ADAPTIVE GALERKIN APPROXIMATION ALGORITHMS FOR PARTIAL DIFFERENTIAL EQUATIONS IN INFINITE DIMENSIONS

CHRISTOPH SCHWAB AND ENDRE SÜLI

Abstract. Space-time variational formulations of infinite-dimensional Fokker–Planck (FP) and Ornstein–Uhlenbeck (OU) equations for functions on a separable Hilbert space H are developed. The well-posedness of these equations in the Hilbert space $\mathrm{L}^2(H,\mu)$ of functions on H, which are square-integrable with respect to a Gaussian measure μ on H, is proved. Specifically, for the infinite-dimensional FP equation, adaptive space-time Galerkin discretizations, based on a tensorized Riesz basis, built from biorthogonal piecewise polynomial wavelet bases in time and the Hermite polynomial chaos in the Wiener–Itô decomposition of $\mathrm{L}^2(H,\mu)$, are introduced and are shown to converge quasioptimally with respect to the nonlinear, best N-term approximation benchmark. As a consequence, the proposed adaptive Galerkin solution algorithms perform quasioptimally with respect to the best N-term approximation in the finite-dimensional case, in particular. All constants in our error and complexity bounds are shown to be independent of the number of "active" coordinates identified by the proposed adaptive Galerkin approximation algorithms.

1. Introduction

Partial differential equations in infinite dimensions arise in a number of relevant applications; most notably as forward and backward Kolmogorov equations for stochastic partial differential equations (see, e.g., [7, 5] and the references therein). Their numerical solution has, to date, received scant attention, however. Approximations to such equations are typically attempted by path simulations in the corresponding stochastic partial differential equation, whose path-wise solutions belong to function spaces over finite-dimensional domains and can be, therefore, approximated by standard discretization techniques, combined with Monte Carlo path sampling. In the present paper, we propose a novel, adaptive approach to the construction of finite-dimensional numerical approximations to the deterministic forward equation in infinite-dimensional spaces, which exhibit certain optimality properties. The proposed approach is based on space-time variational formulations of these equations, which are posed in Gel'fand-triples of Sobolev spaces over a separable Hilbert space H with respect to a Gaussian measure, and on Riesz bases of these spaces, which have been developed for linear and nonlinear parabolic PDEs in finite dimensions in [24, 8, 17].

The structure of the paper is as follows: in the next section we present two space-time variational formulations of linear parabolic equations set in a domain of finite dimension, which we then apply in Section 3 to Fokker–Planck equations that arise in bead-spring chain models for d-dimensional polymeric flow, $d \in \{2, 3\}$, with chains consisting of $K+1$ beads whose kinematics are statistically described by a configuration vector $q \in \mathbb{R}^{Kd}$, $K \gg 1$. The probability density function $\psi = \psi(q,t)$ that is sought as the solution of the associated Fokker–Planck equation is therefore a function of Kd spatial variables with $K \gg 1$ and the time variable t. Our aim is to embed this finite-dimensional problem of potentially very high dimension into an infinite-dimensional problem. We therefore turn to Fokker–Planck equations in infinite-dimensional domains by recapitulating

The research reported in this paper was carried out at the Hausdorff Institute for Mathematics (HIM), Bonn, during the HIM-Trimester on "High Dimensional Approximation", May–August 2011.

CS was partially supported by the European Research Council (ERC) under the FP7 programme ERC AdG 247277. ES was partially supported by the EPSRC Science and Innovation award to the Oxford Centre for Nonlinear PDE (EP/E035027/1).

basic facts about Gaussian measures, before proving in Section 5 well-posedness of the second of our two space-time variational formulations in the infinite-dimensional space $\mathrm{L}^2(H,\mu)$, where H denotes a separable Hilbert space and μ is a Gaussian measure on H. We show in particular that the solution operator of the variational formulation is an isomorphism between suitable solution and data spaces. We then turn our attention to adaptive Galerkin approximations. We outline the general principle in Section 6, where we introduce the idea of conversion of abstract, well-posed operator equations on separable Hilbert spaces to equivalent, infinite matrix-vector problems in the sequence space $\ell^2(\mathbb{N})$. After reviewing some relevant facts concerning N-term approximations in $\ell^2(\mathbb{N})$, we introduce abstract adaptive Galerkin approximation algorithms that *construct* sequences of N-term approximations, which, while not being best N-term approximations, are optimal in the sense that they converge asymptotically at the rate afforded by the best N-term approximation, provided that certain conditions are met by the operators and the Riesz bases used to discretize them. In Section 9 the abstract concepts are applied to infinite-dimensional Fokker–Planck equations. Notably, a Wiener polynomial chaos type Riesz basis in $\mathrm{L}^2(H,\mu)$ and a wavelet basis in time are used for the Galerkin discretization. In Section 10 we verify the abstract assumptions for the specific equations of interest, and in Section 11 we consider the more general setting of nonsymmetric equations with drift, leading to our main result concerning the optimality of adaptive Galerkin discretizations with dimension-independent bounds, in Section 12.

2. Space-time variational formulation of abstract parabolic problems

2.1. **Abstract elliptic equations.** Suppose that $D \subseteq \mathbb{R}^d$ is either a bounded open domain with Lipschitz continuous boundary ∂D or $D = \mathbb{R}^d$. In D, we consider the elliptic partial differential equation

$$Au = f \text{ in } D \tag{2.1}$$

with suitable homogeneous boundary conditions on ∂D when D is bounded, or subject to a decay condition at infinity when $D = \mathbb{R}^d$. In typical weak formulations of elliptic boundary-value problems the solution u is sought in a certain separable Hilbert space $\mathcal{V}$ of functions defined on D (usually a hilbertian Sobolev space), so that $A \in \mathcal{L}(\mathcal{V},\mathcal{V}^*)$ and $f \in \mathcal{V}^*$, where $\mathcal{V}^*$ is the dual space of $\mathcal{V}$, resulting in the following weak formulation of problem (2.1):

$$\text{Given } f \in \mathcal{V}^*, \text{ find } u \in \mathcal{V} \text{ such that: } \mathfrak{a}(u,v) = {}_{\mathcal{V}^*}\langle f, v\rangle_{\mathcal{V}} \quad \forall v \in \mathcal{V}.$$

Here, $\mathfrak{a}(u,v) = {}_{\mathcal{V}^*}\langle Au, v\rangle_{\mathcal{V}}$ and ${}_{\mathcal{V}^*}\langle \cdot,\cdot\rangle_{\mathcal{V}}$ denotes the $\mathcal{V}^* \times \mathcal{V}$ duality pairing. It is also assumed that there exists a separable Hilbert space $\mathcal{H}$, identified with its dual $\mathcal{H}^*$, such that

$$\mathcal{V} \hookrightarrow \mathcal{H} \cong \mathcal{H}^* \hookrightarrow \mathcal{V}^*, \tag{2.2}$$

where the continuous embeddings signified by the symbol $\hookrightarrow$ are dense and compact. All Hilbert spaces considered here are, for the sake of simplicity, assumed to be over the field $\mathbb{R}$ of real numbers. By $(\cdot,\cdot)_{\mathcal{H}}$ we denote the inner product in $\mathcal{H}$, which extends to the duality pairing ${}_{\mathcal{V}^*}\langle \cdot,\cdot\rangle_{\mathcal{V}}$ on $\mathcal{V}^* \times \mathcal{V}$ by continuity.

Example 2.1. *The prototypical example of a problem of the above form is Poisson's equation on a bounded open Lipschitz domain $D \subset \mathbb{R}^d$, subject to a homogeneous Dirichlet boundary condition on ∂D, in which case:*

$$A = -\Delta,\ \mathcal{V} = \mathrm{H}^1_0(D),\ \mathcal{H} = \mathrm{L}^2(D),\ \mathcal{V}^* = \mathrm{H}^{-1}(D) := \mathrm{H}^1_0(D)^*,\ \mathfrak{a}(u,v) = \int_D \nabla u \cdot \nabla v \,\mathrm{d}x.$$

We shall assume that A is selfadjoint, i.e., $A = A^*$ (which implies that the bilinear form $\mathfrak{a}(\cdot,\cdot)$ is *symmetric* on $\mathcal{V} \times \mathcal{V}$) and *coercive* on $\mathcal{V}$, i.e., there exists a real number $\gamma_0 > 0$ such that

$$\forall u \in \mathcal{V}: \qquad \mathfrak{a}(u,u) = {}_{\mathcal{V}^*}\langle Au, u\rangle_{\mathcal{V}} \geq \gamma_0 \|u\|^2_{\mathcal{V}}. \tag{2.3}$$

Our assumption $A \in \mathcal{L}(\mathcal{V},\mathcal{V}^*)$ implies the existence of a positive real number $\gamma_1 = \|A\|_{\mathcal{L}(\mathcal{V},\mathcal{V}^*)} \geq \gamma_0$ such that

$$\forall u,v \in \mathcal{V}: \qquad |\mathfrak{a}(u,v)| \leq \gamma_1 \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}; \tag{2.4}$$

i.e., the bilinear form $\mathfrak{a}$ is *bounded*.

Hence, (2.3) and (2.4) imply that the *energy norm* $\|\cdot\|_{\mathfrak{a}}$, defined on $\mathcal{V}$ by $\|v\|_{\mathfrak{a}} := [\mathfrak{a}(v,v)]^{1/2}$, $v \in \mathcal{V}$, is equivalent on $\mathcal{V}$ with the norm $\|\cdot\|_{\mathcal{V}}$; viz.,

$$\forall v \in \mathcal{V}: \qquad \gamma_0 \|v\|_{\mathcal{V}}^2 \le \|v\|_{\mathfrak{a}}^2 \le \gamma_1 \|v\|_{\mathcal{V}}^2.$$

We recall the following version of the Hilbert–Schmidt theorem (cf. Lemma 15 in [14]).

Theorem 2.2. *Suppose that $\mathcal{H}$ and $\mathcal{V}$ are separable Hilbert spaces, with $\mathcal{V}$ densely and compactly embedded in $\mathcal{H}$. Let $\mathfrak{a} : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ be a nonzero, symmetric, coercive and bounded bilinear form. Then, there exists a sequence of real numbers $(\lambda_n)_{n=1}^{\infty}$ and a sequence of unit $\mathcal{H}$-norm elements $(\varphi_n)_{n=1}^{\infty}$ in $\mathcal{V}$, which solve the following eigenvalue problem: find $\lambda \in \mathbb{R}$ and $\varphi \in \mathcal{V}\setminus\{0\}$ such that*

$$\mathfrak{a}(\varphi, v) = \lambda(\varphi, v)_{\mathcal{H}} \qquad \forall v \in \mathcal{V}.$$

The real numbers λ_n, $n \in \mathbb{N}$, which can be assumed to be in increasing order with respect to the index $n \in \mathbb{N}$, are positive, bounded away from 0, and $\lim_{n\to\infty} \lambda_n = +\infty$.

In addition, the φ_n, $n \in \mathbb{N}$, form an orthonormal system in $\mathcal{H}$ whose closed span in $\mathcal{H}$ is equal to $\mathcal{H}$, and the scaled elements $\varphi_n/\sqrt{\lambda_n}$, $n \in \mathbb{N}$, form an orthonormal system with respect to the inner product defined by the bilinear form $\mathfrak{a}(\cdot,\cdot)$, whose closed span with respect to the norm $\|\cdot\|_{\mathfrak{a}}$ induced by $\mathfrak{a}(\cdot,\cdot)$ is equal to $\mathcal{V}$. Furthermore,

$$h = \sum_{n=1}^{\infty} (h,\varphi_n)_{\mathcal{H}}\varphi_n \qquad \textit{and} \qquad \|h\|_{\mathcal{H}}^2 = \sum_{n=1}^{\infty} (h,\varphi_n)_{\mathcal{H}}^2 \qquad \forall h \in \mathcal{H} \tag{2.5}$$

and

$$v = \sum_{n=1}^{\infty} \mathfrak{a}\left(v, \frac{\varphi_n}{\sqrt{\lambda_n}}\right)\frac{\varphi_n}{\sqrt{\lambda_n}} \qquad \textit{and} \qquad \|h\|_{\mathfrak{a}}^2 = \sum_{n=1}^{\infty} \mathfrak{a}\left(v, \frac{\varphi_n}{\sqrt{\lambda_n}}\right)^2 \qquad \forall v \in \mathcal{V};$$

and, in addition,

$$h \in \mathcal{H} \quad \textit{and} \quad \sum_{n=1}^{\infty} \lambda_n (h,\varphi_n)_{\mathcal{H}}^2 < \infty \iff h \in \mathcal{V}.$$

By virtue of Theorem 2.2, there exists a countable family $\sigma \subset \mathbb{R}_+$ of eigenvalues λ of A, bounded away from zero, whose accumulation point is at $+\infty$, and an $\mathcal{H}$-orthonormal sequence $\{\varphi_\lambda : \lambda \in \sigma\}$ of eigenfunctions $\varphi_\lambda \in \mathcal{V}$, i.e.,

$$A\varphi_\lambda = \lambda\varphi_\lambda, \;\; \lambda \in \sigma, \quad \text{and} \quad (\varphi_\lambda, \varphi_{\lambda'})_{\mathcal{H}} = \delta_{\lambda\lambda'}, \;\; \lambda, \lambda' \in \sigma.$$

We then have by (2.5), for each $v \in \mathcal{H}$, the Fourier series expansion

$$v = \sum_{\lambda\in\sigma} v_\lambda \varphi_\lambda, \quad v_\lambda := (v, \varphi_\lambda)_{\mathcal{H}}$$

and Parseval's identity

$$\|v\|_{\mathcal{H}}^2 = \sum_{\lambda\in\sigma} |v_\lambda|^2.$$

In particular, therefore, the system $\{\varphi_\lambda : \lambda \in \sigma\}$ forms a normalized Riesz basis of $\mathcal{H}$. The next two lemmas whose proofs are elementary show that renormalized versions of $\{\varphi_\lambda : \lambda \in \sigma\}$ constitute Riesz bases in $\mathcal{V}$ and $\mathcal{V}^*$ as well.

Lemma 2.3. *The following two-sided bound holds for each $v \in \mathcal{V}$:*

$$\gamma_0 \|v\|_{\mathcal{V}}^2 \le \sum_{\lambda\in\sigma} \lambda\, |v_\lambda|^2 \le \gamma_1 \|v\|_{\mathcal{V}}^2.$$

For $f \in \mathcal{V}^*$, we have that

$$f = \sum_{\lambda\in\sigma} f_\lambda \varphi_\lambda, \qquad f_\lambda := {}_{\mathcal{V}^*}\langle f, \varphi_\lambda\rangle_{\mathcal{V}}.$$

Lemma 2.4. *The following two-sided bound holds for each $f \in \mathcal{V}^*$:*

$$\frac{1}{\gamma_1} \|f\|_{\mathcal{V}^*}^2 \le \sum_{\lambda\in\sigma} \lambda^{-1} |f_\lambda|^2 \le \frac{1}{\gamma_0} \|f\|_{\mathcal{V}^*}^2.$$

As
$$u = A^{-1} f = \sum_{\lambda\in\sigma} u_\lambda \varphi_\lambda, \quad \text{with} \quad u_\lambda = \lambda^{-1} f_\lambda \ \text{ for } \ \lambda\in\sigma,$$
it follows that
$$\sum_{\lambda\in\sigma} \lambda\, |u_\lambda|^2 = \sum_{\lambda\in\sigma} \lambda^{-1}\, |f_\lambda|^2, \tag{2.6}$$
where, by Lemma 2.3, the left-hand side of the equality belongs to the interval $\|u\|^2_{\mathcal{V}}\,[\gamma_0,\gamma_1]$ and, by Lemma 2.3, the right-hand side belongs to the interval $\|f\|^2_{\mathcal{V}^*}\,\left[\gamma_1^{-1},\gamma_0^{-1}\right]$. Here, for a nonnegative real number α and a nonempty bounded closed interval $[s,t]\subset\mathbb{R}$, the notation $x\in\alpha[s,t]$ and $x\in[s,t]\alpha$ are abbreviations for the two-sided inequality: $\alpha s\le x\le \alpha t$. We deduce from (2.6) in particular that
$$\gamma_0\,\|u\|_{\mathcal{V}} \le \|f\|_{\mathcal{V}^*} \le \gamma_1\,\|u\|_{\mathcal{V}}.$$
Thus, the linear operator A^{-1} is a (bi-Lipschitz) quasi-isometric isomorphism between $\mathcal{V}^*$ and $\mathcal{V}$, when these spaces are equipped with the norms $\|\cdot\|_{\mathcal{V}^*}$ and $\|\cdot\|_{\mathcal{V}}$, respectively.

