldsymbol{j}}(t - s)) \cdot d\beta_{\boldsymbol{i}}(s)$$

for $t \in]\sigma_{\ell,k}, \sigma_{\ell,k+1}] \setminus \{\tau_{\ell+1}\}$ with $\ell = 0, \ldots, m$ and $k = 0, \ldots, m'$. Thus, on each subinterval $[\tau_\ell, \tau_{\ell+1}[$, the process $\overline{X}^{(2)}$ only depends on its initial value $X(\tau_\ell)$ and on the Brownian motions $\beta_{\boldsymbol{i}}$ with $|\boldsymbol{i}|_2 < R$.

Lemma 3. *For (TC)*

$$\int_0^1 E\|X^{(2)}(t) - \overline{X}^{(2)}(t)\|^2 \, dt \preceq \Lambda_1(R)/m^2.$$

Proof. We fix ℓ, and we consider $t \in [\tau_\ell, \tau_{\ell+1}[$. Put

$$\zeta(t) = \sup_{\tau_\ell \leq r \leq t} E\|X^{(2)}(r) - \overline{X}^{(2)}(r)\|^2$$

and

$$\overline{G}(t) = \sum_{k=0}^{m'} G(\sigma_{\ell,k}, \overline{X}^{(2)}(\sigma_{\ell,k})) \cdot 1_{]\sigma_{\ell,k}, \sigma_{\ell,k+1}]}(t).$$

Note that

$$\overline{Y}_{\boldsymbol{j}}^{(2)}(t) = \exp(-\mu_{\boldsymbol{j}}(t - \tau_\ell)) \cdot Y_{\boldsymbol{j}}(\tau_\ell) + \sum_{|\boldsymbol{i}|_2 < R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \int_{\tau_\ell}^t \exp(-\mu_{\boldsymbol{j}}(t - r)) \cdot \langle \overline{G}(r) \cdot h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle \, d\beta_{\boldsymbol{i}}(r).$$

Suppose that $s \in \,]\sigma_{\ell,k}, \sigma_{\ell,k+1}]$ in the sequel. As in the proof of Lemma 2 we derive

$$E\|X^{(2)}(t) - \overline{X}^{(2)}(t)\|^2 \preceq \int_{\tau_\ell}^t E\|G(r, X(r)) - \overline{G}(r)\|^2 \, dr,$$

and

$$E\|G(s, X(s)) - \overline{G}(s)\|^2 \preceq E\|X^{(1)}(s)\|^2 + (s - \sigma_{\ell,k}) \cdot (1 + E\|X^{(2)}(s)\|^2)$$
$$+ E\|X^{(2)}(s) - X^{(2)}(\sigma_{\ell,k})\|^2 + E\|X^{(2)}(\sigma_{\ell,k}) - \overline{X}^{(2)}(\sigma_{\ell,k})\|^2.$$

By a similar argument

$$E\|\overline{X}^{(2)}(s)\|^2 \preceq 1 + E\|\overline{X}^{(2)}(\sigma_{\ell,k})\|^2 \preceq m' + E\|X^{(2)}(\tau_\ell)\|^2.$$

Since $E(\langle X^{(1)}(r), X^{(2)}(v)\rangle) = 0$ for $r, v \in [0,1]$, we have

$$E\|X^{(2)}(r)\|^2 \leq E\|X(r)\|^2$$

and

$$E\|X^{(2)}(r) - X^{(2)}(v)\|^2 \leq E\|X(r) - X(v)\|^2.$$

Thus ζ is bounded and

$$E\|G(s, X(s)) - \overline{G}(s)\|^2 \preceq E\|X^{(1)}(s)\|^2 + \Delta_\ell \cdot (1 + \psi(\sigma_{\ell,k})) + \zeta(s)$$

follows from Lemma 1. The properties of the points $\sigma_{\ell,k}$ and Lemma 2 imply

$$\int_{\sigma_{\ell,1}}^t E\|G(r, X(r)) - \overline{G}(r)\|^2 \, dr$$

$$\preceq \int_{\tau_\ell}^{\tau_{\ell+1}} E\|X^{(1)}(r)\|^2 \, dr + \Delta_\ell \cdot (\tau_{\ell+1} - \tau_\ell) + \sum_{k=1}^{m'} \Delta_\ell^2 \cdot \psi(\sigma_{\ell,k}) + \int_{\sigma_{\ell,1}}^t \zeta(r) \, dr$$

$$\preceq \Lambda_1(R)/m^2 + \int_{\tau_\ell}^t \zeta(s) \, ds,$$

if $t \geq \sigma_{\ell,1}$. On the other hand, if $t < \sigma_{\ell,1}$, then

$$\int_{\tau_\ell}^t E\|G(r, X(r)) - \overline{G}(r)\|^2 \, dr \preceq \Delta_\ell \leq \Lambda_1(R)/m^2,$$

since

$$E\|G(s, X(s)) - \overline{G}(s)\|^2 = E\|G(s, X(s)) - G(\tau_\ell, X(\tau_\ell))\|^2 \preceq 1.$$

Consequently,

$$\zeta(t) \preceq \Lambda_1(R)/m^2 + \int_{\tau_\ell}^t \zeta(r) \, dr,$$

and

$$\sup_{\tau_\ell \leq r < \tau_{\ell+1}} E\|X^{(2)}(r) - \overline{X}^{(2)}(r)\|^2 \preceq \Lambda_1(R)/m^2$$

by Gronwall's Lemma. $\qquad\square$

For $Q = \mathrm{id}$ and $d = 1$ we define $X^{(1)}$, $X^{(2)}$, and $\overline{X}^{(1)}$ as previously, with the specific choice $R = 1$ and $\tau_\ell = \ell/m$. Thus, in particular, $Y_j^{(2)}(0) = \langle \xi, h_j \rangle$ and

$$Y_j^{(2)}(t) = \exp(-\mu_j(t - \tau_\ell)) \cdot Y_j(\tau_\ell)$$

for $t \in \,]\tau_\ell, \tau_{\ell+1}]$. Moreover, we take $\overline{X}^{(2)} = X^{(2)}$.

Lemma 4. *For (ID)*

$$\int_0^1 E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \, dt \preceq 1/m.$$

Proof. Note that (A.4) is valid in the case (ID), too. Thus, for $t \in]\tau_\ell, \tau_{\ell+1}]$,

$$E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \leq \sum_{\boldsymbol{j} \in \mathbb{N}^d} \int_{\tau_\ell}^t \exp(-2\mu_{\boldsymbol{j}}(t-s)) \cdot E\|G(s, X(s)) - G(\tau_\ell, X(\tau_\ell))\|^2 \, ds.$$

From (2.1), (2.2) with $\eta = 1/4$, (2.6), and Lemma 1 we get

$$E\|G(s, X(s)) - G(\tau_\ell, X(\tau_\ell))\|^2 \preceq (s - \tau_\ell)^{1/2} \cdot (1 + E\|X(s)\|^2) + E\|X(s) - X(\tau_\ell)\|^2$$
$$\preceq 1/m^{1/2}.$$

Hence

$$E\|X^{(1)}(t) - \overline{X}^{(1)}(t)\|^2 \preceq 1/m^{1/2} \cdot \sum_{\boldsymbol{j} \in \mathbb{N}^d} \min(1/m, 1/\mu_{\boldsymbol{j}}) \preceq 1/m,$$

see (A.3). $\qquad\qquad\square$

The error of approximating X by $\overline{X} = \overline{X}^{(1)} + \overline{X}^{(2)}$ is therefore bounded as follows,

$$(A.5) \qquad \int_0^1 E\|X(t) - \overline{X}(t)\|^2 \, dt \preceq \begin{cases} \Lambda_1(R) \cdot 1/m^2 & \text{for (TC)} \\ 1/m & \text{for (ID).} \end{cases}$$

The second ingredient to the proof of Proposition 1 is a lower bound for the error of approximating $\overline{X}^{(1)}$ by any method $\widehat{X}$. This bound is a consequence of Lemma 6, which deals with approximation of Ornstein-Uhlenbeck processes, and of the fact that the numbers

$$b_{\boldsymbol{i},\boldsymbol{j}}(t) = \left(E\langle B(t, X(t))h_{\boldsymbol{i}}, h_{\boldsymbol{j}}\rangle^2\right)^{1/2}$$

with $t = \tau_\ell$ are not too small along a diagonal in a set $\mathcal{J}_r \times \mathcal{J}_r$, where

$$\mathcal{J}_r = \{r, r+1, \dots\}^d.$$

Put

$$\overline{b}_{\boldsymbol{i},\boldsymbol{j}} = \int_0^1 b_{\boldsymbol{i},\boldsymbol{j}}(t) \, dt.$$

Lemma 5. *For (TC) and (ID)*

$$\left|\sum_{\ell=0}^m (\tau_{\ell+1} - \tau_\ell) \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell) - \overline{b}_{\boldsymbol{i},\boldsymbol{j}}\right| \preceq 1/m^{1/4}.$$

Furthermore, there exist $\boldsymbol{k} \in \mathbb{N}_0^d$ and $r \in \mathbb{N}$ such that

$$\inf_{\boldsymbol{j} \in \mathcal{J}_r} \overline{b}_{\boldsymbol{j}+\boldsymbol{k},\boldsymbol{j}} > 0.$$