2.2. **Abstract parabolic equations.** Given $T>0$ and the triple (2.2), we consider the abstract parabolic differential equation
$$\partial_t u + Au = f \ \text{ in } \ Q := (0,T)\times D, \tag{2.7}$$
where $A\in\mathcal{L}(\mathcal{V},\mathcal{V}^*)$, with $\|A\|_{\mathcal{L}(\mathcal{V},\mathcal{V}^*)} = \gamma_1>0$, satisfies $A=A^*$, and
$$\exists\gamma_0>0 \quad \forall v\in\mathcal{V}: \qquad \mathfrak{a}(v,v) := {}_{\mathcal{V}^*}\langle Av,v\rangle_{\mathcal{V}} \ge \gamma_0\|v\|^2_{\mathcal{V}}.$$
As our aim is to develop the numerical analysis of space-time adaptive Galerkin discretizations of infinite-dimensional parabolic problems, we begin, following [24], by presenting weak formulations that are amenable to adaptive Galerkin discretizations. To this end, we consider the Bochner spaces $\mathrm{L}^2(0,T;\mathcal{V})$, $\mathrm{L}^2(0,T;\mathcal{H})$ and $\mathrm{L}^2(0,T;\mathcal{V}^*)$ and we define
$$\mathrm{H}^1(0,T;\mathcal{V}) := \{u\in\mathrm{L}^2(0,T;\mathcal{V}) : u'\in\mathrm{L}^2(0,T;\mathcal{V})\},$$
as well as
$$\mathrm{H}^1_{0,\{0\}}(0,T;\mathcal{V}) := \{u\in\mathrm{H}^1(0,T;\mathcal{V}) : u(0)=0 \ \text{ in } \mathcal{V}\}, \tag{2.8a}$$
$$\mathrm{H}^1_{0,\{T\}}(0,T;\mathcal{V}) := \{u\in\mathrm{H}^1(0,T;\mathcal{V}) : u(T)=0 \ \text{ in } \mathcal{V}\}. \tag{2.8b}$$
Here and throughout the rest of the paper u' will signify $\mathrm{d}u/\mathrm{d}t$ or $\partial u/\partial t$, depending on the context.

In the variational formulation of the parabolic problem, an important role is played by the space $\mathcal{X}$ defined by
$$\mathcal{X} := \mathrm{L}^2(0,T;\mathcal{V})\cap\mathrm{H}^1(0,T;\mathcal{V}^*),$$
which we equip with the norm $\|\cdot\|_{\mathcal{X}}$ defined by
$$\|v\|_{\mathcal{X}} := \left(\|v\|^2_{\mathrm{L}^2(0,T;\mathcal{V})} + \|v'\|^2_{\mathrm{L}^2(0,T;\mathcal{V}^*)}\right)^{\frac{1}{2}}.$$
With $\mathcal{V}$, $\mathcal{H}$ and $\mathcal{V}^*$ as in the triple (2.2) the following continuous embedding holds:
$$\mathcal{X}\hookrightarrow \mathrm{C}([0,T];\mathcal{H}) \tag{2.9}$$
(in the sense that any $v\in\mathcal{X}$ is equal almost everywhere to a function that is uniformly continuous as a mapping from the nonempty closed interval $[0,T]$ of the real line into $\mathcal{H}$).

Therefore, for $u\in\mathcal{X}$ and $0\le t\le T$, the values $u(t)$ are well-defined in $\mathcal{H}$ and there exists a constant $C=C(T)>0$ such that
$$\forall u\in\mathcal{X} \quad \forall t\in[0,T]: \qquad \|u(t)\|_{\mathcal{H}} \le C\|u\|_{\mathcal{X}}.$$
In particular, for $u\in\mathcal{X}$ the values $u(0)$ and $u(T)$ are well-defined in $\mathcal{H}$ and
$$\mathcal{X}_{0,\{0\}} := \{u\in\mathcal{X} : u(0)=0 \text{ in } \mathcal{H}\}, \quad \mathcal{X}_{0,\{T\}} := \{u\in\mathcal{X} : u(T)=0 \text{ in } \mathcal{H}\}$$
are closed linear subspaces of $\mathcal{X}$. Henceforth we shall write $\mathcal{Y}=\mathrm{L}^2(0,T;\mathcal{V})$ and we denote by $\mathcal{Y}^*$ the dual space $\mathrm{L}^2(0,T;\mathcal{V}^*)$ of $\mathcal{Y}$, identifying $\mathrm{L}^2(0,T;\mathcal{H})$ with its own dual.

2.3. **First space-time variational formulation.** We consider (2.7) in the special case when

$$Lu := \partial_t u + Au = f \in \mathcal{Y}^* \text{ in } Q := (0,T) \times D, \quad u(0) = u_0 \in \mathcal{H}, \tag{2.10}$$

in conjunction with a suitable homogeneous boundary condition incorporated in the (domain of) definition of the linear operator A. We begin by considering the case when the initial condition is also homogeneous, i.e., $u_0 = 0 \in \mathcal{H}$.

The first space-time variational formulation of the parabolic problem (2.10) is based on the bilinear form

$$(u,v) \in \mathcal{X} \times \mathcal{Y} \mapsto \mathfrak{B}(u,v) := \int_0^T \Big({}_{\mathcal{V}^*}\langle u', v\rangle_{\mathcal{V}} + \mathfrak{a}(u,v)\Big)\,\mathrm{d}t \in \mathbb{R}. \tag{2.11}$$

Then, the space-time weak formulation of the parabolic problem (2.10) with homogeneous initial condition $u_0 = 0$ in $\mathcal{H}$ reads as follows: given $f \in \mathcal{Y}^*$, find $u \in \mathcal{X}_{0,\{0\}}$ such that

$$\mathfrak{B}(u,v) = f(v) \quad \forall v \in \mathcal{Y}. \tag{2.12}$$

2.4. **Second space-time variational formulation.** In (2.12), the initial condition $u(0) = 0$ is incorporated as essential condition in the function space $\mathcal{X}_{0,\{0\}}$ in which the solution to the problem was sought. To accommodate a nonhomogeneous initial condition, one may proceed in (at least) two different ways. If

$$u(0) = u_0 \neq 0 \ \text{ in } \mathcal{H}, \tag{2.13}$$

then one can for example first subtract from u a function $u_p \in \mathcal{X}$ such that $u_p|_{t=0} = u_0$; we may, in particular, choose for this purpose

$$u_p = \mathrm{e}^{-t} u_0, \text{ for } u_0 \in \mathcal{V} \subset \mathcal{H}.$$

The disadvantage of this approach is that $u_0 \in \mathcal{V}$ is needed (instead of $u_0 \in \mathcal{H}$), which can be viewed as an unnecessarily restrictive demand on the regularity of the initial datum u_0.

Alternatively, in order to relax the regularity requirement $u_0 \in \mathcal{V}$ to $u_0 \in \mathcal{H}$, one may impose (2.13) weakly, either by enforcing it via a multiplier as in [24] or through a space-time variational formulation, which incorporates it as a *natural boundary condition*, as follows: we integrate the time derivative by parts using that

$$\int_0^T {}_{\mathcal{V}^*}\langle u', v\rangle_{\mathcal{V}}\,\mathrm{d}t = -\int_0^T {}_{\mathcal{V}^*}\langle v', u\rangle_{\mathcal{V}}\,\mathrm{d}t + (u,v)|_0^T, \quad u,v \in \mathrm{L}^2(0,T;\mathcal{V}) \cap \mathrm{H}^1(0,T;\mathcal{V}^*). \tag{2.14}$$

If $u(0) = u_0 \neq 0$ in $\mathcal{H}$, we require that

$$v(T) = 0. \tag{2.15}$$

Thus, (2.14) and (2.15) lead to the weak formulation (2.18) below. To state it, we recall the spaces $\mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$ and

$$\mathcal{X} = \mathrm{L}^2(0,T;\mathcal{V}) \cap \mathrm{H}^1(0,T;\mathcal{V}^*) \tag{2.16}$$

and the subspaces (cf. also (2.8))

$$\mathcal{X}_{0,\{0\}} := \{u \in \mathcal{X} : \ u(0) = 0 \ \text{ in } \mathcal{H}\}, \tag{2.17a}$$

$$\mathcal{X}_{0,\{T\}} := \{u \in \mathcal{X} : \ u(T) = 0 \ \text{ in } \mathcal{H}\}, \tag{2.17b}$$

equipped with the norm of $\mathcal{X}$; we shall write $\|\cdot\|_{\mathcal{X}_{0,\{0\}}}$ and $\|\cdot\|_{\mathcal{X}_{0,\{T\}}}$ to indicate that the norm of $\mathcal{X}$ is applied to an element of $\mathcal{X}_{0,\{0\}}$ and $\mathcal{X}_{0,\{T\}}$, respectively. Thanks to the continuous embedding (2.9), the spaces $\mathcal{X}_{0,\{0\}}$ and $\mathcal{X}_{0,\{T\}}$ are correctly defined and are closed subspaces of $\mathcal{X}$; in particular the expression (2.14) is meaningful for $u, v \in \mathcal{X}$. The variational form of the parabolic problem (2.10) with weak enforcement of the initial condition, which we shall also refer to as the *space-time adjoint weak formulation*, then reads: given $u_0 \in \mathcal{H}$ and $f \in \mathcal{X}^*_{0,\{T\}}$, find

$$u \in \mathcal{Y} : \quad \mathfrak{B}^*(u,v) = \ell^*(v) \quad \forall v \in \mathcal{X}_{0,\{T\}}. \tag{2.18}$$

Here the bilinear form $\mathfrak{B}^*(\cdot,\cdot)$ is given by

$$\mathfrak{B}^*(u,v) := \int_0^T \big(-{}_{\mathcal{V}^*}\langle v', u\rangle_{\mathcal{V}} + \mathfrak{a}(u,v)\big)\,\mathrm{d}t \tag{2.19}$$

and the linear functional ℓ^* is defined by

$$\ell^*(v) := {}_{\mathcal{X}^*}\langle f, v\rangle_{\mathcal{X}} + (u_0, v(0))_{\mathcal{H}}. \tag{2.20}$$

Clearly, for $f \in \mathcal{X}^*_{0,\{T\}} \simeq \mathrm{L}^2(I;\mathcal{V}^*) + (\mathrm{H}^1_{0,\{T\}})^*(I;\mathcal{V}) \simeq \mathrm{L}^2(I)\otimes\mathcal{V}^* + (\mathrm{H}^1_{0,\{T\}}(I))^*\otimes\mathcal{V}$ and, for any $u_0 \in \mathcal{H}$, the functional ℓ^* in (2.20) is linear and continuous on $\mathcal{X}_{0,\{T\}}$; i.e., there exists a constant $C > 0$ such that

$$\forall v \in \mathcal{X}_{0,\{T\}}: \quad |\ell^*(v)| \le C\big(\|f\|_{\mathcal{X}^*_{0,\{T\}}} + \|u_0\|_{\mathcal{H}}\big)\, \|v\|_{\mathcal{X}_{0,\{T\}}}.$$

2.5. **Well-posedness.** The well-posedness of the variational problems (2.12) and (2.18) is an immediate consequence of the following result.

Theorem 2.5. *Suppose that $\mathcal{V}$ and $\mathcal{H}$ are separable Hilbert spaces over the field $\mathbb{R}$ such that $\mathcal{V} \hookrightarrow \mathcal{H} \simeq \mathcal{H}^* \hookrightarrow \mathcal{V}^*$ with dense embeddings. Assume further that the bilinear form $\mathfrak{a}(\cdot,\cdot) : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ is continuous, i.e.,*

$$\exists M_a > 0 \quad \forall w, v \in \mathcal{V}: \quad |\mathfrak{a}(w,v)| \le M_a \|v\|_{\mathcal{V}}\|w\|_{\mathcal{V}}, \tag{2.21}$$

and coercive in the sense that it satisfies a Garding inequality, i.e., there exists $m_a > 0$ and $\kappa \ge 0$ such that

$$\forall v \in \mathcal{V}: \quad \mathfrak{a}(v,v) \ge m_a\|v\|^2_{\mathcal{V}} - \kappa\|v\|^2_{\mathcal{H}}. \tag{2.22}$$

Under these hypotheses, there exists a positive constant C such that the bilinear form $\mathfrak{B}(\cdot,\cdot)$, as defined in (2.11)*, satisfies the following inf-sup condition:*

$$\inf_{u\in\mathcal{X}_{0,\{0\}}\setminus\{0\}} \ \sup_{v\in\mathcal{Y}\setminus\{0\}} \ \frac{\mathfrak{B}(u,v)}{\|u\|_{\mathcal{X}_{0,\{0\}}}\|v\|_{\mathcal{Y}}} \ge C$$

and

$$\forall v \in \mathcal{Y}\setminus\{0\} : \sup_{u\in\mathcal{X}_{0,\{0\}}} \mathfrak{B}(u,v) > 0.$$

Analogously, there exists a positive constant C such that the bilinear form $\mathfrak{B}^*(\cdot,\cdot)$ in (2.19) *satisfies the following inf-sup condition:*

$$\inf_{u\in\mathcal{Y}\setminus\{0\}} \ \sup_{v\in\mathcal{X}_{0,\{T\}}\setminus\{0\}} \ \frac{\mathfrak{B}^*(u,v)}{\|u\|_{\mathcal{Y}}\|v\|_{\mathcal{X}_{0,\{T\}}}} \ge C$$

and

$$\forall v \in \mathcal{X}_{0,\{T\}}\setminus\{0\} : \qquad \sup_{u\in\mathcal{Y}} \mathfrak{B}^*(u,v) > 0.$$

Furthermore, the bilinear forms $\mathfrak{B}(\cdot,\cdot)$ and $\mathfrak{B}^*(\cdot,\cdot)$ defined in (2.11) *and in* (2.18)*, respectively, induce boundedly invertible, linear operators $B \in \mathcal{L}(\mathcal{X}_{0,\{0\}}, \mathcal{Y}^*)$ and $B^* \in \mathcal{L}(\mathcal{Y}, (\mathcal{X}_{0,\{T\}})^*)$, where the spaces $\mathcal{X}_{0,\{0\}}$, $\mathcal{X}_{0,\{T\}}$ are defined in* (2.16) *and* (2.17)*, and $\mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$.*

The proof is analogous to that of Theorem 5.1 in [24]. The following result is now an immediate consequence of Theorem 2.5.

Theorem 2.6. *For every $u_0 \in \mathcal{H}$ and every $f \in \mathcal{X}^*_{0,\{T\}} = \mathrm{L}^2(I;\mathcal{V}^*) + \mathrm{H}^{-1}_{0,\{T\}}(I;\mathcal{V})$, there exists a unique weak solution $u \in \mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$ of* (2.10) *in the sense that u satisfies* (2.18)*. Moreover, the operator $L : \mathcal{Y} \to \mathcal{X}^*_{0,\{T\}}$ is an isomorphism.*

In the next section we apply these abstract results to a class of Fokker–Planck equations, posed on high-dimensional domains, that arise in bead-spring chain models in kinetic models of dilute polymers (see, for example, [2] and [3]).

3. Fokker–Planck equation

Suppose that K is a fixed positive integer and $d \in \{2,3\}$. Suppose further that $D_k \subset \mathbb{R}^d$ is either a bounded ball in $\mathbb{R}^d$ centred at the origin of $\mathbb{R}^d$ and of fixed radius $\sqrt{b_k}$, where $b_k > 0$, $k = 1,\dots,K$, or that $D_k = \mathbb{R}^d$ for all $k = 1,\dots,K$. By identifying $\mathbb{R}^d$ with a ball of radius $+\infty$, we can unify the two scenarios by taking $b_k \in (0,+\infty]$ and by assuming that the b_k are either finite for all $k = 1,\dots,K$ or infinite for all $k = 1,\dots,K$. We define $D := D_1 \times \cdots \times D_K$. On the interval $[0, \frac{b_k}{2})$ we consider the function $U_k \in \mathrm{C}^1[0, \frac{b_k}{2})$, referred to as a *potential*, such that $U_k(0) = 0$, U_k is strictly monotonic increasing and $\lim_{s\to b_k-} U_k = +\infty$, $k = 1,\dots,K$. We then associate with U_k the *partial Maxwellian*, defined by

$$M_k(q_k) := \frac{1}{\mathcal{Z}_k}\exp\left(-U_k\left(\tfrac{1}{2}|q_k|^2\right)\right), \quad q_k \in D_k,$$

where

$$\mathcal{Z}_k := \int_{D_k} \exp\left(-U_k\left(\tfrac{1}{2}|p_k|^2\right)\right)\,\mathrm{d}p_k,$$

for $k = 1,\dots,K$, and we define the (full) Maxwellian

$$M(q) := M_1(q_1)\cdots M_K(q_K), \qquad q = (q_1^\top,\dots,q_K^\top)^\top \in D_1 \times \cdots \times D_K = D \subseteq \mathbb{R}^{Kd}.$$

Clearly, $M(q) > 0$ on D, $\int_D M(q)\,\mathrm{d}q = 1$ and $\lim_{|q_k|\to b_k} M_k(q_k) = 0$, $k = 1,\dots,K$. When the domains D_k are bounded balls, we shall suppose that there exist positive constants C_{k1} and C_{k2}, and real numbers $\alpha_k > 1$, $k = 1,\dots,K$, such that

$$0 < C_{k1} \le \exp\left(-U_k\left(\tfrac{1}{2}|q_k|^2\right)\right)/(\mathrm{dist}(q_k,\partial D_k))^{\alpha_k} \le C_{k2} < \infty, \quad k = 1,\dots,K.$$

Alternatively, when $D_k = \mathbb{R}^d$ for all $k = 1,\dots,K$, we shall assume that $U_k(s) = s$, $k = 1,\dots,K$.

This section is devoted to the proof of existence and uniqueness of weak solutions to the Fokker–Planck equation

$$\partial_t\psi - \sum_{i,j,=1}^{K} A_{ij}\nabla_{q_j}\cdot\left(M\,\nabla_{q_i}\left(\frac{\psi}{M}\right)\right) = 0, \quad (q,t) \in D\times(0,T], \tag{3.23}$$

subject to the initial condition

$$\psi(q,0) = \psi_0(q), \quad q \in D, \tag{3.24}$$

where $A \in \mathbb{R}^{K\times K}$ is a symmetric positive definite matrix with minimal eigenvalue $\gamma_0 > 0$ and maximal eigenvalue $\gamma_1 > 0$, and $\psi_0 \in \mathrm{L}^1(D)$ is a nonnegative function such that $\int_D \psi_0\,\mathrm{d}q = 1$. As ψ_0 is a probability density function, and this property needs to be propagated during the course of the evolution over the time interval $[0,T]$, the boundary condition on $\partial D \times [0,T]$ needs to be chosen so that $\psi(\cdot,t)$ remains a nonnegative function for all $t \in [0,T]$, and $\int_D \psi(q,t)\,\mathrm{d}q = 1$ for all $t \in [0,T]$. This can be achieved, formally at least, by demanding that

$$\sum_{i=1}^{K} A_{ij}\left(M\,\nabla_{q_i}\left(\frac{\psi}{M}\right)\right)\cdot\frac{q_j}{|q_j|} \to 0 \quad \text{as} \quad |q_j|^2 \to b_j, \qquad \text{for all } j = 1,\dots,K, \tag{3.25}$$

where either $b_j \in (0,\infty)$ for all $j = 1,\dots,K$ when D_j is a bounded ball of radius $\sqrt{b_j}$; or $b_j := +\infty$ for all $j = 1,\dots,K$ when $D_j = \mathbb{R}^d$. By writing

$$\widehat{\psi} := \frac{\psi}{M} \quad \text{and} \quad \widehat{\psi}_0 := \frac{\psi_0}{M},$$

the initial-value problem (3.23), (3.24) can be restated as follows:

$$M\,\partial_t\widehat{\psi} - \sum_{i,j,=1}^{K} A_{ij}\nabla_{q_j}\cdot\left(M\,\nabla_{q_i}\widehat{\psi}\right) = 0, \quad (q,t) \in D\times(0,T], \tag{3.26}$$

subject to the initial condition

$$M(q)\,\widehat{\psi}(q,0) = M(q)\,\widehat{\psi}_0(q), \quad q \in D, \tag{3.27}$$

together with the (formal) boundary condition

$$\sum_{i=1}^{K} A_{ij}\left(M\,\nabla_{q_i}\widehat{\psi}\right)\cdot\frac{q_j}{|q_j|}\to 0 \quad\text{as}\quad |q_j|^2\to b_j, \qquad \text{for all } j=1,\dots,K, \tag{3.28}$$

and the adopted convention that $b_j := +\infty$ when $D_j = \mathbb{R}^d$. We consider the Maxwellian-weighted L^2 space

$$\mathrm{L}^2_M(D) := \{\widehat{\varphi}\in \mathrm{L}^1_{\mathrm{loc}}(D)\,:\,\sqrt{M}\,\widehat{\varphi}\in \mathrm{L}^2(D)\},$$

equipped with the inner product $(\cdot,\cdot)_{\mathrm{L}^2_M(D)}$ and norm $\|\cdot\|_{\mathrm{L}^2_M(D)}$, defined, respectively, by

$$(\widehat{\psi},\widehat{\varphi})_{\mathrm{L}^2_M(D)} := \int_D M(q)\,\widehat{\psi}(q)\,\widehat{\varphi}(q)\,\mathrm{d}q \quad\text{and}\quad \|\widehat{\varphi}\|_{\mathrm{L}^2_M(D)} := (\widehat{\varphi},\widehat{\varphi})^{\frac{1}{2}}_{\mathrm{L}^2_M(D)},$$

and the associated Maxwellian-weighted H^1 space

$$\mathrm{H}^1_M(D) := \{\widehat{\varphi}\in \mathrm{L}^2_M(D)\,:\,\nabla_{q_k}\widehat{\varphi}\in \mathrm{L}^2_M(D), \quad k=1,\dots,K\}$$

equipped with the inner product $(\cdot,\cdot)_{\mathrm{H}^1_M(D)}$ and norm $\|\cdot\|_{\mathrm{H}^1_M(D)}$ defined, respectively, by

$$(\widehat{\psi},\widehat{\varphi})_{\mathrm{H}^1_M(D)} = (\widehat{\psi},\widehat{\varphi})_{\mathrm{L}^2_M(D)} + \sum_{k=1}^{K}(\nabla_{q_k}\widehat{\psi},\nabla_{q_k}\widehat{\varphi})_{[\mathrm{L}^2_M(D)]^d} \quad\text{and}\quad \|\widehat{\varphi}\|_{\mathrm{H}^1_M(D)} := (\widehat{\varphi},\widehat{\varphi})^{\frac{1}{2}}_{\mathrm{H}^1_M(D)}.$$

We note that when D_k are bounded open balls, the open set D, as a Cartesian product of bounded Lipschitz domains, is also a bounded Lipschitz domain (cf. the footnote on p.10 of [14]).

Under our assumptions on the potentials U_k, $k=1,\dots,K$, $\mathrm{L}^2_M(D)$ and $\mathrm{H}^1_M(D)$ are separable Hilbert spaces over the field of real numbers (cf. Theorems 2.7.1, 2.8.1 and 8.10.2 in [19]). Clearly, $\mathrm{H}^1_M(D)$ is continuously embedded in $\mathrm{L}^2_M(D)$. In fact the embedding $\mathrm{H}^1_M(D)\hookrightarrow \mathrm{L}^2_M(D)$ is dense and compact (in the case when of each D_k, $k=1,\dots,K$, is a bounded ball we refer for the proofs of these statements to the Appendix in Barrett & Süli [1], and Appendices B, C, D in the extended version of Barrett & Süli [2]; in the case when $D_k=\mathbb{R}^d$ for each $k=1,\dots,K$, we refer to the Appendices A and D in Barrett & Süli [3]).

Adopting the notations introduced in the previous sections, we take $\mathcal{H} := \mathrm{L}^2_M(D)$, $\mathcal{V} := \mathrm{H}^1_M(D)$, and consider the linear differential operator

$$A\widehat{\varphi} := -\sum_{i,j=1}^{K} A_{ij}\nabla_{q_j}\cdot\left(M\,\nabla_{q_i}\widehat{\varphi}\right), \quad \widehat{\varphi}\in\mathcal{V},$$

that maps $\mathcal{V}$ into its dual space $\mathcal{V}^*$ (with respect to the pivot space $\mathcal{H}$).

We strengthen our original assumption $\psi_0\in \mathrm{L}^1(D)$ by demanding that $\widehat{\psi}_0\in \mathrm{L}^2_M(D)$ (note that $\|\psi_0\|_{\mathrm{L}^1(D)} \le \|\widehat{\psi}_0\|_{\mathrm{L}^2_M(D)} = \|\widehat{\psi}_0\|_{\mathcal{H}}$ for all $\widehat{\psi}_0\in\mathcal{H}=\mathrm{L}^2_M(D)$), and we consider the following weak formulation of the Fokker–Planck initial-boundary-value problem: given $\widehat{\psi}_0\in\mathcal{H} := \mathrm{L}^2_M(D)$, find $\widehat{\psi}\in\mathcal{Y} := \mathrm{L}^2(0,T;\mathcal{V}) = \mathrm{L}^2(0,T;\mathrm{H}^1_M(D))$ such that

$$\mathfrak{B}^*(\widehat{\psi},\widehat{\varphi}) := \int_0^T \left(-{}_{\mathcal{V}^*}\langle\widehat{\varphi}',\widehat{\psi}\rangle_{\mathcal{V}} + \mathfrak{a}(\widehat{\psi},\widehat{\varphi})\right)\mathrm{d}t = (\widehat{\psi}_0,\widehat{\varphi}(0))_{\mathcal{H}} \quad \forall\widehat{\varphi}\in\mathcal{X}_{0,\{T\}}. \tag{3.29}$$

Here, as in the previous sections, $\widehat{\varphi}'$ signifies the derivative of $\widehat{\varphi}$ with respect to the variable t.

On recalling the Brascamp–Lieb inequality (cf. [18], Corollary 2.2) in the case of $b_k<\infty$, $k=1,\dots,K$; and the Poincaré inequality for Gaussian measures (cf. the work of Beckner [4], Theorem 1, inequality (3) with $p=1$) when $b_k=\infty$, $k=1,\dots,K$, we deduce from Theorem 2.6 the existence of a unique weak solution $\widehat{\psi}\in \mathrm{L}^2(0,T;\mathrm{H}^1_M(D))$ to the Fokker–Planck equation (3.26), (3.27). By choosing in the weak formulation (3.29) the test function $\widehat{\varphi}$ from the dense linear subspace $\{\widehat{\varphi}\in \mathrm{C}^\infty([0,T];\mathrm{H}^1_M(D))\,:\,\widehat{\varphi}(\cdot,T)=0\}$ of $\mathcal{X}_{0,\{T\}}$, it then follows from (3.29) that $\widehat{\psi}\in \mathrm{H}^1(0,T;\mathrm{H}^1_M(D)^*)$, and therefore $\widehat{\psi}\in \mathrm{L}^2(0,T;\mathrm{H}^1_M(D))\cap \mathrm{H}^1(0,T;\mathrm{H}^1_M(D)^*)$. Thus, returning to our original variable ψ, we deduce the existence of a unique weak solution ψ to (3.23), (3.24), with $\psi/M\in \mathrm{L}^2(0,T;\mathrm{H}^1_M(D))\cap \mathrm{H}^1(0,T;\mathrm{H}^1_M(D)^*)$. The boundary condition (3.25) (or, equivalently, (3.28)) is imposed weakly, by demanding that the function ψ (or, equivalently, $\widehat{\psi}$) belongs to the appropriate function space for weak solutions.

Henceforth we shall concentrate on the simplest case from the technical point of view, when $D_k = \mathbb{R}^d$ and $U_k(s) = s$ for all $k = 1, \ldots, K$; then,

$$M_k(q_k) = \frac{1}{\mathcal{Z}_k} \exp\left(-\tfrac{1}{2}|q_k|^2\right), \quad q_k \in \mathbb{R}^d, \quad \text{where} \quad \mathcal{Z}_k := \int_{\mathbb{R}^d} \exp\left(-\tfrac{1}{2}|p_k|^2\right) \mathrm{d}p_k, \quad k = 1, \ldots, K,$$

and

$$M(q) = M_1(q_1) \cdots M_K(q_K) = \frac{1}{\mathcal{Z}} \exp\left(-\tfrac{1}{2}|q|^2\right), \quad q = (q_1^\top, \ldots, q_K^\top)^\top \in \mathbb{R}^{Kd},$$

where $|q|^2 = |q_1|^2 + \cdots + |q_K|^2$ and $\mathcal{Z} = \prod_{i=1}^K \mathcal{Z}_k$, $k = 1, \ldots, K$.

We are particularly interested in the case when $K \gg 1$; specifically, our objective is to develop adaptive numerical algorithms with quantitative optimality properties, which are *dimensionally robust*, i.e., whose computational complexity does not exhibit the curse of dimensionality. To this end, we shall embed the weak formulation of the Fokker–Planck initial-boundary-value problem (3.29), with a finite, but arbitrarily large number of independent variables q_k, $k = 1, \ldots, K$, into a problem with countably many variables, q_k, $k = 1, 2, \ldots$, corresponding to formally setting $K = +\infty$. Thus, instead of working on $D = \mathbb{R}^{Kd}$, we shall be interested in the case when $D = \mathbb{R}^\infty$ (cf. the next section for the definition of $\mathbb{R}^\infty$), which we shall equip with a finite measure, the *Gaussian measure* on $\mathbb{R}^\infty$. We shall employ for this purpose the notion of Gaussian measure on a separable Hilbert space. Our exposition in the next section closely follows Sections 1.1, 1.2, 9.1 and 9.2 in the monograph of Da Prato and Zabczyk [12].