Proof. We have

$$\left|\sum_{\ell=0}^m (\tau_{\ell+1} - \tau_\ell) \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell) - \overline{b}_{\boldsymbol{i},\boldsymbol{j}}\right| \leq \sum_{\ell=0}^m \int_{\tau_\ell}^{\tau_{\ell+1}} |b_{\boldsymbol{i},\boldsymbol{j}}(t) - b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell)| \, dt$$

$$\leq \sum_{\ell=0}^m \int_{\tau_\ell}^{\tau_{\ell+1}} \left(E\langle (B(t, X(t)) - B(\tau_\ell, X(\tau_\ell)))h_{\boldsymbol{i}}, h_{\boldsymbol{j}}\rangle^2\right)^{1/2} dt.$$

According to the proofs of Lemma 4 and Lemma 2 we get

$$\left| \sum_{\ell=0}^{m} (\tau_{\ell+1} - \tau_\ell) \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell) - \bar{b}_{\boldsymbol{i},\boldsymbol{j}} \right| \preceq 1/m^{1/4}.$$

for (ID) and

$$\left| \sum_{\ell=0}^{m} (\tau_{\ell+1} - \tau_\ell) \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell) - \bar{b}_{\boldsymbol{i},\boldsymbol{j}} \right| \preceq \tau_1 + \sum_{\ell=1}^{m} \int_{\tau_\ell}^{\tau_{\ell+1}} \left((t - \tau_\ell) \cdot (1 + \psi(\tau_\ell)) \right)^{1/2} dt$$

for (TC). In the latter case it remains to observe that

$$\sum_{\ell=1}^{m} \int_{\tau_\ell}^{\tau_{\ell+1}} \left((t - \tau_\ell) \cdot (1 + \psi(\tau_\ell)) \right)^{1/2} dt \preceq 1/m^{1/2} + 1/m \cdot \sum_{\ell=1}^{m} \left(\int_{(\ell-1)/m}^{\ell/m} \psi(s)\, ds \right)^{1/2} \preceq 1/m^{1/2}.$$

Let $\boldsymbol{m} \in \mathbb{N}_0^d$ and define

$$f_{\boldsymbol{m}}(u) = \prod_{\ell=1}^{d} \cos(m_\ell \pi u_\ell).$$

Moreover, let $|\cdot|_\infty$ denote the sup-norm on $\mathbb{R}^d$. For $\boldsymbol{j} \in \mathcal{J}_r$ and $\boldsymbol{k} \in \mathbb{N}_0^d$

$$\langle B(t,x) h_{\boldsymbol{j}+\boldsymbol{k}}, h_{\boldsymbol{j}} \rangle = \int_D \prod_{\ell=1}^{d} \big(\cos(k_\ell \pi u_\ell) - \cos((2j_\ell + k_\ell) \pi u_\ell) \big) \cdot G(t,x)(u)\, du$$

$$\geq \langle G(t,x), f_{\boldsymbol{k}} \rangle - \sup_{|\boldsymbol{m}|_\infty \geq 2r} \langle G(t,x), f_{\boldsymbol{m}} \rangle.$$

Hence

$$\bar{b}_{\boldsymbol{j}+\boldsymbol{k},\boldsymbol{j}} \geq \int_0^1 \left(E\left(\langle G(t,X(t)), f_{\boldsymbol{k}} \rangle^2 \right) \right)^{1/2} dt - \int_0^1 \left(E\left(\sup_{|\boldsymbol{m}|_\infty \geq 2r} \langle G(t,X(t)), f_{\boldsymbol{m}} \rangle^2 \right) \right)^{1/2} dt,$$

and the linear growth condition together with (2.6) yields

$$\lim_{r \to \infty} \inf_{\boldsymbol{j} \in \mathcal{J}_r} \bar{b}_{\boldsymbol{j}+\boldsymbol{k},\boldsymbol{j}} \geq \int_0^1 \left(E\left(\langle G(t,X(t)), f_{\boldsymbol{k}} \rangle^2 \right) \right)^{1/2} dt.$$

By (2.4), the right-hand side is non-zero for some $\boldsymbol{k} \in \mathbb{N}_0^d$. □

For $t \geq a$, $\mu > 0$, and a standard one-dimensional Brownian motion β we define an Ornstein-Uhlenbeck process by

$$Z(t) = \int_a^t \exp(-\mu(t-s))\, d\beta(s).$$

Lemma 6. *For every $\mu > 0$, $n \in \mathbb{N}$, and all $0 \leq a < s_1 < \cdots < s_n = b < \infty$*

$$\int_a^b E\big(Z(t) - E(Z(t) \mid \beta(s_1), \ldots, \beta(s_n)) \big)^2 dt \succeq (b-a) \cdot \min((b-a)/n, 1/\mu).$$

Proof. Note that

$$Z(t) = \beta(t) - \beta(a) - \mu \cdot \int_a^t \exp(-\mu(t-s)) \cdot (\beta(s) - \beta(a))\, ds.$$

Put $s_0 = a$ and

$$\varepsilon(t) = \left(E\big(Z(t) - E(Z(t) \mid \beta(s_1), \ldots, \beta(s_n)) \big)^2 \right)^{1/2}$$

for $t \in [a, b]$. We turn to the reproducing kernel Hilbert space corresponding to the Brownian motion $(\beta(s) - \beta(a))_{s \in [a,b]}$ and obtain

$$\varepsilon(t) = \sup_{v} \left(v(t) - \mu \cdot \int_a^t \exp(-\mu(t-s)) \cdot v(s) \, ds \right),$$

where

$$v' \in L_2([a, b]), \quad \int_a^b (v'(s))^2 \, ds \leq 1, \quad v(s_0) = \cdots = v(s_n) = 0.$$

See, e.g., Ritter (2000, p. 58). Integration by parts yields

$$\varepsilon(t) = \sup_{\varphi} \int_a^t \exp(-\mu(t-s)) \cdot \varphi(s) \, ds,$$

where

(A.6) $\qquad \varphi \in L_2([a, b]), \quad \int_a^b \varphi^2(s) \, ds \leq 1, \quad \int_{s_{\ell-1}}^{s_\ell} \varphi(s) \, ds = 0$ for $\ell = 1, \ldots, n$.

Assume that $t \in \,]s_{\ell-1}, s_\ell[$. Let $0 < \delta < t - s_{\ell-1}$, and define $\rho > 0$ by

$$\rho^2 = \frac{s_\ell - t}{\delta \cdot (s_\ell - t + \delta)}.$$

Moreover, define φ by

$$\varphi(s) = \begin{cases} \rho & \text{if } t - \delta \leq s \leq t \\ -\rho \cdot \delta/(s_\ell - t) & \text{if } t < s \leq s_\ell \\ 0 & \text{otherwise.} \end{cases}$$

Then (A.6) is satisfied, and therefore

$$\varepsilon(t) \geq \int_a^t \exp(-\mu(t-s)) \cdot \varphi(s) \, ds = \frac{(s_\ell - t)^{1/2}}{\mu} \cdot \frac{1 - \exp(-\mu\delta)}{(\delta \cdot (s_\ell - t + \delta))^{1/2}}.$$

Assume that t is near the center of the ℓ-th subinterval, i.e.,

$$|t - (s_\ell + s_{\ell-1})/2| \leq (s_\ell - s_{\ell-1})/3,$$

and let x^* denote the positive solution of $1 + 2x^* = \exp(x^*)$. Take

$$\delta = \begin{cases} t - s_{\ell-1} & \text{if } (t - s_{\ell-1}) \cdot \mu < x^* \\ x^*/\mu & \text{if } (t - s_{\ell-1}) \cdot \mu \geq x^* \end{cases}$$

to obtain

$$\varepsilon(t) \geq c \cdot (s_\ell - s_{\ell-1})^{1/2}$$

in the first case and

$$\varepsilon(t) \geq c \cdot \mu^{-1/2}$$

in the second case for some constant $c > 0$.

Thus

$$\int_a^b \varepsilon^2(t) \, dt \geq c \cdot 2/3 \cdot \sum_{\ell=1}^n (s_\ell - s_{\ell-1}) \cdot \min((s_\ell - s_{\ell-1}), 1/\mu)$$

$$\geq c/3 \cdot (b - a) \cdot \min\left((b - a)/n, 1/\mu\right).$$

$\hfill \square$

Now we use the assumption that $\lambda_{\boldsymbol{i}} = \lambda(|\boldsymbol{i}|_2)$ with a non-increasing regularly varying function λ of negative index (in the trace class case). Actually, for Proposition 1 to hold only the following consequence of this property is needed.

Lemma 7. *For every $c_1 > 0$ there exist $R_0, c_2 > 0$ such that*

$$\lambda(R) \leq c_2 \cdot \lambda(\widetilde{R})$$

if $R_0 \leq R \leq \widetilde{R} \leq c_1 \cdot R$.

Proof. See Bingham, Goldie, Teugels (1987, Thm. 1.5.6). $\qquad\square$

Proof of Proposition 1. Let $n \in \mathbb{N}_0^{\mathbb{N}^d}$ with $|n|_1 \leq N$, and consider any method $\widehat{X}$ that uses $n_{\boldsymbol{i}}$ evaluations of $\beta_{\boldsymbol{i}}$ for $\boldsymbol{i} \in \mathbb{N}^d$. Let $\mathfrak{T}_{\boldsymbol{i}} \subseteq [0, 1]$ denote the corresponding sets of knots.