4. Gaussian measures

Suppose that H is a separable Hilbert space with norm $\|\cdot\|_H$ and inner product $(\cdot,\cdot)_H$ over the field of real numbers. We denote by $\mathrm{L}(H)$ the space of all bounded linear operators from H into H equipped with the associated operator norm $\|\cdot\|_{\mathrm{L}(H)}$, and we denote by $\mathrm{L}^+(H)$ the subset of H consisting of all nonnegative symmetric linear operators. Finally, $\mathfrak{B}(H)$ will signify the σ-algebra of all Borel subsets of H. A bounded linear operator $R \in \mathrm{L}(H)$ is said to be of *trace class* if there exist two sequences $(a_k)_{k=1}^\infty$ and $(b_k)_{k=1}^\infty$ in H such that

$$Rh = \sum_{k=1}^\infty (h, a_k)_H b_k, \quad h \in H, \tag{4.1}$$

and

$$\sum_{k=1}^\infty \|a_k\|_H \, \|b_k\|_H < \infty.$$

The set of all elements of $\mathrm{L}(H)$ that are trace class will be denoted by $\mathrm{L}_1(H)$. We note that if $R \in \mathrm{L}_1(H)$, then R is a compact linear operator on H. The set $\mathrm{L}_1(H)$, endowed with the usual operations of addition and scalar multiplication, is a Banach space with the norm

$$\|R\|_{\mathrm{L}_1(H)} := \inf\left\{\sum_{k=1}^\infty \|a_k\|_H \, \|b_k\|_H \,:\, Rh = \sum_{k=1}^\infty (h, a_k)_H b_k, \quad h \in H, \quad (a_k)_{k=1}^\infty, \, (b_k)_{k=1}^\infty \subset H\right\}.$$

Assuming that $R \in \mathrm{L}_1(H)$, the *trace* $\mathrm{Tr}\, R$ of R is defined by the formula

$$\mathrm{Tr}\, R = \sum_{k=1}^\infty (Re_k, e_k)_H,$$

where at this stage $(e_k)_{k=1}^\infty$ is any complete orthonormal basis in H. In particular if $R \in \mathrm{L}_1(H)$ is expressed by (4.1), then

$$\mathrm{Tr}\, R = \sum_{k=1}^\infty (a_k, b_k)_H.$$

The definition of *trace* being independent of the choice of the basis, we have that $|\mathrm{Tr} R| \leq \|R\|_{\mathrm{L}_1(H)}$. The next proposition (cf. Pietsch [22], Theorem 8.3.3) sharpens these statements further.

Proposition 4.1. *Suppose that R is a compact selfadjoint linear operator on a separable Hilbert space H, and that $(\lambda_k)_{k=1}^{\infty}$ are its eigenvalues (repeated according to their multiplicity).*

Then, $R \in \mathrm{L}_1(H)$ if, and only if, $\sum_{k=1}^{\infty} |\lambda_k| < +\infty$. Furthermore, $\|R\|_{\mathrm{L}_1(H)} = \sum_{k=1}^{\infty} |\lambda_k|$ and $\mathrm{Tr}\, R = \sum_{k=1}^{\infty} \lambda_k$. In particular if $\lambda_k \geq 0$ for all $k \in \mathbb{N}$, then $\mathrm{Tr}\, R = \|R\|_{\mathrm{L}_1(H)}$.

Let $\mathbb{R}^\infty$ denote linear space of all sequences $x = (x_k)_{k=1}^{\infty}$ of real numbers, equipped with the metric

$$\mathfrak{d}(x,y) := \sum_{k=1}^{\infty} 2^{-k} \frac{|x_k - y_k|}{1 + |x_k - y_k|}, \quad x, y \in \mathbb{R}^\infty.$$

Let further $\ell^2(\mathbb{N})$ denote the Hilbert space of all sequences $x = (x_k)_{k=1}^{\infty} \in \mathbb{R}^\infty$ such that

$$\|x\|_{\ell^2(\mathbb{N})} := \left(\sum_{k=1}^{\infty} |x_k|^2 \right)^{1/2} < +\infty;$$

the inner product on $\ell^2(\mathbb{N})$ is defined by $(x,y)_{\ell^2(\mathbb{N})} := \sum_{k=1}^{\infty} x_k y_k$, for $x, y \in \ell^2(\mathbb{N})$.

Let us further consider, for $a \in \mathbb{R}$ and for $\lambda > 0$, the Gaussian measure on $\mathbb{R}$ with mean a and standard deviation λ defined by

$$N_{a,\lambda}(\mathrm{d}x) := \frac{1}{\sqrt{2\pi\lambda}}\, \mathrm{e}^{-\frac{(x-a)^2}{2\lambda}}\, \mathrm{d}x.$$

The following result is stated as Theorem 1.2.1 in [12].

Theorem 4.2. *Suppose that $a \in H$ and $Q \in \mathrm{L}_1^+(H)$. Then, there exists a unique probability measure μ on $(H, \mathfrak{B}(H))$ such that*

$$\int_H \mathrm{e}^{\imath (h,x)_H} \mu(\mathrm{d}x) = \mathrm{e}^{\imath (a,h)_H} \mathrm{e}^{-\frac{1}{2}(Qh,h)_H}, \quad h \in H.$$

Moreover, μ is the restriction to H (identified with the Hilbert space $\ell^2(\mathbb{N})$) of the product measure

$$\bigotimes_{k=1}^{\infty} \mu_k = \bigotimes_{k=1}^{\infty} N_{a_k,\lambda_k},$$

defined on $(\mathbb{R}^\infty, \mathfrak{B}(\mathbb{R}^\infty))$.

We shall write $\mu = N_{a,Q}$, and we call a the *mean* and the trace class operator Q the *covariance operator* of μ. The measure μ will be referred to as a *Gaussian measure on H* with mean a and covariance operator Q. If the law of a random variable is a Gaussian measure, then the random variable is said to be *Gaussian*. In particular, Theorem 4.2 implies that a random variable X with values in H is Gaussian if, and only if, for any $h \in H$ the real-valued random variable $(h, X)_H$ is Gaussian.

For $\mu = N_{a,Q}$, we denote by $\mathrm{L}^2(H,\mu)$ the Hilbert space of all square-integrable (equivalence classes of) functions from H into $\mathbb{R}$ with inner product

$$(u,v)_{\mathrm{L}^2(H,\mu)} = \int_H u(x)\, v(x)\, \mu(\mathrm{d}x), \quad u, v \in \mathrm{L}^2(H,\mu)$$

and norm

$$\|u\|_{\mathrm{L}^2(H,\mu)} = (u,u)^{\frac{1}{2}}_{\mathrm{L}^2(H,\mu)}, \quad u \in \mathrm{L}^2(H,\mu).$$

Analogously, we shall denote by $\mathrm{L}^2(H,\mu;H)$ the Hilbert space of all square-integrable (equivalence classes of) functions from H into H with inner product

$$(u,v)_{\mathrm{L}^2(H,\mu;H)} = \int_H (u(x), v(x))_H\, \mu(\mathrm{d}x), \quad u, v \in \mathrm{L}^2(H,\mu;H)$$

and norm

$$\|u\|_{\mathrm{L}^2(H,\mu;H)} = (u,u)^{\frac{1}{2}}_{\mathrm{L}^2(H,\mu;H)} \quad u \in \mathrm{L}^2(H,\mu;H).$$

Throughout the rest of the paper, $\mu = N_Q := N_{0,Q}$, where $Q \in \mathrm{L}_1^+(H)$. Moreover, we shall assume that $\mathrm{Ker}(Q) = \{0\}$. We shall also suppose that there exists a complete orthonormal system $(e_k)_{k=1}^{\infty}$ in H and a sequence $(\lambda_k)_{k=1}^{\infty}$ of positive real numbers, the eigenvalues of Q, such that

$Qe_k = \lambda_k e_k$, $k \in \mathbb{N}$. The eigenvalues λ_k are assumed to be enumerated in decreasing order (and repeated according to their multiplicity), and accumulating only at zero. For $x \in H$, we shall then write $x_k = (x, e_k)_H$, $k \in \mathbb{N}$.

The subspace $Q^{1/2}(H) = \{Q^{1/2}h : h \in H\}$ is called the *reproducing kernel* of the Gaussian measure $N_Q := N_{0,Q}$. If $\mathrm{Ker}(Q) = \{0\}$ as has been assumed, then $Q^{1/2}(H)$ is dense in H. In fact, if $x_0 \in H$ is such that $(Q^{1/2}h, x_0)_H = 0$ for all $h \in H$, then $Q^{1/2}x_0 = 0$, and therefore $Qx_0 = 0$, which implies that $x_0 = 0$.

For $Q \in \mathrm{L}_1^+(H)$ such that $\mathrm{Ker}(Q) = \{0\}$, we introduce the isomorphism W from H into $\mathrm{L}^2(H, N_Q)$ as follows: for $f \in Q^{1/2}(H)$ let $W_f \in \mathrm{L}^2(H; N_Q)$ be defined by

$$W_f(x) = (Q^{-1/2}f, x)_H, \quad x \in H.$$

Let $\mu = N_Q$. We define Hermite polynomials in $\mathrm{L}^2(H, \mu)$. Let us consider to this end the set Γ of all mappings $\gamma : n \in \mathbb{N} \to \gamma_n \in \{0\} \cup \mathbb{N}$, such that $|\gamma| := \sum_{k=1}^{\infty} \gamma_k < +\infty$. Clearly $\gamma \in \Gamma$ if, and only if, $\gamma_n = 0$ for all, except possibly finitely many, $n \in \mathbb{N}$. For any $\gamma \in \Gamma$ we define the *Hermite polynomial*

$$H_\gamma(x) := \prod_{k=1}^{\infty} H_{\gamma_k}(W_{e_k}(x)), \quad x \in H, \tag{4.2}$$

where the factors in the product on the right-hand side are defined by

$$H_n(\xi) = \frac{(-1)^n}{\sqrt{n!}}\,\mathrm{e}^{\frac{\xi^2}{2}}\,\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\left(\mathrm{e}^{-\frac{\xi^2}{2}}\right), \quad \xi \in \mathbb{R}, \quad n \in \{0\} \cup \mathbb{N}; \tag{4.3}$$

H_n is the classical univariate Hermite polynomial of degree n on $\mathbb{R}$. Note that $H_0 \equiv 1$, so that for each $\gamma \in \Gamma$, the countable product $H_\gamma(x)$ contains only finitely many nontrivial factors and is, therefore, well-defined. Moreover, with the Gaussian measure $\mu = N_Q$ being a countable product measure, we have that

$$\forall \gamma, \gamma' \in \Gamma : \quad (H_\gamma, H_{\gamma'})_{\mathrm{L}^2(H,\mu)} = \delta_{\gamma,\gamma'}. \tag{4.4}$$

We shall denote by $E(H)$ the linear space spanned by all exponential functions, that is all functions $\varphi : x \in H \mapsto \varphi(x) \in \mathbb{R}$ of the form

$$\varphi(x) = \mathrm{e}^{(h,x)_H}, \quad h \in H.$$

It follows from Proposition 1.2.5 in Da Prato and Zabczyk [12] that $E(H)$ is dense in $\mathrm{L}^2(H, \mu)$. On account of the separability of H, $\mathrm{L}^2(H, \mu)$ is separable.

For any $k \in \mathbb{N}$ we consider the partial derivative in the direction e_k (with e_k as above), defined by

$$\mathrm{D}_k\varphi(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(\varphi(x + \varepsilon e_k) - \varphi(x)\right), \quad x \in H,\ \varphi \in E(H).$$

Thus, if $\varphi(x) = \mathrm{e}^{(h,x)_H}$ with $h \in H$, then clearly

$$\mathrm{D}_k\varphi(x) = \mathrm{e}^{(h,x)_H}\, h_k, \quad \text{where } h_k = (h, e_k)_H.$$

Let Λ_0 denote the linear span of $\{H_\gamma \otimes e_k : \gamma \in \Gamma, k \in \mathbb{N}\}$, with H_γ and e_k as above and $\mu = N_Q$, and define the linear operator

$$\mathrm{D} : E(H) \subset \mathrm{L}^2(H, \mu) \to \mathrm{L}^2(H, \mu; H) \quad \text{by} \quad \mathrm{D}\varphi(x) := \sum_{k=1}^{\infty} \mathrm{D}_k\varphi(x)\, e_k, \quad x \in H.$$

Thanks to Proposition 9.2.2 in Da Prato and Zabczyk [12], D_k is a closable linear operator for all $k \in \mathbb{N}$. If φ belongs to the domain of the closure of D_k, which we shall still denote by D_k, we shall say that $\mathrm{D}_k\varphi$ belongs to $\mathrm{L}^2(H, \mu)$.

Analogously, by Proposition 9.2.4 in Da Prato and Zabczyk [12], D is a closable linear operator. If φ belongs to the domain of the closure of D, which we shall still denote by D, we shall say that $\mathrm{D}\varphi$ belongs to $\mathrm{L}^2(H, \mu; H)$.

For $\mu = N_Q$, let us denote by $\mathrm{W}^{1,2}(H,\mu)$ the linear space of all functions $\varphi \in \mathrm{L}^2(H,\mu)$ such that $\mathrm{D}\varphi \in \mathrm{L}^2(H,\mu;H)$, equipped with the inner product

$$(\varphi,\psi)_{\mathrm{W}^{1,2}(H,\mu)} := (\varphi,\psi)_{\mathrm{L}^2(H,\mu)} + \int_H (\mathrm{D}\varphi(x), \mathrm{D}\psi(x))_H \, \mu(\mathrm{d}x)$$

and norm $\|\varphi\|_{\mathrm{W}^{1,2}(H,\mu)} = (\varphi,\varphi)^{\frac{1}{2}}_{\mathrm{W}^{1,2}(H,\mu)}$. Then, the *Sobolev space* $\mathrm{W}^{1,2}(H,\mu)$ is complete, and is therefore a separable Hilbert space. In particular, for any $\varphi \in \mathrm{L}^2(H,\mu)$, we have that

$$\varphi = \sum_{\gamma\in\Gamma} \varphi_\gamma H_\gamma, \quad \text{where } \varphi_\gamma := (\varphi, H_\gamma)_{\mathrm{L}^2(H,\mu)}.$$

The next theorem, proved independently by Da Prato, Malliavin and Nualart [11] and Peszat [21] provides an analogous characterization of functions belonging to $\mathrm{W}^{1,2}(H,\mu)$ in terms of the complete orthonormal basis $(H_\gamma)_{\gamma\in\Gamma}$.

Theorem 4.3. *A function $\varphi \in \mathrm{L}^2(H,\mu)$ belongs to $\mathrm{W}^{1,2}(H,\mu)$ if, and only if,*

$$\sum_{\gamma\in\Gamma} \langle \gamma, \lambda^{-1}\rangle |\varphi_\gamma|^2 < +\infty, \tag{4.5}$$

where $\varphi_\gamma := (\varphi, H_\gamma)_{\mathrm{L}^2(H,\mu)}$, $\langle\gamma,\lambda^{-1}\rangle := \sum_{k=1}^\infty \gamma_k \lambda_k^{-1}$, and $(\lambda_k)_{k=1}^\infty$ is the sequence of (positive) eigenvalues (repeated according to their multiplicity) of the covariance operator $Q \in \mathrm{L}_1^+(H)$, $\mathrm{Ker}(Q) = \{0\}$. Moreover, if (4.5) holds, then

$$\|\varphi\|^2_{\mathrm{W}^{1,2}(H,\mu)} = \|\varphi\|^2_{\mathrm{L}^2(H,\mu)} + \sum_{\gamma\in\Gamma} \langle\gamma,\lambda^{-1}\rangle |\varphi_\gamma|^2.$$

Identifying $\mathrm{L}^2(H,\mu)$ with its own dual $\mathrm{L}^2(H,\mu)^*$, we obtain

$$\varphi \in (\mathrm{W}^{1,2}(H,\mu))^* \iff \sum_{\gamma\in\Gamma} \langle\gamma,\lambda\rangle |\varphi_\gamma|^2 < +\infty.$$

Furthermore, the embedding of $\mathrm{W}^{1,2}(H,\mu)$ into $\mathrm{L}^2(H,\mu)$ is compact.

For a proof of these statements we refer, for example, to [12, Theorem 9.2.12]. As $E(H)$ is dense in both $\mathrm{L}^2(H,\mu)$ and $\mathrm{W}^{1,2}(H,\mu)$, the embedding of $\mathrm{W}^{1,2}(H,\mu)$ into $\mathrm{L}^2(H,\mu)$ is also dense.