Clearly

$$e(\widehat{X}) \geq \left(\int_0^1 E\|\overline{X}(t) - \widehat{X}(t)\|^2 \, dt \right)^{1/2} - \left(\int_0^1 E\|X(t) - \overline{X}(t)\|^2 \, dt \right)^{1/2},$$

so that, because of (A.5), we only have to provide a suitable lower bound for the error of approximating $\overline{X}$ by $\widehat{X}$.

Let $R \geq 1$ and $m \in \mathbb{N}$. We introduce the σ-algebras

$$\mathfrak{A}_\ell = \sigma\big(\mathfrak{F}_{\tau_\ell} \cup \sigma(\{\beta_{\boldsymbol{i}}(s) - \beta_{\boldsymbol{i}}(\tau_\ell) : |\boldsymbol{i}|_2 < R, \ s \in \,]\tau_\ell, 1]\})\big)$$

and

$$\mathfrak{B}_\ell = \sigma\big(\{\beta_{\boldsymbol{i}}(s) - \beta_{\boldsymbol{i}}(\tau_\ell) : |\boldsymbol{i}|_2 \geq R, \ s \in (\mathfrak{T}_{\boldsymbol{i}} \cap \,]\tau_\ell, \tau_{\ell+1}]) \cup [\tau_{\ell+1}, 1]\}\big)$$

for $\ell = 0, \ldots, m$. Let $t \in [\tau_\ell, \tau_{\ell+1}]$ in the sequel, put $\overline{Y}_{\boldsymbol{j}}(t) = \overline{Y}_{\boldsymbol{j}}^{(1)}(t) + \overline{Y}_{\boldsymbol{j}}^{(2)}(t)$, and define

$$U_{\boldsymbol{i},\boldsymbol{j},\ell}(t) = \int_{\tau_\ell}^t \exp(-\mu_{\boldsymbol{j}}(t - s)) \, d\beta_{\boldsymbol{i}}(s)$$

In both cases, (TC) and (ID), we have

    (a) $\overline{X}^{(2)}(t)$ is measurable w.r.t. $\mathfrak{A}_\ell$,
    (b) $\beta_{\boldsymbol{i}}(s)$ is measurable w.r.t. $\sigma(\mathfrak{A}_\ell \cup \mathfrak{B}_\ell)$ for all $\boldsymbol{i} \in \mathbb{N}^d$ and $s \in \mathfrak{T}_{\boldsymbol{i}}$,
    (c) $\mathfrak{A}_\ell$ and $\mathfrak{B}_\ell \cup \sigma(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t))$ are independent if $|\boldsymbol{i}|_2 \geq R$.

By (a) and (b)

$$E\|\overline{X}(t) - \widehat{X}(t)\|^2 = \sum_{\boldsymbol{j} \in \mathbb{N}^d} E\big(\overline{Y}_{\boldsymbol{j}}(t) - \langle \widehat{X}(t), h_{\boldsymbol{j}} \rangle\big)^2 \geq \sum_{\boldsymbol{j} \in \mathbb{N}^d} E\big(\overline{Y}_{\boldsymbol{j}}^{(1)}(t) - E(\overline{Y}_{\boldsymbol{j}}^{(1)} \mid \sigma(\mathfrak{A}_\ell \cup \mathfrak{B}_\ell))\big)^2.$$

Furthermore, by (c),

$$E\big(\overline{Y}_{\boldsymbol{j}}^{(1)}(t) \mid \sigma(\mathfrak{A}_\ell \cup \mathfrak{B}_\ell)\big) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}}^{1/2} \cdot \big\langle B(\tau_\ell, X(\tau_\ell))h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \big\rangle \cdot E(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t) \mid \mathfrak{B}_\ell)$$

and

$$E\big(\overline{Y}_{\boldsymbol{j}}^{(1)}(t) - E(\overline{Y}_{\boldsymbol{j}}^{(1)}(t) \mid \sigma(\mathfrak{A}_\ell \cup \mathfrak{B}_\ell))\big)^2 = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot b_{\boldsymbol{i},\boldsymbol{j}}^2(\tau_\ell) \cdot E\left(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t) - E(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t) \mid \mathfrak{B}_\ell)\right)^2.$$

Let

$$n_{\boldsymbol{i},\ell} = \#\big((\mathfrak{T}_{\boldsymbol{i}} \cap \,]\tau_\ell, \tau_{\ell+1}]) \cup \{\tau_{\ell+1}\}\big)$$

and

$$\delta_\ell = \tau_{\ell+1} - \tau_\ell.$$

By Lemma 6

$$\int_{\tau_\ell}^{\tau_{\ell+1}} E\left(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t) - E(U_{\boldsymbol{i},\boldsymbol{j},\ell}(t)\,|\,\mathfrak{B}_\ell)\right)^2 dt \succeq \delta_\ell \cdot \min(\delta_\ell/n_{\boldsymbol{i},\ell}, 1/\mu_{\boldsymbol{j}}).$$

We therefore conclude that

$$(A.7) \qquad \int_0^1 E\|\overline{X}(t) - \widehat{X}(t)\|^2\,dt \succeq \sum_{|\boldsymbol{i}|_2 \geq R}\sum_{\boldsymbol{j}\in\mathbb{N}^d}\sum_{\ell=0}^{m} \lambda_{\boldsymbol{i}} \cdot b_{\boldsymbol{i},\boldsymbol{j}}^2(\tau_\ell) \cdot \delta_\ell \cdot \min(\delta_\ell/n_{\boldsymbol{i},\ell}, 1/\mu_{\boldsymbol{j}}).$$

Fix $\boldsymbol{i}$ and $\boldsymbol{j}$, put

$$L^1 = \{\ell : 1/\mu_{\boldsymbol{j}} > \delta_\ell/n_{\boldsymbol{i},\ell}\}, \qquad L^2 = \{\ell : 1/\mu_{\boldsymbol{j}} \leq \delta_\ell/n_{\boldsymbol{i},\ell}\},$$

and note that $\sum_{\ell=0}^{m} n_{\boldsymbol{i},\ell} \preceq n_{\boldsymbol{i}} + m$. Moreover,

$$\sum_{\ell\in L^1} b_{\boldsymbol{i},\boldsymbol{j}}^2(\tau_\ell) \cdot \delta_\ell^2/n_{\boldsymbol{i},\ell} \succeq 1/(n_{\boldsymbol{i}} + m) \cdot \left(\sum_{\ell\in L^1} \delta_\ell \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell)\right)^2$$

and

$$\sum_{\ell\in L^2} b_{\boldsymbol{i},\boldsymbol{j}}^2(\tau_\ell) \cdot \delta_\ell M \succeq \left(\sum_{\ell\in L^2} \delta_\ell \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell)\right)^2.$$

Thus

$$(A.8) \quad \sum_{\ell=0}^{m} b_{\boldsymbol{i},\boldsymbol{j}}^2(\tau_\ell) \cdot \delta_\ell \cdot \min(\delta_\ell/n_{\boldsymbol{i},\ell}, 1/\mu_{\boldsymbol{j}}) \succeq \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{j}}, 1/m) \cdot \left(\sum_{\ell=0}^{m} \delta_\ell \cdot b_{\boldsymbol{i},\boldsymbol{j}}(\tau_\ell)\right)^2.$$

Due to Lemma 5 there exist $\boldsymbol{k} \in \mathbb{N}_0^d$ and $m_0, r \in \mathbb{N}$ such that

$$(A.9) \qquad \inf_{\boldsymbol{j}\in\mathscr{J}_r} \inf_{m\geq m_0} \sum_{\ell=0}^{m} \delta_\ell \cdot b_{\boldsymbol{j}+\boldsymbol{k},\boldsymbol{j}}(\tau_\ell) > 0.$$

Put $R^* = R \cdot (r + \max(k_1, \ldots, k_d))$. Combining (A.7), (A.8), and (A.9) we obtain

$$\int_0^1 E\|\overline{X}(t) - \widehat{X}(t)\|^2\,dt \succeq \sum_{\boldsymbol{i}\in\mathscr{J}_{R^*}} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}-k}, 1/m)$$

for $m \geq m_0$.