5. Fokker–Planck equation in countably many dimensions

For $k \in \mathbb{N}$, let us denote by D_k the set $\mathbb{R}^d$ equipped with the Gaussian measure

$$\mu_k(\mathrm{d}q_k) = N_{a_k,\Sigma_k}(\mathrm{d}q_k) = \frac{1}{(2\pi)^{d/2}[\det(\Sigma_k)]^{1/2}} \exp\left(-\tfrac{1}{2}(q_k - a_k)^\top \Sigma_k^{-1}(q_k - a_k)\right) \mathrm{d}q_k$$

with *mean* $a_k \in \mathbb{R}^d$ and positive definite *covariance matrix* $\Sigma_k \in \mathbb{R}^{d\times d}$. We shall assume henceforth that $a_k = 0$ for all $k \in \mathbb{N}$ and that the covariance operator Q, represented by the bi-infinite block-diagonal matrix

$$\Sigma = \mathrm{diag}(\Sigma_1, \Sigma_2, \dots),$$

with $d \times d$ diagonal blocks Σ_k, $k = 1, 2, \dots$, is trace class. We define

$$D := \bigtimes_{k=1}^\infty D_k$$

so that

$$q := (q_1^\top, q_2^\top, \dots)^\top \in D, \quad q_k \in D_k, \quad k = 1, 2, \dots.$$

We equip the domain D with the product measure

$$\mu := \bigotimes_{k=1}^\infty \mu_k = \bigotimes_{k=1}^\infty N_{0,\Sigma_k}.$$

Let $\mathbf{A} = (A_{ij})_{i,j=1}^{\infty} \in \mathbb{R}^{\infty\times\infty}$ be a symmetric infinite matrix, i.e., $A_{ij} = A_{ji}$ for all $i,j \in \mathbb{N}$. Suppose further that there exists a real number $\gamma_0 > 0$ such that

$$\sum_{i,j=1}^{\infty} A_{ij}\,\xi_i\xi_j \geq \gamma_0\|\xi\|^2_{\ell^2(\mathbb{N})} \quad \text{for all} \quad \xi = (\xi_i)_{i=1}^{\infty} \in \ell^2(\mathbb{N}) \tag{5.1}$$

and a real number $\gamma_1 > 0$ such that

$$\left|\sum_{i,j=1}^{\infty} A_{ij}\,\xi_i\eta_j\right| \leq \gamma_1\|\xi\|_{\ell^2(\mathbb{N})}\,\|\eta\|_{\ell^2(\mathbb{N})} \quad \text{for all} \quad \xi = (\xi_i)_{i=1}^{\infty} \in \ell^2(\mathbb{N}) \text{ and } \eta = (\eta_i)_{i=1}^{\infty} \in \ell^2(\mathbb{N}). \tag{5.2}$$

Example 5.1. *As an example of an infinite matrix* $\mathbf{A}$ *that satisfies* (5.1) *and* (5.2)*, we mention*

$$\mathbf{A}[\epsilon] = \mathrm{tridiag}\{(\epsilon_j, 1, \epsilon_j) : j = 1, 2, \dots\}, \tag{5.3}$$

where the sequence $\epsilon = (\epsilon_j)_{j=1}^{\infty}$ *is assumed to satisfy* $\|\epsilon\|_{\ell^\infty(\mathbb{N})} < 1/2$*. Then, the matrix* $\mathbf{A}[\epsilon]$ *in* (5.3) *satisfies* (5.1) *with* $\gamma_0 = 1 - 2\|\epsilon\|_{\ell^\infty(\mathbb{N})}$ *and* (5.2) *with* $\gamma_1 = 1 + 2\|\epsilon\|_{\ell^\infty(\mathbb{N})}$*.*

Example 5.2. *More generally, we may consider* $\mathbf{A}$ *of block-diagonal form* $\mathbf{A} = \mathrm{diag}\{\mathbf{A}_{11}, \mathbf{A}_{22}\}$*, where* $\mathbf{A}_{11} \in \mathbb{R}^{d\times d}$ *is symmetric, positive definite, and* $\mathbf{A}_{22}$ *is as in Example 5.1. Then,* (5.1) *and* (5.2) *hold with constants* $0 < \gamma_0 \leq \gamma_1 < \infty$ *depending on the spectrum of* $\mathbf{A}_{11}$ *and on* d*, however.*

In order to state the space-time variational formulation of the Fokker–Planck equation using the notation introduced in Section 2, we select $\mathcal{H} = \mathrm{L}^2(D;\mu)$ and $\mathcal{V} = \mathrm{W}^{1,2}(D,\mu)$, and we write $\mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$. Guided by the abstract framework in Sections 2.3, 2.4, we consider the space

$$\mathcal{X} = \mathrm{L}^2(0,T;\mathcal{V}) \cap \mathrm{H}^1(0,T;\mathcal{V}^*) \tag{5.4}$$

and its subspaces $\mathcal{X}_{0,\{0\}}$, $\mathcal{X}_{0,\{T\}}$ as in (2.17), equipped with the norm of $\mathcal{X}$, which are closed due to (2.9). With these spaces, the *space-time adjoint weak formulation of the infinite-dimensional Fokker–Planck equation* reads as follows: given $\widehat{\psi}_0 \in \mathcal{H}$ and $g \in \mathcal{X}^*_{0,\{T\}}$, find

$$\widehat{\psi} \in \mathcal{Y}: \quad \mathfrak{B}^*(\widehat{\psi}, \widehat{\varphi}) = \ell^*(\widehat{\varphi}) \quad \forall \widehat{\varphi} \in \mathcal{X}_{0,\{T\}}, \tag{5.5}$$

where the bilinear form $\mathfrak{B}^*(\cdot,\cdot) : \mathcal{Y} \times \mathcal{X}_{0,\{T\}} \to \mathbb{R}$ is defined by

$$\mathfrak{B}^*(\widehat{\psi}, \widehat{\varphi}) := \int_0^T \big(-\, {}_{\mathcal{V}^*}\langle \widehat{\varphi}', \widehat{\psi}\rangle_{\mathcal{V}} + \mathfrak{a}(\widehat{\psi}, \widehat{\varphi})\big)\,\mathrm{d}t, \tag{5.6}$$

with

$$\mathfrak{a}(\widehat{\psi}, \widehat{\varphi}) := \sum_{i,j=1}^{\infty} A_{ij}(\nabla_{q_i}\widehat{\psi}, \nabla_{q_j}\widehat{\varphi})_{[\mathrm{L}^2(D,\mu)]^d},$$

and the linear functional ℓ^* is defined by

$$\ell^*(\widehat{\varphi}) := (\widehat{\psi}_0, \widehat{\varphi}(0))_{\mathrm{L}^2(D,\mu)}. \tag{5.7}$$

We note that $\widetilde{\varphi} := \mathrm{e}^t\widehat{\varphi} \in \mathcal{X}_{0,\{T\}}$ if, and only if, $\widehat{\varphi} \in \mathcal{X}_{0,\{T\}}$; analogously, $\widetilde{\psi} := \mathrm{e}^{-t}\widehat{\psi} \in \mathcal{Y}$ if, and only if, $\widehat{\psi} \in \mathcal{Y}$. Thus, by selecting $\widehat{\varphi} = \mathrm{e}^{-t}\widetilde{\varphi}$ as test function in (5.5), we deduce that $\widehat{\psi} \in \mathcal{Y}$ is a solution to (5.5) if, and only if, $\widetilde{\psi} := \mathrm{e}^{-t}\widehat{\psi} \in \mathcal{Y}$ solves

$$\widetilde{\psi} \in \mathcal{Y}: \quad \mathfrak{B}^*(\widetilde{\psi}, \widetilde{\varphi}) = \ell^*(\widetilde{\varphi}) \quad \forall \widetilde{\varphi} \in \mathcal{X}_{0,\{T\}},$$

where the linear functional ℓ^* is defined as in (5.7) (note that $\widetilde{\psi}(0) = \widehat{\psi}(0)$), and the bilinear form $\mathfrak{B}^*(\cdot,\cdot) : \mathcal{Y} \times \mathcal{X}_{0,\{T\}} \to \mathbb{R}$ is defined by (5.6), except that now

$$\mathfrak{a}(\widetilde{\psi}, \widetilde{\varphi}) := \sum_{i,j=1}^{\infty} A_{ij}(\nabla_{q_i}\widetilde{\psi}, \nabla_{q_j}\widetilde{\varphi})_{[\mathrm{L}^2(D,\mu)]^d} + (\widetilde{\psi}, \widetilde{\varphi})_{\mathrm{L}^2(D,\mu)}.$$

The solution $\widehat{\psi} \in \mathcal{Y}$ to (5.5) is then directly recovered from $\widetilde{\psi}$ by setting $\widehat{\psi} := \mathrm{e}^t\widetilde{\psi}$. We shall therefore assume (without loss of generality) in what follows that, whenever the problem (5.5) is referred to, it is the bilinear form $\mathfrak{a}$ defined by

$$\mathfrak{a}(\widehat{\psi}, \widehat{\varphi}) := \sum_{i,j=1}^{\infty} A_{ij}(\nabla_{q_i}\widehat{\psi}, \nabla_{q_j}\widehat{\varphi})_{[\mathrm{L}^2(D,\mu)]^d} + (\widehat{\psi}, \widehat{\varphi})_{\mathrm{L}^2(D,\mu)} \tag{5.8}$$

that is used in the definition of $\mathfrak{B}^*$, rather than the bilinear form $\mathfrak{a}$ defined above following (5.6).

By adopting this simple convention, the well-posedness of the infinite-dimensional Fokker–Planck equation is now an immediate consequence of Theorem 2.5; in particular, the following result holds.

Theorem 5.3. *For the bilinear form $\mathfrak{B}^*(\cdot,\cdot)$ in (5.6), where $\mathfrak{a}(\cdot,\cdot)$ is defined by* (5.8)*, there exists a positive constant C such that*

$$\inf_{\widehat{\psi}\in\mathcal{Y}\setminus\{0\}} \sup_{\widehat{\varphi}\in\mathcal{X}_{0,\{T\}}\setminus\{0\}} \frac{\mathfrak{B}^*(\widehat{\psi}, \widehat{\varphi})}{\|\widehat{\psi}\|_{\mathcal{Y}}\|\widehat{\varphi}\|_{\mathcal{X}_{0,\{T\}}}} \geq C \tag{5.9}$$

and

$$\forall\widehat{\varphi} \in \mathcal{X}_{0,\{T\}} \setminus \{0\} : \qquad \sup_{\widehat{\psi}\in\mathcal{Y}} \mathfrak{B}^*(\widehat{\psi}, \widehat{\varphi}) > 0. \tag{5.10}$$

Furthermore, for each $\widehat{\psi}_0 \in \mathcal{H}$ and each $g \in \mathcal{X}^*_{0,\{T\}}$, there exists a unique $\widehat{\psi} \in \mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$ that satisfies (5.5); and the linear operator $B^* \in \mathcal{L}(\mathcal{Y}, \mathcal{X}^*_{0,\{T\}})$ induced by $\mathfrak{B}^*(\cdot,\cdot)$ is an isomorphism.

Following [24], we shall next develop a class of adaptive Galerkin discretization algorithms for (5.5), along the lines of adaptive wavelet discretizations of boundedly invertible operator equations considered in [9, 10]. These algorithms exhibit, in particular, *"stability by adaptivity"*, i.e., their stability follows directly from the stability (5.9), (5.10) of the continuous, infinite-dimensional problem through suitable Riesz bases of the spaces $\mathcal{Y}$ and $\mathcal{X}_{0,\{T\}}$, which we shall construct. Notably, in the present context, these algorithms are *dimensionally robust by design*: as we shall prove, they deliver a sequence of approximate solutions with finitely supported coefficient vectors, i.e., with only finitely many variables q_k being 'active' in each iteration. We will establish certain optimality properties for these finitely supported Galerkin solutions, with all constants in the error and complexity bounds being absolute, i.e., independent of the number of active variables. We first develop the necessary concepts in an abstract setting, before applying them to the Fokker–Planck equation (5.5) in space-time variational form.

6. Well-posed operator equations as infinite matrix problems

Let us denote for a moment by $\mathcal{X}, \mathcal{Y}$ two generic separable Hilbert spaces over $\mathbb{R}$, and let us assume that we have available a Riesz basis $\Psi^{\mathcal{X}} = \{\psi^{\mathcal{X}}_\lambda : \lambda \in \nabla_{\mathcal{X}}\}$ for $\mathcal{X}$, meaning that the *synthesis operator*

$$s_{\Psi^{\mathcal{X}}} : \ell^2(\nabla_{\mathcal{X}}) \to \mathcal{X} : \mathbf{c} \mapsto \mathbf{c}^\top\Psi^{\mathcal{X}} := \sum_{\lambda\in\nabla_{\mathcal{X}}} c_\lambda \psi^{\mathcal{X}}_\lambda$$

is boundedly invertible. By identifying $\ell^2(\nabla_{\mathcal{X}})$ with its dual, the adjoint of $s_{\Psi^{\mathcal{X}}}$, known as the *analysis operator*, is

$$s^*_{\Psi^{\mathcal{X}}} : \mathcal{X}^* \to \ell^2(\nabla_{\mathcal{X}}) : g \mapsto [g(\psi^{\mathcal{X}}_\lambda)]_{\lambda\in\nabla_{\mathcal{X}}}.$$

Similarly, let $\Psi^{\mathcal{Y}} = \{\psi^{\mathcal{Y}}_\lambda : \lambda \in \nabla_{\mathcal{Y}}\}$ be a Riesz basis for $\mathcal{Y}$, with synthesis operator $s_{\Psi^{\mathcal{Y}}}$ and its adjoint $s^*_{\Psi^{\mathcal{Y}}}$. Here, $\nabla_{\mathcal{X}}$ and $\nabla_{\mathcal{Y}}$ are countable sets of (multi-)indices λ.