Let $\boldsymbol{i}(R^*) = (R^*, \ldots, R^*) \in \mathbb{N}^d$ and $\boldsymbol{i} \in \mathbb{N}^d$ with $|\boldsymbol{i}|_2 \geq R$. Clearly $\boldsymbol{i} + \boldsymbol{i}(R^*) \in \mathscr{J}_{R^*}$ and

$$|\boldsymbol{i} + \boldsymbol{i}(R^*) - \boldsymbol{k}|_2^2 \preceq |\boldsymbol{i} + \boldsymbol{i}(R^*)|_2^2 \asymp |\boldsymbol{i}|_2^2.$$

Therefore $\mu_{\boldsymbol{i}+\boldsymbol{i}(R^*)-\boldsymbol{k}} \asymp \mu_{\boldsymbol{i}}$ and $\lambda_{\boldsymbol{i}+\boldsymbol{i}(R^*)} \succeq \lambda_{\boldsymbol{i}}$ if $R \geq R_0$ according to Lemma 7. Hence

$$\int_0^1 E\|\overline{X}(t) - \widehat{X}(t)\|^2\,dt \succeq \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}+\boldsymbol{i}(R^*)}, 1/\mu_{\boldsymbol{i}}, 1/m)$$

for $m \geq m_0$ and $R \geq R_0$. It remains to observe that $|\widetilde{n}|_1 \leq N$ for $\widetilde{n} \in \mathbb{N}_0^{\mathbb{N}^d}$ given by $\widetilde{n}_{\boldsymbol{i}} = n_{\boldsymbol{i}+\boldsymbol{i}(R^*)}$. $\qquad\square$

A.2. **Proof of Theorem 1 in the cases (TC) with $\gamma \in [d, 2d[$ and (ID).** The lower bounds from Proposition 1 involve an optimization problem, and for the analysis of the latter we introduce the nonlinear scalar equation

$$(A.10) \qquad 1/N \cdot \Lambda_2(R_N) = \Lambda_3(R_N)$$

with $N \in \mathbb{N}$, where

$$(A.11) \qquad \Lambda_2(R) = \left(\int_1^R \lambda^{1/2}(r) \cdot r^{d-1} \, dr \right)^2$$

and

$$(A.12) \qquad \Lambda_3(R) = \int_R^\infty \lambda(r) \cdot r^{d-3} \, dr$$

for $R \geq 1$. Clearly $\Lambda_3(R) < \infty$. Hence (A.10) is uniquely solvable for every $N \in \mathbb{N}$. Furthermore, $R_N < R_{N+1}$ and $(R_N)_N$ converges to infinity.

Note that

$$(A.13) \qquad \Lambda_3(R) \asymp R^{d-2} \cdot \lambda(R)$$

for $\gamma \in [d, \infty[$. Moreover, if $\gamma > d$,

$$(A.14) \qquad \Lambda_1(R) \asymp R^d \cdot \lambda(R)$$

and, if $\gamma < 2d$,

$$(A.15) \qquad \Lambda_2(R) \asymp R^{2d} \cdot \lambda(R).$$

See Bingham, Goldie, Teugels (1987, Thm. 1.5.11).

First we study the case (TC) with $\gamma \in [d, 2d[$. We show that

$$(A.16) \qquad e(N) \succeq (\Lambda_3(R_N))^{1/2},$$

and then we determine the asymptotic behaviour of $\Lambda_3(R_N)$.

Lemma 8. *There exists a sequence $(\widetilde{R}_N)_N$ that converges to infinity and moreover satisfies*

$$\lim_{N \to \infty} \frac{\Lambda_2(\widetilde{R}_N)}{\Lambda_2(R_N)} = 0$$

and

$$\lim_{N \to \infty} \frac{\Lambda_1(\widetilde{R}_N)}{\widetilde{R}_N^4 \cdot \Lambda_3(R_N)} = 0.$$

Proof. Let $0 < \varepsilon < 1$ and $0 < \delta < 2d - \gamma$. Then, if R is sufficiently large,

$$\frac{\lambda(R^{1-\varepsilon})}{\lambda(R)} \preceq R^{\varepsilon(\gamma+\delta)}$$

see Bingham, Goldie, Teugels (1987, Thm. 1.5.6). Hence, by (A.15),

$$\frac{\Lambda_2(R^{1-\varepsilon})}{\Lambda_2(R)} \preceq R^{-\varepsilon(2d-\gamma-\delta)}$$

which implies

$$\lim_{R \to \infty} \frac{\Lambda_2(R^{1-\varepsilon})}{\Lambda_2(R)} = 0.$$

Now we specifically take

$$\varepsilon = \frac{1}{\gamma - d + 4},$$

and we claim that the conclusions of the lemma hold for $\widetilde{R}_N = R_N^{1-\varepsilon}$.

We first consider the case $\gamma = d$. Then

$$\lambda(R) \succeq R^{-\gamma-\delta},$$

see Bingham, Goldie, Teugels (1987, Thm. 1.5.6), so that, by (A.13),

$$\frac{\Lambda_1(R^{1-\varepsilon})}{R^{4(1-\varepsilon)} \cdot \Lambda_3(R)} \asymp \frac{\Lambda_1(R^{1-\varepsilon})}{R^{d+1} \cdot \lambda(R)} \preceq R^{\delta-1} \cdot \Lambda_1(R^{1-\varepsilon}).$$

Take $\delta < 1$ to obtain

$$\lim_{R\to\infty} \frac{\Lambda_1(R^{1-\varepsilon})}{R^{4(1-\varepsilon)} \cdot \Lambda_3(R)} = 0.$$

In the case $d < \gamma < 2d$ we assume

$$\delta < \gamma - d + 4,$$

additionally. Then, by (A.14) and (A.13),

$$\frac{\Lambda_1(R^{1-\varepsilon})}{R^{4(1-\varepsilon)} \cdot \Lambda_3(R)} \asymp R^{\varepsilon(4-d)-2} \cdot \frac{\lambda(R^{1-\varepsilon})}{\lambda(R)} \preceq R^{\varepsilon(4-d+\gamma+\delta)-2},$$

which implies

$$\lim_{R\to\infty} \frac{\Lambda_1(R^{1-\varepsilon})}{R^{4(1-\varepsilon)} \cdot \Lambda_3(R)} = 0.$$

$\square$

Proof of Theorem 1 in the case (TC) with $\gamma \in [d, 2d[$. In order to establish the estimate (A.16) we apply Proposition 1 and Lemma 8 with

$$R = \widetilde{R}_N$$

and

$$m = \widetilde{R}_N^2 \geq m_0,$$

which holds if N is sufficiently large. For $n \in \mathbb{N}_0^{\mathbb{N}^d}$ we define

$$f(n) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}, 1/m) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}),$$

and we claim that

(A.17)
$$\inf_{|n|_1 \leq N} f(n) \succeq \inf_{K \geq R} \left(1/N \cdot \left(\Lambda_2^{1/2}(K) - \Lambda_2^{1/2}(R) \right)^2 + \Lambda_3(K) \right).$$

Fix $n \in \mathbb{N}_0^{\mathbb{N}^d}$ with $|n|_1 \leq N$. If $n_{\boldsymbol{i}} \leq \mu_{\boldsymbol{i}}$ for every $\boldsymbol{i} \in \mathbb{N}^d$ with $|\boldsymbol{i}|_2 \geq R$ then

$$f(n) = \sum_{|\boldsymbol{i}|_2 \geq R} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \asymp \Lambda_3(R).$$

Otherwise we may assume $n_{\boldsymbol{i}} > \mu_{\boldsymbol{i}}$ for every $\boldsymbol{i}$ with $|\boldsymbol{i}|_2 \geq R$ and $n_{\boldsymbol{i}} > 0$. Observing the monotonicity properties of $\mu_{\boldsymbol{i}}$ and $\lambda_{\boldsymbol{i}}$ we may further assume that

$$\{\boldsymbol{i} \in \mathbb{N}^d : R \leq |\boldsymbol{i}|_2 < K\} \subseteq \{\boldsymbol{i} \in \mathbb{N}^d : n_{\boldsymbol{i}} > \mu_{\boldsymbol{i}}\} \subseteq \{\boldsymbol{i} \in \mathbb{N}^d : R \leq |\boldsymbol{i}|_2 \leq K + 1\}$$

for some $K \geq R$. In this case

$$f(n) \geq \sum_{R \leq |\boldsymbol{i}|_2 < K} \lambda_{\boldsymbol{i}}/n_{\boldsymbol{i}} + \sum_{|\boldsymbol{i}|_2 > K+1} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \geq 1/N \cdot \left(\sum_{R \leq |\boldsymbol{i}|_2 < K} \lambda_{\boldsymbol{i}}^{1/2} \right)^2 + \sum_{|\boldsymbol{i}|_2 > K+1} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}}.$$

Furthermore,

$$\sum_{|\boldsymbol{i}|_2 > K+1} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \succeq \Lambda_3(K).$$

For $K \leq R + \sqrt{d} + 1$ it suffices to use $f(n) \succeq \Lambda_3(K) \succeq \Lambda_3(R)$. Otherwise we have

$$\sum_{R \leq |\boldsymbol{i}|_2 < K} \lambda_{\boldsymbol{i}}^{1/2} \geq \int_{R+\sqrt{d}}^{K-1} \lambda^{1/2}(r) \cdot r^{d-1} \, dr \asymp \Lambda_2^{1/2}(K-1) - \Lambda_2^{1/2}(R).$$

This completes the proof of (A.17).

We provide a lower bound for the right-hand side in (A.17). For every $K \geq R$

$$1/N \cdot \left(\Lambda_2^{1/2}(K) - \Lambda_2^{1/2}(R) \right)^2 + \Lambda_3(K) \succeq 1/(2N) \cdot \Lambda_2(K) - 1/N \cdot \Lambda_2(R) + \Lambda_3(K)$$
$$\geq 1/2 \cdot (1/N \cdot \Lambda_2(K) + \Lambda_3(K)) - 1/N \cdot \Lambda_2(R).$$

Note that

$$1/N \cdot \Lambda_2(K) + \Lambda_3(K) \geq \Lambda_3(R_N)$$

for every $K \geq 1$. Hence

$$\inf_{K \geq R} \left(1/N \cdot \left(\Lambda_2^{1/2}(K) - \Lambda_2^{1/2}(R) \right)^2 + \Lambda_3(K) \right) \succeq 1/2 \cdot \Lambda_3(R_N) - 1/N \cdot \Lambda_2(\widetilde{R}_N).$$

Now Proposition 1 and (A.17) yield

$$e^2(N) \succeq 1/2 \cdot \Lambda_3(R_N) - 1/N \cdot \Lambda_2(\widetilde{R}_N) - c \cdot \Lambda_1(\widetilde{R}_N)/\widetilde{R}_N^4$$
$$\asymp \Lambda_3(R_N) \cdot \left(1 - \frac{2 \cdot \Lambda_2(\widetilde{R}_N)}{\Lambda_2(R_N)} - \frac{2c \cdot \Lambda_1(\widetilde{R}_N)}{\widetilde{R}_N^4 \cdot \Lambda_3(R_N)} \right).$$

Use Lemma 8 to complete the proof of (A.16).

It remains to determine the asymptotic behaviour of $\Lambda_3(R_N)$. By (A.13) and (A.15)

$$1 = \frac{N \cdot \Lambda_3(R_N)}{\Lambda_2(R_N)} \asymp \frac{N}{R_N^{d+2}}.$$

Thus $R_N \asymp N^{1/(d+2)}$ and

$$\Lambda_3(R_N) \asymp R_N^{d-\gamma-2} \cdot L(R_N) \asymp N^{-1+\frac{2d-\gamma}{d+2}} \cdot L(N^{1/(d+2)}) = e_*^2(N),$$

which finishes the proof. □

Proof of Theorem 1 in the case (ID). We proceed similar to the case (TC) with $\gamma < 2d$. We apply Proposition 1 with $R_0 \in \mathbb{N}$ and

$$m = N^{2/3}.$$

It suffices to show that

(A.18) $$\inf_{|n|_1 \leq N} f(n) \succeq N^{-1/3}$$

for $N \geq \max(m_0^{3/2}, \mu_{R_0})$, where

$$f(n) = \sum_{\boldsymbol{i}=R_0}^{\infty} \cdot \min(1/n_{\boldsymbol{i}}, 1/\mu_{\boldsymbol{i}}, 1/m).$$

Let $n \in \mathbb{N}_0^{\mathbb{N}}$ with $|n|_1 \leq N$. Without loss of generality we may assume

$$\{\boldsymbol{i} \in \mathbb{N} : n_{\boldsymbol{i}} > \max(\mu_{\boldsymbol{i}}, m)\} = \{\boldsymbol{i} \in \mathbb{N} : n_{\boldsymbol{i}} > 0\} = \{R_0, \ldots, K\}$$

for some $K \in \mathbb{N}$ with $K \geq R_0$. In this case

$$f(n) \geq \sum_{\boldsymbol{i}=R_0}^{K} 1/n_{\boldsymbol{i}} + \sum_{\boldsymbol{i}=K+1}^{\infty} 1/\max(\mu_{\boldsymbol{i}}, m) \asymp K^2/N + \sum_{\boldsymbol{i}=K+1}^{\infty} 1/\max(\boldsymbol{i}^2, m).$$

For $(K+1)^2 \geq m$ we get

$$f(n) \succeq K^2/N + 1/K \succeq N^{-1/3}.$$

For $(K+1)^2 < m$ we get

$$f(n) \succeq K^2/N - K/N^{2/3} + N^{-1/3} \geq 3/4 \cdot N^{-1/3}.$$

This completes the proof of (A.18). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

APPENDIX B. PROOF OF THEOREM 2

Throughout this section we consider a stochastic heat equation (5.1) with additive noise that satisfies the smoothness assumption (5.2). Furthermore, $\widehat{X}$ is an algorithm of the form (5.3) and based on the knots

$$t_{\ell,\boldsymbol{i}} = \ell/n_{\boldsymbol{i}}, \qquad \ell = 1, \ldots, n_{\boldsymbol{i}}, \ \boldsymbol{i} \in \mathcal{I}.$$

In the case (ID) we take

$$\mathcal{J} = \mathcal{I}.$$

In the case (TC) we use the hyperbolic cross

$$\mathcal{J} = \left\{\boldsymbol{j} \in \mathbb{N}^d : \prod_{\ell=1}^{d} j_\ell \leq J\right\},$$

where $J \geq 1$ is chosen in such a way that

$$\mathcal{I} \subseteq \mathcal{J}$$

and

$$\text{(B.1)} \qquad (\ln J)^{d-1} \cdot J^{-\frac{\min(\gamma, 2d)-d+2}{d}} \preceq \sum_{\boldsymbol{i} \in \mathcal{I}} \lambda_{\boldsymbol{i}}/n_{\boldsymbol{i}} + \sum_{\boldsymbol{i} \notin \mathcal{I}} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}}.$$

We provide an upper bound for the error of $\widehat{X}$, see Appendix B.1 for the proof. In Appendix B.2 we derive Theorem 2 from this upper bound.

Proposition 2. *For (TC) with $\xi \in C^{(1,\ldots,1)}(\overline{D})$ and for (ID) with $\xi \in H$*

$$e(\widehat{X}) \preceq \left(\sum_{\boldsymbol{i} \in \mathcal{I}} \lambda_{\boldsymbol{i}}/n_{\boldsymbol{i}} + \sum_{\boldsymbol{i} \notin \mathcal{I}} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}}\right)^{1/2}.$$

B.1. **Proof of Proposition 2.** Clearly, the error of $\widehat{X}$ satisfies

(B.2) $$e^2(\widehat{X}) = \sum_{\boldsymbol{j} \in \mathcal{J}} \int_0^1 E\left(Y_{\boldsymbol{j}}(t) - \widehat{Y}_{\boldsymbol{j}}(t)\right)^2 dt + \sum_{\boldsymbol{j} \notin \mathcal{J}} \int_0^1 E\left(Y_{\boldsymbol{j}}(t)\right)^2 dt$$

with

(B.3) $$E\left(Y_{\boldsymbol{j}}(t) - \widehat{Y}_{\boldsymbol{j}}(t)\right)^2 = \sum_{\boldsymbol{i} \in \mathcal{I}} \lambda_{\boldsymbol{i}} \cdot E\left(Z_{\boldsymbol{i},\boldsymbol{j}}(t) - \widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}(t)\right)^2 + \sum_{\boldsymbol{i} \notin \mathcal{I}} \lambda_{\boldsymbol{i}} \cdot E\left(Z_{\boldsymbol{i},\boldsymbol{j}}(t)\right)^2$$

and

(B.4) $$E\left(Y_{\boldsymbol{j}}(t)\right)^2 = \exp(-2\mu_{\boldsymbol{j}}t) \cdot \langle \xi, h_{\boldsymbol{j}} \rangle^2 + \sum_{\boldsymbol{i} \in \mathbb{N}^d} \lambda_{\boldsymbol{i}} \cdot E\left(Z_{\boldsymbol{i},\boldsymbol{j}}(t)\right)^2.$$

Put

$$B_{\boldsymbol{i},\boldsymbol{j}}(t) = \langle B(t)h_{\boldsymbol{i}}, h_{\boldsymbol{j}} \rangle = \int_D h_{\boldsymbol{i}}(u) \cdot h_{\boldsymbol{j}}(u) \cdot \mathcal{G}(t, u) \, du,$$

and note that $B_{\boldsymbol{i},\boldsymbol{j}} \in C^1([0,1])$ due to (5.2). For $\boldsymbol{i} \in \mathcal{I}$ we let $\widehat{\beta}_{\boldsymbol{i}}$ denote the piecewise linear interpolation of $\beta_{\boldsymbol{i}}$ at the knots $t_{\ell,\boldsymbol{i}}$. We have

(B.5) $$\widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}(t) = B_{\boldsymbol{i},\boldsymbol{j}}(t) \cdot \widehat{\beta}_{\boldsymbol{i}}(t) - \int_0^t \frac{\partial}{\partial s}\left(\exp(-\mu_{\boldsymbol{j}}(t-s)) \cdot B_{\boldsymbol{i},\boldsymbol{j}}(s)\right) \cdot \widehat{\beta}_{\boldsymbol{i}}(s) \, ds.$$

Decay properties of $B_{\boldsymbol{i},\boldsymbol{j}}$ and $B'_{\boldsymbol{i},\boldsymbol{j}}$ are crucial in our analysis. Put

(B.6) $$\delta_{\boldsymbol{i},\boldsymbol{j}} = \prod_{\substack{\ell=1 \\ i_\ell \neq j_\ell}}^d 1/|i_\ell - j_\ell|,$$

such that $\delta_{\boldsymbol{i},\boldsymbol{i}} = 1$ in particular.