Now let $B \in \mathcal{L}(\mathcal{X}, \mathcal{Y}^*)$ be boundedly invertible; then, also its adjoint $B^* \in \mathcal{L}(\mathcal{Y}, \mathcal{X}^*)$ is boundedly invertible and, for every $f \in \mathcal{Y}^*$, $f^* \in \mathcal{X}^*$ the *operator equations*

$$Bu = f, \qquad B^*u^* = f^*$$

admit unique solutions $u \in \mathcal{X}$, $u^* \in \mathcal{Y}$. Writing $u = s_{\Psi^{\mathcal{X}}}\mathbf{u}$ and $u^* = s_{\Psi^{\mathcal{Y}}}\mathbf{u}^*$, these operator equations are *equivalent* to *infinite matrix-vector problems*

$$\mathbf{B}\mathbf{u} = \mathbf{f}, \quad \mathbf{B}^*\mathbf{u}^* = \mathbf{f}^*, \tag{6.1}$$

where the "load vectors" $\mathbf{f} = s^*_{\Psi^{\mathcal{Y}}} f = [f(\psi^{\mathcal{Y}}_\lambda)]_{\lambda\in\nabla_{\mathcal{Y}}} \in \ell^2(\nabla_{\mathcal{Y}})$, $\mathbf{f}^* = s^*_{\Psi^{\mathcal{X}}} f = [f(\psi^{\mathcal{X}}_\lambda)]_{\lambda\in\nabla_{\mathcal{X}}} \in \ell^2(\nabla_{\mathcal{X}})$; and the *system matrix* $\mathbf{B}$ given by

$$\mathbf{B} = s^*_{\Psi^{\mathcal{Y}}} B s_{\Psi^{\mathcal{X}}} = [(B\psi^{\mathcal{X}}_\mu)(\psi^{\mathcal{Y}}_\lambda)]_{\lambda\in\nabla_{\mathcal{Y}},\mu\in\nabla_{\mathcal{X}}} \in \mathcal{L}(\ell^2(\nabla_{\mathcal{X}}), \ell^2(\nabla_{\mathcal{Y}}))$$

and with $\mathbf{B}^*$ defined analogously, are boundedly invertible. We consider the associated bilinear forms

$$\mathfrak{B}(\cdot,\cdot) \;:\; \mathcal{X}\times\mathcal{Y}\to\mathbb{R} : (w,v)\mapsto (Bw)(v), \quad \mathfrak{B}^*(\cdot,\cdot) \;:\; \mathcal{Y}\times\mathcal{X}\to\mathbb{R} : (v,w)\mapsto (B^*v)(w),$$

and introduce the notations

$$\mathbf{B} = \mathfrak{B}(\Psi^{\mathcal{X}}, \Psi^{\mathcal{Y}}), \quad \mathbf{f} = f(\Psi^{\mathcal{Y}}), \quad \mathbf{B}^* = \mathfrak{B}^*(\Psi^{\mathcal{Y}}, \Psi^{\mathcal{X}}), \quad \mathbf{f}^* = f^*(\Psi^{\mathcal{X}}).$$

With the *Riesz constants*

$$\Lambda^{\mathcal{X}}_{\Psi^{\mathcal{X}}} := \|s_{\Psi^{\mathcal{X}}}\|_{\ell^2(\nabla_{\mathcal{X}})\to\mathcal{X}} = \sup_{0\neq\mathbf{c}\in\ell^2(\nabla_{\mathcal{X}})} \frac{\|\mathbf{c}^\top\Psi^{\mathcal{X}}\|_{\mathcal{X}}}{\|\mathbf{c}\|_{\ell^2(\nabla_{\mathcal{X}})}},$$

$$\lambda^{\mathcal{X}}_{\Psi^{\mathcal{X}}} := \|s^{-1}_{\Psi^{\mathcal{X}}}\|^{-1}_{\mathcal{X}\to\ell^2(\nabla_{\mathcal{X}})} = \inf_{0\neq\mathbf{c}\in\ell^2(\nabla_{\mathcal{X}})} \frac{\|\mathbf{c}^\top\Psi^{\mathcal{X}}\|_{\mathcal{X}}}{\|\mathbf{c}\|_{\ell^2(\nabla_{\mathcal{X}})}},$$

and $\Lambda^{\mathcal{Y}}_{\Psi^{\mathcal{Y}}}$ and $\lambda^{\mathcal{Y}}_{\Psi^{\mathcal{Y}}}$ defined analogously, we obviously have that

$$\|\mathbf{B}\|_{\ell^2(\nabla_{\mathcal{X}})\to\ell^2(\nabla_{\mathcal{Y}})} \leq \|B\|_{\mathcal{X}\to\mathcal{Y}^*}\Lambda^{\mathcal{X}}_{\Psi^{\mathcal{X}}}\Lambda^{\mathcal{Y}}_{\Psi^{\mathcal{Y}}}, \tag{6.2}$$

$$\|\mathbf{B}^{-1}\|_{\ell^2(\nabla_{\mathcal{Y}})\to\ell^2(\nabla_{\mathcal{X}})} \leq \frac{\|B^{-1}\|_{\mathcal{Y}^*\to\mathcal{X}}}{\lambda^{\mathcal{X}}_{\Psi^{\mathcal{X}}}\lambda^{\mathcal{Y}}_{\Psi^{\mathcal{Y}}}}. \tag{6.3}$$

7. Best N-term approximations and approximation classes

We say that $\mathbf{u}_N \in \ell^2(\nabla_{\mathcal{X}})$ is a *best N-term approximation* of $\mathbf{u}\in\ell^2(\nabla_{\mathcal{X}})$, if it is the best possible approximation to u in the norm of $\ell^2(\nabla_{\mathcal{X}})$, given a budget of N coefficients. Determining the best N-term approximation $\mathbf{u}_N$ of $\mathbf{u}$ requires searching the infinite vector $\mathbf{u}$, which is not feasible, i.e., actually locating $\mathbf{u}_N$ may not be practically feasible. Nevertheless, rates of convergence afforded by best N-term approximations serve as a *benchmark* for concrete numerical schemes. To this end, we collect all $\mathbf{u}\in\ell^2(\nabla_{\mathcal{X}})$ admitting a best N-term approximation converging with rate $s>0$ in the approximation class $\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{X}})) := \{\mathbf{v}\in\ell^2(\nabla_{\mathcal{X}}) : \|\mathbf{v}\|_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{X}}))} < \infty\}$, where

$$\|\mathbf{v}\|_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{X}}))} := \sup_{\varepsilon>0} \varepsilon \times [\min\{N\in\mathbb{N}_0 : \|\mathbf{v}-\mathbf{v}_N\|_{\ell^2(\nabla_{\mathcal{X}})} \leq \varepsilon\}]^s,$$

which consists of all $\mathbf{v}\in\ell^2(\nabla_{\mathcal{X}})$ whose best N-term approximations converge to $\mathbf{v}$ with rate s.

Generally, best N-term approximations cannot be realized in practice, in particular not in situations where the vector $\mathbf{u}$ to be approximated is only defined implicitly as the solution of an infinite matrix-vector problem, such as (6.1). We will now design, following [24], a space-time adaptive Galerkin discretization method for the infinite-dimensional Fokker–Planck equation (5.5) that produces a sequence of approximations to $\mathbf{u}$, which, whenever $\mathbf{u}\in\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{X}}))$ for some $s>0$, converge to $\mathbf{u}$ *with this rate* $s>0$ in computational complexity that is linear with respect to N, the cardinality of the support of the "active" coefficients in the computed, finitely supported approximation to $\mathbf{u}$.

8. Adaptive Galerkin methods

Let $s>0$ be such that $\mathbf{u}\in\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{X}}))$. In [9] and [10], adaptive wavelet Galerkin methods for solving elliptic operator equations (6.1) were introduced; the methods considered in both papers are iterative methods. To be able to bound their complexity, one needs a suitable bound on the complexity of an approximate matrix-vector product in terms of the prescribed tolerance. We formalize this idea through the notion of s^*-admissibility.

Definition 8.1. *The infinite matrices* $\mathbf{B}\in\mathcal{L}(\ell^2(\nabla_{\mathcal{X}}),\ell^2(\nabla_{\mathcal{Y}}))$, $\mathbf{B}^*\in\mathcal{L}(\ell^2(\nabla_{\mathcal{Y}}),\ell^2(\nabla_{\mathcal{X}}))$ *are* s^*-admissible if there exist routines

$$\mathbf{APPLY}_{\mathbf{B}}[\mathbf{w},\varepsilon]\to\mathbf{z}, \quad \mathbf{APPLY}_{\mathbf{B}^*}[\tilde{\mathbf{w}},\varepsilon]\to\tilde{\mathbf{z}}$$

that yield, for any prescribed tolerance $\varepsilon > 0$ and any finitely supported $\mathbf{w} \in \ell^2(\nabla_{\mathcal{X}})$ and $\tilde{\mathbf{w}} \in \ell^2(\nabla_{\mathcal{Y}})$, finitely supported vectors $\mathbf{z} \in \ell^2(\nabla_{\mathcal{Y}})$ and $\tilde{\mathbf{z}} \in \ell^2(\nabla_{\mathcal{X}})$ such that

$$\|\mathbf{B}\mathbf{w} - \mathbf{z}\|_{\ell^2(\nabla_{\mathcal{Y}})} + \|\mathbf{B}^*\tilde{\mathbf{w}} - \tilde{\mathbf{z}}\|_{\ell^2(\nabla_{\mathcal{X}})} \le \varepsilon,$$

and for which, for any $\bar{s} \in (0, s^*)$, there exists an *admissibility constant* $a_{\mathbf{B},\bar{s}}$ such that

$$\#\mathrm{supp}\,\mathbf{z} \le a_{\mathbf{B},\bar{s}}\,\varepsilon^{-1/\bar{s}}\|\mathbf{w}\|^{1/\bar{s}}_{\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{X}}))},$$

and likewise for $\tilde{\mathbf{z}}$, $\tilde{\mathbf{w}}$. The number of arithmetic operations and storage locations used by the call $\mathbf{APPLY_B}[\mathbf{w}, \varepsilon]$ is bounded by some absolute multiple of

$$a_{\mathbf{B},\bar{s}}\,\varepsilon^{-1/\bar{s}}\|\mathbf{w}\|^{1/\bar{s}}_{\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{X}}))} + \#\mathrm{supp}\,\mathbf{w} + 1.$$

The design of $\mathbf{APPLY_B}[\mathbf{w}, \varepsilon]$ and $\mathbf{APPLY_{B^*}}[\tilde{\mathbf{w}}, \varepsilon]$ for the system matrices $\mathbf{B}$ and $\mathbf{B}^*$ arising from the infinite-dimensional Fokker–Planck equation (5.5) is the major technical building block in the adaptive Galerkin discretization of (5.5).

In order to approximate $\mathbf{u}$ one should be able to approximate $\mathbf{f}^*$. To this end, we shall assume in what follows the availability of the following routine.

$\mathbf{RHS_{f^*}}[\varepsilon] \to \mathbf{f}^*_\varepsilon$: *For a given $\varepsilon > 0$, the routine yields a finitely supported $\mathbf{f}^*_\varepsilon \in \ell^2(\nabla_{\mathcal{X}})$ with*

$$\|\mathbf{f}^* - \mathbf{f}^*_\varepsilon\|_{\ell^2(\nabla_{\mathcal{X}})} \le \varepsilon \quad \textit{and} \quad \#\mathrm{supp}\,\mathbf{f}^*_\varepsilon \lesssim \min\{N : \|\mathbf{f}^* - \mathbf{f}^*_N\| \le \varepsilon\},$$

with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS_{f^*}}[\varepsilon]$ being bounded by some absolute multiple of $\#\mathrm{supp}\,\mathbf{f}^*_\varepsilon + 1$.

The availability of the routines $\mathbf{APPLY_B}$ and $\mathbf{RHS_f}$ has the following implications.

Proposition 8.2. *Let $\mathbf{B}$ in (6.1) be s^*-admissible. Then, for any $\bar{s} \in (0, s^*)$, we have that*

$$\|\mathbf{B}\|_{\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{X}}))\to\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{Y}}))} \le a^{\bar{s}}_{\mathbf{B},\bar{s}}.$$

For $\mathbf{z}_\varepsilon := \mathbf{APPLY_B}[\mathbf{w}, \varepsilon]$, we have that

$$\|\mathbf{z}_\varepsilon\|_{\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{Y}}))} \le a^{\bar{s}}_{\mathbf{B},\bar{s}}\|\mathbf{w}\|_{\mathcal{A}^{\bar{s}}_\infty(\ell^2(\nabla_{\mathcal{X}}))}.$$

Analogous statements hold for $\mathbf{B}^*$ in (6.1).

A proof of Proposition 8.2 can be given along the lines of the arguments presented in [9, 10]. Using the definition of $\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))$ and the properties of $\mathbf{RHS_f}$, we have the following corollary.

Corollary 8.3. *Suppose that, in (6.1), $\mathbf{B}^*$ is s^*-admissible and $\mathbf{u} \in \mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))$ for $s < s^*$; then, for $\mathbf{f}^*_\varepsilon = \mathbf{RHS_{f^*}}[\varepsilon]$,*

$$\#\mathrm{supp}\,\mathbf{f}^*_\varepsilon \lesssim a_{\mathbf{B}^*,s}\,\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))},$$

with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS_{f^*}}[\varepsilon]$ being bounded by some absolute multiple of

$$a_{\mathbf{B}^*,s}\,\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))} + 1.$$

Remark 8.4. *Besides $\|\mathbf{f}^* - \mathbf{f}^*_\varepsilon\|_{\ell^2(\nabla_{\mathcal{Y}})} \le \varepsilon$, the complexity bounds in Corollary 8.3, with $a_{\mathbf{B}^*,s}$ signifying a constant that depends only on $\mathbf{B}^*$ and s, are essential for the use of $\mathbf{RHS_{f^*}}$ in the adaptive Galerkin methods.*

The following corollary of Proposition 8.2 from [24] can be used for example for the construction of valid **APPLY** and **RHS** routines in case the adaptive Galerkin algorithms are applied to a preconditioned system.

Corollary 8.5. *Suppose that $\mathbf{B}^* \in \mathcal{L}(\ell^2(\nabla_{\mathcal{Y}}), \ell^2(\nabla_{\mathcal{X}}))$, $\mathbf{C} \in \mathcal{L}(\ell^2(\nabla_{\mathcal{X}}), \ell^2(\nabla_{\mathcal{Z}}))$ are both s^*-admissible; then, so is $\mathbf{CB}^* \in \mathcal{L}(\ell^2(\nabla_{\mathcal{Y}}), \ell^2(\nabla_{\mathcal{Z}}))$. A valid routine $\mathbf{APPLY_{CB^*}}$ is*

$$[\mathbf{w}, \varepsilon] \mapsto \mathbf{APPLY_C}\big[\mathbf{APPLY_{B^*}}[\mathbf{w}, \varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2\big], \tag{8.1}$$

with admissibility constant $a_{\mathbf{CB}^*,\bar{s}} \lesssim a_{\mathbf{B}^*,\bar{s}}(\|\mathbf{C}\|^{1/\bar{s}} + a_{\mathbf{C},\bar{s}})$ for $\bar{s} \in (0, s^*)$.

For some $s^* > s$, let $\mathbf{C} \in \mathcal{L}(\ell^2(\nabla_\mathcal{Y}), \ell^2(\nabla_\mathcal{Z}))$ be s^*-admissible. Then, for

$$\mathbf{RHS}_{\mathbf{Cf}^*}[\varepsilon] := \mathbf{APPLY}_\mathbf{C}[\mathbf{RHS}_{\mathbf{f}^*}[\varepsilon/(2\|\mathbf{C}\|)], \varepsilon/2], \tag{8.2}$$

we have that

$$\#\mathrm{supp}\,\mathbf{RHS}_{\mathbf{Cf}^*}[\varepsilon] \lesssim a_{\mathbf{B}^*,s}(\|\mathbf{C}\|^{1/s} + a_{\mathbf{C},s})\,\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))}$$

and $\|\mathbf{Cf}^* - \mathbf{RHS}_{\mathbf{Cf}^*}[\varepsilon]\|_{\ell^2(\nabla_\mathcal{Z})} \le \varepsilon$, with the number of arithmetic operations and storage locations used by the call $\mathbf{RHS}_{\mathbf{Cf}^*}[\varepsilon]$ being bounded by some absolute multiple of

$$a_{\mathbf{B}^*,s}(\|\mathbf{C}\|^{1/s} + a_{\mathbf{C},s})\,\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{X}))} + 1.$$

Remark 8.6. *The properties of $\mathbf{RHS}_{\mathbf{Cf}^*}$ stated in Corollary 8.5 show that $\mathbf{RHS}_{\mathbf{Cf}^*}$ is a valid routine for approximating $\mathbf{Cf}^*$ in the sense of Remark 8.4.*

In the special case when $\mathbf{B}$ is *symmetric positive definite*, i.e., $\nabla_\mathcal{X} = \nabla_\mathcal{Y}$ and $\mathbf{B} = \mathbf{B}^* > 0$, the two Galerkin methods considered in the papers [9, 10] were shown to be *quasi-optimal* in the following sense.

Theorem 8.7. *Suppose that, in (6.1), $\mathbf{B}^*$ is s^*-admissible; then, for any $\varepsilon > 0$, the two adaptive Galerkin methods from [9, 10] produce approximations $\mathbf{u}_\varepsilon$ to $\mathbf{u}$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell^2(\nabla_\mathcal{Y})} \le \varepsilon$. Suppose that, in (6.1), for some $s > 0$ we have $\mathbf{u} \in \mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))$; then $\#\mathrm{supp}\,\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))}$ and if, moreover, $s < s^*$, then the number of arithmetic operations and storage locations required by a call of either of these adaptive solvers with tolerance $\varepsilon > 0$ is bounded by some multiple of*

$$\varepsilon^{-1/s}(1 + a_{\mathbf{B},s})\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))} + 1.$$

The multiples depend on s only when s tends to 0 or ∞, and on $\|\mathbf{B}\|$ and $\|\mathbf{B}^{-1}\|$ when they tend to infinity.