Lemma 9. *For* $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{N}^d$

$$\sup_{s \in [0,1]} \left(|B_{\boldsymbol{i},\boldsymbol{j}}(s)| + |B'_{\boldsymbol{i},\boldsymbol{j}}(s)|\right) \preceq \delta_{\boldsymbol{i},\boldsymbol{j}}$$

Proof. For simplicity assume that $i_\ell \neq j_\ell$ iff $1 \leq \ell \leq k$, where $0 \leq k \leq d$. By partial integration

$$\int_{[0,1]^k} \prod_{\ell=1}^k \sin(i_\ell \pi u_\ell) \cdot \sin(j_\ell \pi u_\ell) \cdot \mathcal{G}(s, u) \, d(u_1, \ldots, u_k)$$

$$= \int_{[0,1]^k} \prod_{\ell=1}^k \left(\frac{\sin((i_\ell + j_\ell)\pi u_\ell)}{(i_\ell + j_\ell) \cdot \pi} - \frac{\sin((i_\ell - j_\ell)\pi u_\ell)}{(i_\ell - j_\ell) \cdot \pi}\right) \cdot \frac{\partial}{\partial u_1} \cdots \frac{\partial}{\partial u_k} \mathcal{G}(s, u) \, d(u_1, \ldots, u_k).$$

Hence

$$\left|\int_D h_{\boldsymbol{i}}(u) \cdot h_{\boldsymbol{j}}(u) \cdot \mathcal{G}(s, u) \, du\right| \leq \frac{2^d}{\pi^k} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}} \cdot \int_D \left|\frac{\partial}{\partial u_1} \cdots \frac{\partial}{\partial u_k} \mathcal{G}(s, u)\right| du.$$

Clearly, the same estimate holds with $\frac{\partial}{\partial s}\mathcal{G}(s, u)$ in place of $\mathcal{G}(s, u)$. $\qquad\square$

Lemma 10. *For* $\boldsymbol{i} \in \mathbb{N}^d$

$$\sum_{\boldsymbol{j} \in \mathbb{N}^d} \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq 1$$

and

$$\sum_{\boldsymbol{j}\in\mathbb{N}^d} 1/\mu_{\boldsymbol{j}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq 1/\mu_{\boldsymbol{i}}.$$

Proof. The first estimate holds, since

$$\sum_{\boldsymbol{j}\in\mathbb{N}^d} \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \leq \left(2 \cdot \sum_{j=1}^{\infty} 1/j^2 \right)^d.$$

Furthermore, for every $\ell \in \{1,\dots,d\}$,

$$\sum_{\boldsymbol{j}\in\mathbb{N}^d} \frac{i_\ell^2}{\mu_{\boldsymbol{j}}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \sum_{\substack{j=1 \\ j\neq i_\ell}}^{\infty} \frac{i_\ell^2}{j^2} \cdot \frac{1}{(i_\ell - j)^2} \leq \sum_{j=1}^{\lceil i_\ell/2 \rceil} j^{-2} + \sum_{j=\lceil i_\ell/2 \rceil + 1}^{i_\ell - 1} (i_\ell - j)^{-2} + \sum_{j=i_\ell+1}^{\infty} (i_\ell - j)^{-2} \preceq 1,$$

and hereby the second estimate follows. $\square$

Lemma 11. *Let $\boldsymbol{j} \in \mathbb{N}^d$. In the case (TC) with $\gamma \in [d, 2d[$*

$$\sum_{\boldsymbol{i}\in\mathbb{N}^d} \lambda_{\boldsymbol{i}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \lambda_{\boldsymbol{j}} + \prod_{\ell=1}^{d} j_\ell^{-\gamma/d}.$$

In the case (TC) with $\gamma \in]2d, \infty[$

$$\sum_{\boldsymbol{i}\in\mathbb{N}^d} \lambda_{\boldsymbol{i}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \prod_{\ell=1}^{d} j_\ell^{-2}.$$

Proof. Fix $\boldsymbol{j} \in \mathbb{N}^d$. First, we assume that $\gamma < 2d$. Let

$$A_S = \{ \boldsymbol{i} \in \mathbb{N}^d : i_\ell = j_\ell \text{ iff } \ell \notin S \}$$

for $S \subseteq \{1,\dots,d\}$. It suffices to show that

(B.7) $$\sum_{\boldsymbol{i}\in A_S} \lambda_{\boldsymbol{i}} \cdot \prod_{\ell\in S}(i_\ell - j_\ell)^{-2} \preceq \lambda_{\boldsymbol{j}} + \prod_{\ell=1}^{d} j_\ell^{-\gamma/d}$$

for every S. Obviously, (B.7) holds if $S = \emptyset$. Inductively, we proceed as follows. Assume that $\#S = s \geq 1$ and that (B.7) holds for every proper subset of S. Without loss of generality we may assume that $S = \{1,\dots,s\}$. Put

$$a = \left(\sum_{\ell=s+1}^{d} j_\ell^2 \right)^{1/2},$$

and let

$$B = \{ (i_2,\dots,i_s) \in \mathbb{N}^{s-1} : i_\ell \neq j_\ell \text{ for every } \ell \}$$

if $s \geq 2$ and $B = \{0\}$ otherwise. Then

$$\sum_{\boldsymbol{i}\in A_S} \lambda_{\boldsymbol{i}} \cdot \prod_{\ell\in S}(i_\ell - j_\ell)^{-2} = \Sigma_{\leq} + \Sigma_{>},$$

where

$$\Sigma_{\leq} = \sum_{\boldsymbol{i}\in B} \sum_{i_1\leq j_1/2} \lambda\left((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2} \right) \cdot \prod_{\ell=1}^{s}(i_\ell - j_\ell)^{-2}$$

and

$$\Sigma_> = \sum_{\boldsymbol{i} \in B} \sum_{\substack{i_1 > j_1/2 \\ i_1 \neq j_1}} \lambda\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \cdot \prod_{\ell=1}^s (i_\ell - j_\ell)^{-2}.$$

To derive an upper bound for $\Sigma_\leq$ we distinguish the cases

$$|\boldsymbol{i}|_2^2 + a^2 \geq j_1^2$$

and

$$|\boldsymbol{i}|_2^2 + a^2 < j_1^2.$$

In the first case we use Lemma 7 to obtain

$$\lambda\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \preceq \lambda\big((j_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big)$$

for every $i_1 \in \mathbb{N}$. In the second case we take $\varepsilon = 1 - \gamma/(2d)$ and use (4.1) to obtain

$$\lambda\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) = L\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \cdot (i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{-\gamma/2}$$

$$\preceq (i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{\varepsilon/2} \cdot \prod_{\ell=1}^s i_\ell^{-\gamma/d} \cdot \prod_{\ell=s+1}^d j_\ell^{-\gamma/d}$$

$$\preceq j_1^\varepsilon \cdot \prod_{\ell=1}^s i_\ell^{-\gamma/d} \cdot \prod_{\ell=s+1}^d j_\ell^{-\gamma/d}$$

for every $i_1 \leq j_1/2$ since $\varepsilon > 0$, see Bingham, Goldie, Teugels (1987, Thm. 1.5.6). Hence

$$(\text{B.8}) \qquad \lambda\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \preceq \lambda\big((j_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) + j_1^\varepsilon \cdot \prod_{\ell=1}^s i_\ell^{-\gamma/d} \cdot \prod_{\ell=s+1}^d j_\ell^{-\gamma/d}$$

for all $\boldsymbol{i} \in B$ and $i_1 \leq j_1/2$. Furthermore, by hypothesis,

$$(\text{B.9}) \qquad \sum_{\boldsymbol{i} \in B} \sum_{i_1 \leq j_1/2} \lambda\big((j_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \cdot \prod_{\ell=1}^s (i_\ell - j_\ell)^{-2}$$

$$\preceq \sum_{\boldsymbol{i} \in B} \lambda\big((j_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \cdot \prod_{\ell=2}^s (i_\ell - j_\ell)^{-2} \preceq \lambda_{\boldsymbol{j}} + \prod_{\ell=1}^d j_\ell^{-\gamma/d}.$$

On the other hand

$$(\text{B.10}) \qquad \sum_{\boldsymbol{i} \in B} \sum_{i_1 \leq j_1/2} j_1^\varepsilon \cdot \prod_{\ell=1}^s i_\ell^{-\gamma/d} \prod_{\ell=s+1}^d j_\ell^{-\gamma/d} \cdot \prod_{\ell=1}^s (i_\ell - j_\ell)^{-2} \preceq \ln j_1 \cdot j_1^{-2+\varepsilon} \cdot \prod_{\ell=2}^d j_\ell^{-\gamma/d},$$

where the logarithmic term is needed to cover the case $\gamma = d$. Note that $-2 + \varepsilon < -\gamma/d$. Combining (B.8), (B.9), and (B.10) we therefore get

$$\Sigma_\leq \preceq \lambda_{\boldsymbol{j}} + \prod_{\ell=1}^d j_\ell^{-\gamma/d}.$$

The sum $\Sigma_>$ can be bounded as follows. By Lemma 7

$$\lambda\big((i_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big) \preceq \lambda\big((j_1^2 + |\boldsymbol{i}|_2^2 + a^2)^{1/2}\big)$$

for every $i_1 > j_1/2$. Thus, by hypothesis, we get

$$\Sigma_> \preceq \lambda_{\boldsymbol{j}} + \prod_{\ell=1}^d j_\ell^{-\gamma/d}$$

analogously to (B.9), which completes the proof of (B.7).