The method from [10] consists of the application of a damped Richardson iteration to $\mathbf{Bu} = \mathbf{f}$, where the required residual computations are approximated using calls of $\mathbf{APPLY}_\mathbf{B}$ and $\mathbf{RHS}_\mathbf{f}$ within tolerances that decrease linearly with the iteration counter.

The method from [9] produces a sequence $\Xi_0 \subset \Xi_1 \subset \cdots \subset \nabla_\mathcal{X}$, together with a corresponding sequence of (approximate) finitely supported Galerkin solutions $\mathbf{u}_i \in \ell^2(\Xi_i)$. The coefficients of approximate residuals $\mathbf{f} - \mathbf{Bu}_i$ are used as indicators how to expand Ξ_i to Ξ_{i+1} so that the latter gives rise to an improved Galerkin approximation.

Both methods rely on a recurrent coarsening of the approximation vectors, where small coefficients are removed in order to keep an optimal balance between accuracy and support length; in [15], a modification of the algorithm introduced in [9] was proposed, which does not require the coarsening step.

The key to why s^*-admissibility of $\mathbf{B}$ can be expected is the observation that for a large class of operators the stiffness matrix with respect to suitable wavelet bases is close to a computable sparse matrix, in the sense of the following definition.

Definition 8.8. $\mathbf{B} \in \mathcal{L}(\ell^2(\nabla_\mathcal{X}), \ell^2(\nabla_\mathcal{Y}))$ *is s^*-*computable *if, for each $N \in \mathbb{N}$, there exists a $\mathbf{B}^{[N]} \in \mathcal{L}(\ell^2(\nabla_\mathcal{X}), \ell^2(\nabla_\mathcal{Y}))$ having in each 'column' at most N nonzero entries whose joint computation takes an absolute multiple of N operations, such that the* computability constants

$$c_{\mathbf{B},\bar{s}} := \sup_{N \in \mathbb{N}} N\|\mathbf{B} - \mathbf{B}^{[N]}\|^{1/\bar{s}}_{\ell^2(\nabla_\mathcal{X}) \to \ell^2(\nabla_\mathcal{Y})} \tag{8.3}$$

are finite for any $\bar{s} \in (0, s^*)$. The notion of s^*-computability of $\mathbf{B}^*$ is defined analogously.

Theorem 8.9. *An s^*-computable $\mathbf{B}$ is s^*-admissible. Moreover, for $\bar{s} < s^*$, $a_{\mathbf{B},\bar{s}} \lesssim c_{\mathbf{B},\bar{s}}$ where the constant in this estimate depends only on $\bar{s} \downarrow 0$, $\bar{s} \uparrow s^*$, and on $\|\mathbf{B}\| \to \infty$.*

This theorem is proved by constructing a suitable $\mathbf{APPLY}_\mathbf{B}$ routine as in [9, §6.4] (a log factor in the complexity estimate there due to sorting was removed later through the use of approximate sorting, see [13] and the references there).

Remark 8.10. *Theorem 8.7 requires that* $\mathbf{B}$ *is* s^**-admissible for an* $s^* > s$ *when* $\mathbf{u} \in \mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{X}))$*. Generally this value of* s *is unknown, and so the condition on* s^* *should be interpreted in the sense that* s^* *has to be larger than any* s *for which membership of the solution* $\mathbf{u}$ *in* $\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{X}))$ *can be expected.*

The approach from [10] applies also to the saddle point variational principle (5.5) whenever one has a linearly convergent stationary iterative scheme available for the matrix-vector problem $\mathbf{B}^*\mathbf{u} = \mathbf{f}^*$. There is, unfortunately, no such scheme available for a general boundedly invertible $\mathbf{B}^*$. In particular, for the stiffness matrices $\mathbf{B}^*$ resulting from the space-time saddle-point formulation (5.5) of the infinite-dimensional Fokker–Planck equation no directly applicable scheme is available.

A remedy proposed in [10] is to apply adaptive Galerkin discretizations to the *normal equation*

$$\mathbf{B}\mathbf{B}^*\mathbf{u} = \mathbf{B}\mathbf{f}^*. \tag{8.4}$$

By Theorem 5.3, the operator $\mathbf{B}\mathbf{B}^* \in \mathcal{L}(\ell^2(\nabla_\mathcal{Y}), \ell^2(\nabla_\mathcal{Y}))$ is boundedly invertible, symmetric positive definite, with

$$\|\mathbf{B}\mathbf{B}^*\|_{\ell^2(\nabla_\mathcal{Y})\to\ell^2(\nabla_\mathcal{Y})} \leq \|\mathbf{B}^*\|^2_{\ell^2(\nabla_\mathcal{Y})\to\ell^2(\nabla_\mathcal{X})}, \; \|(\mathbf{B}\mathbf{B}^*)^{-1}\|_{\ell^2(\nabla_\mathcal{Y})\to\ell^2(\nabla_\mathcal{Y})} \leq \|(\mathbf{B}^*)^{-1}\|^2_{\ell^2(\nabla_\mathcal{X})\to\ell^2(\nabla_\mathcal{Y})}.$$

Now let $\mathbf{u} \in \mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{X}))$, and for some $s^* > s$, let $\mathbf{B}$ and $\mathbf{B}^*$ be s^*-admissible. By Corollary 8.5, with $\mathbf{B}^*$ in place of $\mathbf{C}$, a valid $\mathbf{RHS}_{\mathbf{Bf}^*}$ routine is given by (8.2), and $\mathbf{B}\mathbf{B}^*$ is s^*-admissible with a valid $\mathbf{APPLY}_{\mathbf{BB}^*}$ routine given by (8.1). In the context of (5.5), one execution of $\mathbf{APPLY}_{\mathbf{BB}^*}$ corresponds to one (approximate) sweep over the "primal" problem, followed by one (approximate) sweep over the dual problem, respectively. A combination of Theorem 8.7 and Corollary 8.5 yields the following result obtained in [24].

Theorem 8.11. *For any* $\varepsilon > 0$*, the adaptive wavelet methods from* [10] *or from* [9] *and* [15] *applied to the normal equations* (8.4) *using the above* $\mathbf{APPLY}_{\mathbf{BB}^*}$ *and* $\mathbf{RHS}_{\mathbf{B}f^*}$ *routines produce an approximation* $\mathbf{u}_\varepsilon$ *to* $\mathbf{u}$ *with* $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{\ell^2(\nabla_\mathcal{Y})} \leq \varepsilon$*.*

Suppose that for some $s > 0$*,* $\mathbf{u} \in \mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))$*; then* $\#\mathrm{supp}\,\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))}$*, with the constant in this bound only being dependent on* s *when it tends to* 0 *or* ∞*, and on* $\|\mathbf{B}^*\|$ *and* $\|(\mathbf{B}^*)^{-1}\|$ *when they tend to infinity.*

Suppose that $s < s^*$*; then, the number of arithmetic operations and storage locations required by a call of either of these adaptive wavelet methods with tolerance* ε *is bounded by some multiple of*

$$1 + \varepsilon^{-1/s}(1 + a_{\mathbf{B}^*,s}(1 + a_{\mathbf{B},s}))\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_\mathcal{Y}))},$$

where this multiple only depends on s *when it tends to* 0 *or* ∞*, and on* $\|\mathbf{B}^*\|$ *and* $\|(\mathbf{B}^*)^{-1}\|$ *when they tend to infinity.*

9. Infinite-dimensional Fokker–Planck equation as an infinite matrix vector equation

We apply the foregoing abstract concepts to the space-time variational formulation (5.5) of the infinite-dimensional Fokker–Planck equation. To construct Riesz bases for $\mathcal{X}_{0,\{T\}}$ and $\mathcal{Y}$, we use that

$$\mathcal{X}_{0,\{T\}} = (\mathrm{L}^2(0,T)\otimes\mathcal{V}) \cap (\mathrm{H}^1_{0,\{T\}}(0,T)\otimes\mathcal{V}^*) \quad \text{and} \quad \mathcal{Y} = \mathrm{L}^2(0,T)\otimes\mathcal{V}$$

with the spaces $\mathcal{X}_{0,\{T\}}$ as defined in (5.4), and $\mathcal{Y} = \mathrm{L}^2(0,T;\mathcal{V})$; recall that $\mathcal{H} = \mathrm{L}^2(D,\mu)$, $\mathcal{V} = \mathrm{W}^{1,2}(D,\mu)$. Let

$$\Upsilon = \{H_\gamma : \gamma \in \Gamma\} \subset \mathcal{V}, \quad \text{with} \quad H_\gamma(q) = \prod_{k=1}^\infty H_{\gamma_k}\left(\frac{q_k}{\sqrt{\lambda_k}}\right),$$

denote the polynomial chaos basis (4.2). By Theorem 4.3, Υ is a normalized Riesz basis for $[\mathrm{L}^2(H,\mu)]^d$, which, when renormalized by the scaling sequence $(\langle\gamma,\lambda^{-1}\rangle)_{\gamma\in\Gamma}$ in $\mathcal{V}$ and by the scaling sequence $(\langle\gamma,\lambda\rangle)_{\gamma\in\Gamma}$ in $\mathcal{V}^*$, is a Riesz basis for $\mathcal{V}$ and $\mathcal{V}^*$, respectively. Let

$$\Theta = \{\theta_\lambda : \lambda \in \nabla_t\} \subset \mathrm{H}^1_{0,\{T\}}(0,T)$$

be a family of functions that is a normalized Riesz basis for $\mathrm{L}^2(0,T)$, which, renormalized in $\mathrm{H}^1(0,T)$, is a Riesz basis for $\mathrm{H}^1_{0,\{T\}}(0,T)$. From [16, Prop. 1 and 2], it follows that then the collection $\Theta\otimes\Upsilon$, normalized in $\mathcal{X}$, i.e., the collection

$$\left\{(t,q)\mapsto \frac{\theta_\lambda(t)H_\gamma(q)}{\sqrt{\langle\gamma,\lambda^{-1}\rangle+\|\theta_\lambda\|^2_{\mathrm{H}^1(0,T)}\langle\gamma,\lambda\rangle}} : (\lambda,\gamma)\in\nabla_{\mathcal{X}}:=\nabla_t\times\Gamma\right\},$$

is a Riesz basis for $\mathcal{X}_{0,\{T\}}$, and that $\Theta\otimes\Upsilon$ normalized in $\mathcal{Y}$, i.e., the collection

$$\left\{(t,q)\mapsto \frac{\theta_\lambda(t)H_\gamma(q)}{\|H_\gamma\|_{\mathcal{V}}} : (\lambda,\gamma)\in\nabla_{\mathcal{Y}}:=\nabla_t\times\Gamma\right\},$$

is a Riesz basis for $\mathcal{Y}$. Moreover, denoting the Riesz basis for $\mathcal{V}^*$ consisting of the collection Υ normalized in $\mathcal{V}^*$ by $[\Upsilon]_{\mathcal{V}^*}$, and similarly for the other collections and spaces, with the notations introduced in Section 6 we have that

$$\begin{aligned}
\Lambda^{\mathcal{X}}_{[\Theta\otimes\Upsilon]_{\mathcal{X}}} &\le \max(\Lambda^{\mathrm{L}^2(0,T)}_{\Theta}\Lambda^{\mathcal{V}}_{[\Upsilon]_{\mathcal{V}}},\Lambda^{\mathrm{H}^1(0,T)}_{[\Theta]_{\mathrm{H}^1(0,T)}}\Lambda^{\mathcal{V}^*}_{[\Upsilon]_{\mathcal{V}^*}}),\\
\lambda^{\mathcal{X}}_{[\Theta\otimes\Upsilon]_{\mathcal{X}}} &\ge \min(\lambda^{\mathrm{L}^2(0,T)}_{\Theta}\lambda^{\mathcal{V}}_{[\Upsilon]_{\mathcal{V}}},\lambda^{\mathrm{H}^1(0,T)}_{[\Theta]_{\mathrm{H}^1(0,T)}}\lambda^{\mathcal{V}^*}_{[\Upsilon]_{\mathcal{V}^*}}),\\
\Lambda^{\mathcal{Y}}_{[\Theta\otimes\Upsilon]_{\mathcal{Y}}} &\le \Lambda^{\mathrm{L}^2(0,T)}_{\Theta}\Lambda^{\mathcal{V}}_{[\Upsilon]_{\mathcal{V}}},\\
\lambda^{\mathcal{Y}}_{[\Theta\otimes\Upsilon]_{\mathcal{Y}}} &\ge \lambda^{\mathrm{L}^2(0,T)}_{\Theta}\lambda^{\mathcal{V}}_{[\Upsilon]_{\mathcal{V}}}.
\end{aligned}$$

Denoting by $\|\Upsilon\|_{\mathcal{V}}$ the infinite diagonal matrix with diagonal entries $\|H_\gamma\|_{\mathcal{V}}$ where $\gamma\in\Gamma$, and similarly for the other collections and spaces, the stiffness or system matrix $\mathbf{B}^*$ corresponding to the variational form (5.5) and the Riesz bases $[\Theta\otimes\Upsilon]_{\mathcal{Y}}$, $[\Theta\otimes\Upsilon]_{\mathcal{X}}$ for $\mathcal{Y}$ and $\mathcal{X}_{0,\{T\}}$ is given by the infinite matrix

$$\begin{aligned}
\mathbf{B}^* &= \mathfrak{B}^*([\Theta\otimes\Upsilon]_{\mathcal{Y}},[\Theta\otimes\Upsilon]_{\mathcal{X}})\\
&= \left[\mathrm{Id}_t\otimes\|\Upsilon\|^{-1}_{\mathcal{V}}\right]\circ\left[-(\Theta,\Theta')_{\mathrm{L}^2(0,T)}\otimes(\Upsilon,\Upsilon)_{\mathcal{H}}+\int_0^T\mathfrak{a}(\Theta\otimes\Upsilon,\Theta\otimes\Upsilon)\,\mathrm{d}t\right]\circ\left[\|\Theta\otimes\Upsilon\|^{-1}_{\mathcal{X}}\right],
\end{aligned}$$

where Id_t denotes the identity operator with respect to the t variable and the symbol $\circ$ signifies composition of operators.

Writing the solution $\widehat{\psi}$ of (5.5) as $\widehat{\psi}=\mathbf{u}^\top[\Theta\otimes\Upsilon]_{\mathcal{X}}$, we deduce that $\mathbf{u}$ is the solution of the infinite matrix-vector equation $\mathbf{B}^*\mathbf{u}=\mathbf{f}^*$ with

$$\mathbf{f}^* := \left[(\widehat{\psi}_0,\Upsilon)_{\mathcal{H}}\right]. \tag{9.1}$$

Introducing the infinite matrices

$$\mathbf{D}_1 := (\|\Theta\|_{\mathrm{H}^1(0,T)}\otimes\|\Upsilon\|_{\mathcal{V}^*})\|\Theta\otimes\Upsilon\|^{-1}_{\mathcal{X}} \quad\text{and}\quad \mathbf{D}_2 := (\mathrm{Id}_t\otimes\|\Upsilon\|_{\mathcal{V}})\|\Theta\otimes\Upsilon\|^{-1}_{\mathcal{X}},$$

both being diagonal, with entries in modulus less than 1, the infinite matrix operator $\mathbf{B}^*$ can be written as

$$\mathbf{B}^* = \left[-(\Theta,[\Theta]'_{\mathrm{H}^1(0,T)})_{\mathrm{L}^2(0,T)}\otimes([\Upsilon]_{\mathcal{V}^*},[\Upsilon]_{\mathcal{V}})_{\mathcal{H}}\,\mathbf{D}_1+\int_0^T\mathfrak{a}(\Theta\otimes[\Upsilon]_{\mathcal{V}},\Theta\otimes[\Upsilon]_{\mathcal{V}})\,\mathrm{d}t\,\mathbf{D}_2\right]. \tag{9.2}$$

10. s^*-admissibility of $\mathbf{B}$ and $\mathbf{B}^*$ from (9.2) and of its adjoint

By Theorem 8.9, the s^*-admissibility of $\mathbf{B}$ and $\mathbf{B}^*$ follows from their s^*-computability. To verify the s^*-computability of tensor products of possibly infinite matrices, it suffices to analyze the s^*-computability of the factors, according to the following result from [24, 20].