Now we turn to the case $\gamma > 2d$, and we put $\varepsilon = \gamma/2 - d$. Since $\varepsilon > 0$, the slowly varying function L in (4.1) satisfies

$$L(|\boldsymbol{i}|_2) \preceq |\boldsymbol{i}|_2^\varepsilon$$

for every $\boldsymbol{i} \in \mathbb{N}^d$, see Bingham, Goldie, Teugels (1987, Thm. 1.5.6). Hence

$$\sum_{\boldsymbol{i}\in\mathbb{N}^d} \lambda_{\boldsymbol{i}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \sum_{\boldsymbol{i}\in\mathbb{N}^d} |\boldsymbol{i}|_2^{-(\gamma-\varepsilon)} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2$$

$$\preceq \sum_{\boldsymbol{i}\in\mathbb{N}^d} \prod_{\ell=1}^d i_\ell^{-(\gamma-\varepsilon)/d} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2$$

$$= \prod_{\ell=1}^d \left(\sum_{i_\ell\in\mathbb{N}} i_\ell^{-(\gamma-\varepsilon)/d} \cdot \max(|i_\ell - j_\ell|^{-2}, 1) \right).$$

Observe that $(\gamma - \varepsilon)/d > 2$. Thus

$$\sum_{i_\ell\in\mathbb{N}} i_\ell^{-(\gamma-\varepsilon)/d} \cdot \max(|i_\ell - j_\ell|^{-2}, 1)$$

$$\preceq j_\ell^{-(\gamma-\varepsilon)/d} + \sum_{i_\ell\le j_\ell/2} j_\ell^{-2} \cdot i_\ell^{-(\gamma-\varepsilon)/d} + \sum_{\substack{i_\ell>j_\ell/2 \\ i_\ell\ne j_\ell}} j_\ell^{-(\gamma-\varepsilon)/d} \cdot (i_\ell - j_\ell)^{-2}$$

$$\preceq j_\ell^{-2},$$

which finishes the proof. $\qquad\square$

Now we provide an estimate for approximation of $Z_{\boldsymbol{i},\boldsymbol{j}}(t)$ by its conditional expectation $\widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}(t)$. Put

$$\Delta_{\boldsymbol{i}} = \max_{\ell=1,\dots,n_{\boldsymbol{i}}} (t_{\ell,\boldsymbol{i}} - t_{\ell-1,\boldsymbol{i}})$$

for $\boldsymbol{i} \in \mathcal{I}$, where $t_{0,\boldsymbol{i}} = 0$, and note that $\widehat{Z}_{\boldsymbol{i},\boldsymbol{j}} = 0$ if $\boldsymbol{i} \notin \mathcal{I}$.

Lemma 12. *For $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{N}^d$*

$$\sup_{t\in[0,1]} E\left(Z_{\boldsymbol{i},\boldsymbol{j}}^2(t) \right) \preceq 1/\mu_{\boldsymbol{j}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2.$$

For $\boldsymbol{i} \in \mathcal{I}$ and $\boldsymbol{j} \in \mathbb{N}^d$

$$\sup_{t\in[0,1]} E\left(Z_{\boldsymbol{i},\boldsymbol{j}}(t) - \widehat{Z}_{\boldsymbol{i},\boldsymbol{j}}(t) \right)^2 \preceq \Delta_{\boldsymbol{i}} \cdot \delta_{\boldsymbol{i},\boldsymbol{j}}^2.$$

Proof. Use

$$E\left(Z_{\boldsymbol{i},\boldsymbol{j}}^2(t) \right) = \int_0^t \exp(-2\mu_{\boldsymbol{j}}(t-s)) \cdot B_{\boldsymbol{i},\boldsymbol{j}}^2(s)\,ds \preceq 1/\mu_{\boldsymbol{j}} \cdot \sup_{s\in[0,t]} B_{\boldsymbol{i},\boldsymbol{j}}^2(s)$$

and Lemma 9 to derive the first estimate.

Put

$$f(s) = \exp(-\mu_{\boldsymbol{j}}(t-s)) \cdot B_{\boldsymbol{i},\boldsymbol{j}}(s).$$

Use (B.5) and the analogous formula for $Z_{i,j}$ to obtain

$$E\left(Z_{i,j}(t) - \widehat{Z}_{i,j}(t)\right)^2 \preceq B_{i,j}^2(t) \cdot E\left(\beta_i(t) - \widehat{\beta}_i(t)\right)^2 + E\left(\int_0^t f'(s) \cdot \left(\beta_i(s) - \widehat{\beta}_i(s)\right) ds\right)^2.$$

Recall that β_i and $\widehat{\beta}_i$ are Gaussian. Moreover,

$$E\left(\beta_i(t) - \widehat{\beta}_i(t)\right)^2 \preceq \Delta_i.$$

Therefore

$$E\left(\int_0^t f'(s) \cdot \left(\beta_i(s) - \widehat{\beta}_i(s)\right) ds\right)^2 \preceq \Delta_i \cdot \left(\int_0^t |f'(s)| \, ds\right)^2,$$

and hereby

$$E\left(Z_{i,j}(t) - \widehat{Z}_{i,j}(t)\right)^2 \preceq \Delta_i \cdot \left(B_{i,j}^2(t) + \left(\int_0^t |f'(s)| \, ds\right)^2\right).$$

Clearly,

$$\int_0^t |f'(s)| \, ds \leq \sup_{s \in [0,1]} \left(|B_{i,j}(s)| + |B'_{i,j}(s)|\right).$$

Use Lemma 9 to complete the proof of the second estimate. $\qquad\square$

Proof of Proposition 2. Lemma 12 together with (B.2), (B.3), and (B.4) yields

$$e^2(\widehat{X}) \preceq \Sigma_1 + \Sigma_2,$$

where

$$\Sigma_1 = \sum_{j \in \mathcal{J}} \left(\sum_{i \in \mathcal{I}} \lambda_i/n_i \cdot \delta_{i,j}^2 + \sum_{i \notin \mathcal{I}} \lambda_i/\mu_j \cdot \delta_{i,j}^2\right)$$

and

$$\Sigma_2 = \sum_{j \notin \mathcal{J}} \left(1/\mu_j \cdot \langle \xi, h_j \rangle^2 + \sum_{i \in \mathbb{N}^d} \lambda_i/\mu_j \cdot \delta_{i,j}^2\right).$$

By Lemma 10

$$\Sigma_1 \preceq \sum_{i \in \mathcal{I}} \lambda_i/n_i + \sum_{i \notin \mathcal{I}} \lambda_i/\mu_i,$$

and it remains to establish the analogous estimate for Σ_2.

In the case (ID)

$$\Sigma_2 \preceq \sum_{j \notin \mathcal{I}} 1/\mu_j \leq \sum_{i \in \mathcal{I}} \lambda_i/n_i + \sum_{i \notin \mathcal{I}} \lambda_i/\mu_i$$

by means of the first estimate from Lemma 10 and $\mathcal{I} = \mathcal{J}$.

In the case (TC), by assumption on ξ,

$$\langle \xi, h_j \rangle^2 \preceq \prod_{\ell=1}^d j_\ell^{-2}.$$

Put $\theta = (\min(\gamma, 2d) - d + 2)/d$. Then

$$\sum_{j \notin \mathcal{J}} 1/\mu_j \cdot \langle \xi, h_j \rangle^2 \preceq \sum_{j \notin \mathcal{J}} |j|_2^2 \cdot \prod_{\ell=1}^d j_\ell^{-2} \preceq \sum_{j \notin \mathcal{J}} \prod_{\ell=1}^d j_\ell^{-2-2/d} \preceq \sum_{j \notin \mathcal{J}} \prod_{\ell=1}^d j_\ell^{-(\theta+1)}.$$

Lemma 11 yields

$$\sum_{\boldsymbol{j}\notin\mathcal{J}}\sum_{\boldsymbol{i}\in\mathbb{N}^d}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{j}}\cdot\delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \sum_{\boldsymbol{i}\notin\mathcal{J}}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} + \sum_{\boldsymbol{j}\notin\mathcal{J}}|\boldsymbol{j}|_2^{-2}\cdot\prod_{\ell=1}^d j_\ell^{-(\theta-2/d+1)}$$

$$\preceq \sum_{\boldsymbol{i}\notin\mathcal{J}}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} + \sum_{\boldsymbol{j}\notin\mathcal{J}}\prod_{\ell=1}^d j_\ell^{-(\theta+1)}$$

if $\gamma < 2d$ and

$$\sum_{\boldsymbol{j}\notin\mathcal{J}}\sum_{\boldsymbol{i}\in\mathbb{N}^d}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{j}}\cdot\delta_{\boldsymbol{i},\boldsymbol{j}}^2 \preceq \sum_{\boldsymbol{j}\notin\mathcal{J}}\prod_{\ell=1}^d j_\ell^{-(\theta+1)}$$

if $\gamma > 2d$. Since $\mathcal{I}\subseteq\mathcal{J}$, we conclude that

$$\Sigma_2 \preceq \sum_{\boldsymbol{i}\notin\mathcal{I}}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} + \sum_{\boldsymbol{j}\notin\mathcal{J}}\prod_{\ell=1}^d j_\ell^{-(\theta+1)}$$

in both cases. Furthermore,

$$\sum_{\boldsymbol{j}\notin\mathcal{J}}\prod_{\ell=1}^d j_\ell^{-(\theta+1)} \preceq \frac{(\ln J)^{d-1}}{J^\theta},$$

see Papageorgiou, Wasilkowski (1990, Section 2.2). Thus by property (B.1) of J

$$\Sigma_2 \preceq \sum_{\boldsymbol{i}\in\mathcal{I}}\lambda_{\boldsymbol{i}}/n_{\boldsymbol{i}} + \sum_{\boldsymbol{i}\notin\mathcal{I}}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}},$$

which completes the proof. $\square$

B.2. **Proof of Theorem 2.** Now we suppose that $\mathcal{I} = \mathcal{I}_N$ and $n = n_N$ as defined in Section 5. Recall the definition of $e_*(N)$ from Section 4.

Lemma 13. *For (TC) and (ID)*

$$\sum_{\boldsymbol{i}\in\mathcal{I}_N}\lambda_{\boldsymbol{i}}/n_{\boldsymbol{i},N} + \sum_{\boldsymbol{i}\notin\mathcal{I}_N}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \asymp e_*^2(N).$$

Proof. For (ID) the lemma trivially holds. In the case (TC) with $\gamma < 2d$ we have

$$\sum_{\boldsymbol{i}\in\mathcal{I}_N}\lambda_{\boldsymbol{i}}/n_{\boldsymbol{i},N} \preceq N^{-1+\frac{d-\gamma/2}{d+2}}\cdot\left(L(N^{1/(d+2)})\right)^{1/2}\cdot\sum_{|\boldsymbol{i}|_2\leq N^{1/(d+2)}}\lambda_{\boldsymbol{i}}^{1/2}$$

$$\asymp N^{-1+\frac{d-\gamma/2}{d+2}}\cdot\left(L(N^{1/(d+2)})\right)^{1/2}\cdot\left(\Lambda_2(N^{1/(d+2)})\right)^{1/2}$$

$$= e_*^2(N)\cdot N^{\frac{\gamma/2-d}{d+2}}\cdot\left(L(N^{1/(d+2)})\right)^{-1/2}\cdot\left(\Lambda_2(N^{1/(d+2)})\right)^{1/2}$$

and

$$\sum_{\boldsymbol{i}\notin\mathcal{I}_N}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} = \sum_{|\boldsymbol{i}|_2 > N^{1/(d+2)}}\lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \asymp \Lambda_3(N^{1/(d+2)})$$

with Λ_2 and Λ_3 according to (A.11) and (A.12). Moreover,

$$\left(\Lambda_2(N^{1/(d+2)})\right)^{1/2} \asymp N^{\frac{d-\gamma/2}{d+2}}\cdot\left(L(N^{1/(d+2)})\right)^{1/2}$$

by (A.15) and

$$\Lambda_3(N^{1/(d+2)}) \asymp N^{-1+\frac{2d-\gamma}{d+2}} \cdot L(N^{1/(d+2)}) = e_*^2(N)$$

by (A.13). This completes the proof in the case $\gamma < 2d$.

In the case (TC) with $\gamma > 2d$

$$\sum_{\boldsymbol{i} \in \mathcal{I}_N} \lambda_{\boldsymbol{i}}/n_{\boldsymbol{i},N} + \sum_{\boldsymbol{i} \notin \mathcal{I}_N} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}} \preceq 1/N \cdot \sum_{|\boldsymbol{i}|_2 \leq N^{1/(d+2)}} \lambda_{\boldsymbol{i}}^{1/2} + \sum_{|\boldsymbol{i}|_2 > N^{1/(d+2)}} \lambda_{\boldsymbol{i}}/\mu_{\boldsymbol{i}}$$

$$\asymp e_*^2(N) \cdot \left(\Lambda_2(N^{1/(d+2)})\right)^{1/2} + \Lambda_3(N^{1/(d+2)})$$

$$\asymp e_*^2(N) + \Lambda_3(N^{1/(d+2)}),$$

since $\sup_{R \geq 1} \Lambda_2(R) < \infty$. Furthermore, by (A.13) and since L is slowly varying,

$$\Lambda_3(N^{1/(d+2)}) \asymp N^{\frac{d-2-\gamma}{d+2}} \cdot L(N^{1/(d+2)}) = e_*^2(N) \cdot N^{\frac{-(\gamma-2d)}{d+2}} \cdot L(N^{1/(d+2)}) \preceq e_*^2(N).$$

On the other hand

$$\sum_{\boldsymbol{i} \in \mathcal{I}_N} \lambda_{\boldsymbol{i}}/n_{\boldsymbol{i},N} \geq \lambda_{\boldsymbol{i}^*}/n_{\boldsymbol{i}^*,N} \asymp 1/N \asymp e_*^2(N)$$

for $\boldsymbol{i}^* = (1, \ldots, 1)$. $\qquad \square$

Finally, we suppose that $J = J_N$ and $\mathcal{J} = \mathcal{J}_n$ as defined in Section 5. Then $\mathcal{I}_N \subseteq \mathcal{J}_N$. Moreover, $\ln J_N \asymp \ln N$, and it is straightforward to verify that

$$J_N^{-\frac{\min(\gamma, 2d)-d+2}{d}} \leq e_*^2(N) \cdot (\ln N)^{-(d-1)}.$$

Hence Lemma 13 yields (B.1).

We thus obtain Theorem 2 from Proposition 2 and Lemma 13.

References

Allen, E. J., Novosel, S. J., Zhang, Z. (1998), Finite element and difference approximation of some linear stochastic partial differential equations, Stochastics Stochastics Rep. **64**, 117–142.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1987), Regular Variation, Cambridge Univ. Press, Cambridge.

Da Prato, G., and Zabczyk, J. (1992), Stochastic Equations in Infinite Dimensions, Cambridge Univ. Press, Cambridge.

Davie, A. M., Gaines, J. (2001), Convergence of numerical schemes for the solution of parabolic partial differential equations, Math. Comp. **70**, 121–134.

Du, Q., Zhang T. (2002), Numerical approximation of some linear stochastic partial differential equations driven by special additive noises, SIAM J. Numer. Anal. **40**, 1421–1445.

Grecksch, W., Kloeden, P. E. (1996), Time-discretised Galerkin approximations of parabolic stochastic PDEs, Bull. Austr. Math. Soc. **54**, 79–85.

Gyöngy, I. (1998), Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise I, Potential Analysis **9**, 1–25.

Gyöngy, I. (1999), Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise II, Potential Analysis **11**, 1–37.

Gyöngy, I., Nualart, D. (1997), Implicit scheme for stochastic parabolic partial differential equations driven by space-time white noise, Potential Analysis **7**, 725–757.

Hausenblas, E. (2002), Numerical analysis of semilinear stochastic evolution equations in Banach spaces, J. Comput. Appl. Math. **147**, 485–516.

Hausenblas, E. (2003), Approximation for semilinear stochastic evolution equations, Potential Analysis **18**, 141–186.

Hofmann, N., Müller-Gronbach, T., Ritter, K. (2001), The optimal discretization of stochastic differential equations, J. Complexity **17**, 117–153.

Kallianpur, G., and Xiong, J. (1995), Stochastic Differential Equations in Infinite Dimensional Spaces, Lect. Notes–Monogr. Ser. Vol. **26**, Inst. Math. Statist., Hayward.

Kloeden, P. E., Shott, S. (2001), Linear-implicit strong schemes for Ito-Galerkin approximations of stochastic PDEs, J. Appl. Math. Stochastic Anal. **14**, 47–53.

Lord, G. J., Rougemont, J. (2003), A numerical scheme for stochastic PDEs with Gevrey regularity, Technical Report, Department of Mathematics, Heriot-Watt University, Edinburgh.

Manthey, R., and Zausinger, T. (1999), Stochastic evolution equations in $L_\rho^{2\nu}$, Stochastics Stochastics Rep. **66**, 37–85.

Müller-Gronbach, T. (2002), Strong Approximation of Systems of Stochastic Differential Equations. Habilitationsschrift, TU Darmstadt.

Novak, E. (1988), Deterministic and Stochastic Error Bounds in Numerical Analysis, Lect. Notes in Math. **1349**, Springer-Verlag, Berlin.

Papageorgiou, A., Wasilkowski, G. W. (1900), On the average complexity of multivariate problems, J. Complexity **6**, 1–23.

Printems, J. (2001), On the discretization in time of parabolic stochastic partial differential equations, Monte Carlo Meth. Appl. **7**, 359–368.

Ritter, K. (2000), Average-Case Analysis of Numerical Problems, Lect. Notes in Math. **1733**, Springer-Verlag, Berlin.

Shardlow, T. (1999), Numerical methods for stochastic parabolic PDEs, Numer. Funct. Anal. Optim. **20**, 121–145.

Temlyakov, V. N. (1994), Approximation of Periodic Functions, Nova Science, New York.

Traub, J. F., Wasilkowski, G. W., Woźniakowski, H. (1988), Information-Based Complexity, Academic Press, New York.

Yan, Y. (2003a), The finite element method for a linear stochastic parabolic partial differential equation driven by additive noise, Preprint 2003-07, Chalmers Finite Element Center, Chalmers University, Göteborg.

Yan, Y. (2003b), A finite element method for a nonlinear stochastic parabolic equation, Preprint 2003-08, Chalmers Finite Element Center, Chalmers University, Göteborg.

Yoo, H. (2000), Semi-discretization of stochastic partial differential equations on $\mathbb{R}^1$ by a finite-difference method, Math. Comp. **69**, 653–666.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK, FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT MAGDEBURG, POSTFACH 4120, 39016 MAGDEBURG, GERMANY
  E-mail address: gronbach@mail.math.uni-magdeburg.de

FACHBEREICH MATHEMATIK, TECHNISCHE UNIVERSITÄT DARMSTADT, SCHLOSSGARTENSTRASSE 7, 64289 DARMSTADT, GERMANY
  E-mail address: ritter@mathematik.tu-darmstadt.de