Proposition 10.1. *Let, for some $s^*>0$, $\mathbf{D}$ and $\mathbf{E}$ be s^*-computable. Then,*

(a) *$\mathbf{D}\otimes\mathbf{E}$ is s^*-computable with computability constant satisfying, for $0<\bar{s}<\tilde{s}<s^*$, $c_{\mathbf{D}\otimes\mathbf{E},\bar{s}}\lesssim(c_{\mathbf{D},\tilde{s}}\,c_{\mathbf{E},\tilde{s}})^{\tilde{s}/\bar{s}}$; and*

(b) *for any $\varepsilon\in(0,s^*)$, $\mathbf{D}\otimes\mathbf{E}$ is $(s^*-\varepsilon)$–computable, with computability constant $c_{\mathbf{D}\otimes\mathbf{E},\bar{s}}$ satisfying, for $0<\bar{s}<s^*-\varepsilon<\tilde{s}<s^*$, $c_{\mathbf{D}\otimes\mathbf{E},\bar{s}}\lesssim\max(c_{\mathbf{D},\tilde{s}},1)\max(c_{\mathbf{E},\tilde{s}},1)$.*

The constants absorbed in the $\lesssim$ symbol in the bounds on the computability constants in (a) and (b) are only dependent on $\tilde{s} \downarrow 0$, $\tilde{s} \to \infty$ and $\tilde{s} - \bar{s} \downarrow 0$.

In view of the representation (9.2) of $\mathbf{B}^*$, using Corollary 8.5 and Proposition 10.1, part (a), for proving s^*-admissibility of $\mathbf{B}$ and $\mathbf{B}^*$ it suffices to show that $(\Upsilon, \Upsilon)_{\mathcal{H}}$ and $\int_0^T \mathfrak{a}(\Theta \otimes [\Upsilon]_{\mathcal{V}}, \Theta \otimes [\Upsilon]_{\mathcal{V}})\,\mathrm{d}t$ and its adjoint are s^*-admissible. To prove this, by Theorem 8.9 we need to verify that these objects are s^*-computable. This will follow from the fact that $([\Theta]'_{\mathrm{H}^1(0,T)}, \Theta)_{\mathrm{L}^2(0,T)}$, $(\Theta, [\Theta]'_{\mathrm{H}^1(0,T)})_{\mathrm{L}^2(0,T)}$, $([\Upsilon]_{\mathcal{V}^*}, [\Upsilon]_{\mathcal{V}})_{\mathcal{H}}$ and $([\Upsilon]_{\mathcal{V}}, [\Upsilon]_{\mathcal{V}^*})_{\mathcal{H}}$ are s^*-computable and, in view of the definition of $\mathfrak{a}(\cdot,\cdot)$ in (5.6), from a sparsity assumption on the coefficient matrix $\mathbf{A}$.

10.1. **Choice of Riesz bases Θ and Υ.** We have already assumed that $\Theta = \{\theta_\lambda : \lambda \in \nabla_t\}$ is a normalized Riesz basis of $\mathrm{L}^2(0,T)$, which, when renormalized in $\mathrm{H}^1(0,T)$ is a Riesz basis for $\mathrm{H}^1_{0,\{T\}}(0,T)$. We will now select the basis Θ to be a *wavelet basis* that satisfies certain additional assumptions. Specifically, we shall assume that the basis functions θ_λ of Θ are:

(t1) *local*, i.e., $\sup_{x\in[0,1],\ell\in\mathbb{N}_0} \#\{|\lambda| = \ell : x \in \operatorname{supp}\theta_\lambda\} < \infty$ and $\operatorname{supp}\theta_\lambda \lesssim 2^{-|\lambda|}$;

(t2) *piecewise polynomial of order d_t*, where by "piecewise" we mean that the singular support consists of a set of points whose cardinality is uniformly bounded;

(t3) *globally continuous*, specifically $\|\theta_\lambda\|_{W^k_\infty(0,1)} \lesssim 2^{|\lambda|(\frac{1}{2}+k)}$ for $k \in \{0,1\}$;

(t4) for $|\lambda| > 0$, have $\tilde{d}_t \geq d_t$ *vanishing moments.*

The assumptions (t1)–(t4) can be met by wavelet constructions (see, e.g. [8] and the references therein).

As a Riesz basis in D we choose the (countable) "polynomial chaos" basis $\Upsilon = \{H_\gamma : \gamma \in \Gamma\}$. According to Theorem 4.3, Υ is an orthonormal Riesz basis of $\mathrm{L}^2(D,\mu)$ that, normalized in $\mathcal{V} = \mathrm{W}^{1,2}(D,\mu)$ or its dual $\mathcal{V}^*$, is a Riesz basis for these spaces, respectively. Indeed, on denoting by $\mathbf{D}^\lambda$ the diagonal matrix

$$\mathbf{D}^\lambda = \operatorname{diag}\{(\langle\gamma, \lambda^{-1}\rangle)^{1/2} : \gamma \in \Gamma\}, \tag{10.1}$$

we observe that the diagonal entries of $\mathbf{D}^\lambda$ relate to the $\mathcal{V}$ and $\mathcal{V}^*$ norms of $[\Upsilon]_{\mathcal{H}}$ as follows:

$$\mathbf{D}^\lambda \simeq \|[\Upsilon]_{\mathcal{H}}\|_{\mathcal{V}}, \quad (\mathbf{D}^\lambda)^{-1} \simeq \|[\Upsilon]_{\mathcal{H}}\|_{\mathcal{V}^*}.$$

Therefore, the collection $[\Upsilon]_{\mathcal{V}} := (\mathbf{D}^\lambda)^{-1}[\Upsilon]_{\mathcal{H}} = \{(D^\lambda_\gamma)^{-1} H_\gamma : \gamma \in \Gamma\}$ is a Riesz basis of $\mathcal{V} = \mathrm{W}^{1,2}(D,\mu)$. This follows readily from the $\mathrm{L}^2(D,\mu)$ orthonormality of Υ and from the identity (cf. [12, (9.2.11)])

$$\mathrm{D}_k H_\gamma = \gamma_k^{1/2} \lambda_k^{-1/2} H_{\gamma_k - 1}(W_{e_k}) H^{(k)}_\gamma, \qquad \text{for} \quad k = 1,2,\ldots \quad \text{and} \quad \gamma \in \Gamma \quad \text{with} \quad \gamma_k \geq 1; \tag{10.2}$$

with the notational convention $H^{(k)}_\gamma := \prod_{j\neq k} H_{\gamma_j}$ for $\gamma \in \Gamma$. If $\gamma_k = 0$, then $\mathrm{D}_k H_\gamma = 0$.

By duality, the family $[\Upsilon]_{\mathcal{V}^*} := \mathbf{D}^\lambda [\Upsilon]_{\mathcal{H}}$ is a Riesz basis of $\mathcal{V}^*$.

10.2. **s^*-computability of $([\Theta]'_{\mathrm{H}^1(0,T)}, \Theta)_{\mathrm{L}^2(0,T)}$ and of its adjoint.** By (t1), (t2) and (t4), for each $\lambda \in \nabla_t$ and $\ell \in \mathbb{N}_0$, the number of $\mu \in \nabla_t$ with $|\mu| = \ell$ and $\int_0^T \theta'_\lambda \theta_\mu\,\mathrm{d}t \neq 0$ or $\int_0^T \theta'_\mu \theta_\lambda\,\mathrm{d}t \neq 0$ is bounded, uniformly in λ and ℓ. Indeed, $\int_0^T \theta'_\lambda \theta_\mu\,\mathrm{d}t$ can only be nonzero when θ_μ does not vanish on the singular support of θ_λ, and using integration by parts, $\int_0^T \theta'_\mu \theta_\lambda\,\mathrm{d}t$ can only be nonzero when θ_μ does not vanish on the singular support of θ_λ or at $t \in \{0,T\}$.

As a consequence of Θ being of order $d_t \geq 1$ we have that

$$2^{|\lambda|} = 2^{|\lambda|}\|\theta_\lambda\|_{\mathrm{L}^2(0,T)} = 2^{|\lambda|}\|(\mathrm{Id} - Q_{|\lambda|-1})\theta_\lambda\|_{\mathrm{L}^2(0,T)} \lesssim \|\theta_\lambda\|_{\mathrm{H}^1(0,T)}.$$

Using (t1) and (t3), we infer that

$$\|\theta_\lambda\|^{-1}_{\mathrm{H}^1(0,T)} \left| \int_0^T \theta'_\lambda \theta_\mu\,\mathrm{d}t \right| \lesssim 2^{-|\lambda|} 2^{-\max(|\lambda|,|\mu|)} 2^{|\lambda|(\frac{1}{2}+1)} 2^{\frac{1}{2}|\mu|} = 2^{-\frac{1}{2}\left||\lambda|-|\mu|\right|}. \tag{10.3}$$

Finally, we note that any entry of $\langle[\Theta]'_{\mathrm{H}^1(0,T)}, \Theta\rangle_{\mathrm{L}^2(0,T)}$ can be evaluated in closed form in $\mathcal{O}(1)$ work and memory. Schur's Lemma (cf. [23], p.6, ¶2 and [6], p.449, Theorem B) now implies that $([\Theta]'_{\mathrm{H}^1(0,T)}, \Theta)_{\mathrm{L}^2(0,T)}$ and $(\Theta, [\Theta]'_{\mathrm{H}^1(0,T)})_{\mathrm{L}^2(0,T)}$ are ∞-computable.

Note that from $\operatorname{diam}(\operatorname{supp}\theta_\lambda) \lesssim 2^{-|\lambda|}$ (cf. (t1)) and from and (t3) we deduce that $\|\theta_\lambda\|_{\mathrm{H}^1(0,T)} \lesssim 2^{|\lambda|}$, and thus that

$$\|\theta_\lambda\|_{\mathrm{H}^1(0,T)} \eqsim 2^{|\lambda|}.$$

Remark 10.2. *Suppose that instead of* (t3)*, the θ_λ belong also to $\mathrm{C}^{r_t}(0,T)$ for some $r_t \in \mathbb{N}$ (necessarily with $r_t \leq d_t - 2$), i.e., that $\|\theta_\lambda\|_{\mathrm{W}^{s,\infty}(0,1)} \lesssim 2^{|\lambda|(\frac{1}{2}+s)}$ for $s \in \{0, r_t+1\}$. Then, by subtracting a suitable polynomial of order r_t from θ'_λ in* (10.3)*, and using that θ_μ has $\tilde{d}_t \geq d_t \geq r_t$ vanishing moments one deduces that, for $|\lambda| \leq |\mu|$,*

$$\|\theta_\lambda\|^{-1}_{\mathrm{H}^1(0,T)} \left| \int_0^T \theta'_\lambda \theta_\mu \,\mathrm{d}t \right| \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{1}{2}+r_t)}.$$

Similarly, for $|\lambda| \geq |\mu|$, using integration by parts one obtains that

$$\|\theta_\lambda\|^{-1}_{\mathrm{H}^1(0,T)} \left| \int_0^T \theta'_\lambda \theta_\mu \,\mathrm{d}t \right| \lesssim 2^{-\left||\lambda|-|\mu|\right|(\frac{3}{2}+r_t)}$$

if $t \mapsto \theta_\lambda(t)\,\theta_\mu(t)$ vanishes at $t \in \{0, T\}$. Since the wavelets in time do not satisfy Dirichlet boundary conditions, there are, however, $\lambda, \mu \in \nabla_t$ with $|\lambda| \geq |\mu|$ for which $t \mapsto \theta_\lambda(t)\,\theta_\mu(t)$ does not vanish at the boundary. For those entries (10.3) *cannot be improved.*

Remark 10.3. *In* [8]*, multiresolutions $\Theta = (\theta_\lambda)_{\lambda\in\nabla_t}$ were constructed for which the matrices $([\Theta]'_{\mathrm{H}^1(0,T)}, \Theta)_{\mathrm{L}^2(0,T)}$ are* sparse without compression.

10.3. **s^*-computability of $(\Upsilon, \Upsilon)_{\mathcal{H}}$ and $([\Upsilon]_{\mathcal{V}^*}, [\Upsilon]_{\mathcal{V}})_{\mathcal{H}}$ and of its adjoint.** By the $\mathrm{L}^2(H,\mu)$-orthonormality (4.4) of the family Υ of Hermite polynomials, both $(\Upsilon, \Upsilon)_{\mathcal{H}}$ and $([\Upsilon]_{\mathcal{V}^*}, [\Upsilon]_{\mathcal{V}})_{\mathcal{H}}$ and their adjoints are bi-infinite, diagonal matrices that are therefore ∞-computable.

10.4. **s^*-computability of $\int_0^T \mathfrak{a}(\Theta \otimes [\Upsilon]_{\mathcal{V}}, \Theta \otimes [\Upsilon]_{\mathcal{V}})\,\mathrm{d}t$ and of its adjoint.** We have that

$$\int_0^T \mathfrak{a}(\Theta \otimes [\Upsilon]_{\mathcal{V}}, \Theta \otimes [\Upsilon]_{\mathcal{V}})\,\mathrm{d}t = (\Theta, \Theta)_{\mathrm{L}^2(0,T)} \otimes \mathfrak{a}([\Upsilon]_{\mathcal{V}}, [\Upsilon]_{\mathcal{V}}). \tag{10.4}$$

Since the form $\mathfrak{a}(\cdot,\cdot)$ defined in (5.8) is symmetric, by Proposition 10.1 it suffices to investigate s^*-computability of the two factors on the right-hand side of (10.4) in order to deduce s^*-computability of $\mathbf{B}$ and $\mathbf{B}^*$.

As was noted in [24, Sec. 8.5], under the assumptions (t1)–(t4) above, $(\Theta, \Theta)_{\mathrm{L}^2(0,T)}$ is ∞-computable, so it remains to address the s^*-computability of $\mathfrak{a}([\Upsilon]_{\mathcal{V}}, [\Upsilon]_{\mathcal{V}})$.

In view of the definition (5.8) of the bilinear form $\mathfrak{a}(\cdot,\cdot)$ appearing in (5.6) and noting Definition 8.8, the s^*-computability of the infinite matrix $\mathbf{G} = \mathfrak{a}([\Upsilon]_{\mathcal{V}}, [\Upsilon]_{\mathcal{V}})$ depends on the structure of the infinite matrix $\mathbf{A} = (A_{ij})_{i,j=1}^\infty$ in (5.6). For the sake of simplicity of the exposition we shall suppose that $d = 1$, so that $D = \mathbb{R} \times \mathbb{R} \times \cdots$. Thanks to (10.4),

$$\mathfrak{a}([\Upsilon]_{\mathcal{V}}, [\Upsilon]_{\mathcal{V}}) = (G_{\gamma,\gamma'})_{\gamma,\gamma\in\Gamma}$$

where, on recalling the definition (5.8) of the symmetric bilinear form $\mathfrak{a}(\cdot,\cdot)$, the entries $G_{\gamma\gamma'}$ of the symmetric matrix $\mathbf{G}$ are, for $\gamma, \gamma' \in \Gamma$, given by

$$G_{\gamma\gamma'} = \sum_{i,j\geq 1} A_{ij} (D^\lambda_\gamma)^{-1} (\mathrm{D}_{q_i} H_\gamma, \mathrm{D}_{q_j} H_{\gamma'})_{\mathrm{L}^2(D,\mu)} (D^\lambda_{\gamma'})^{-1} + \delta_{\gamma\gamma'}.$$

Clearly, if $\gamma = 0$, then $G_{\gamma\gamma'} = \delta_{\gamma\gamma'}$ for all $\gamma' \in \Gamma$; similarly, if $\gamma' = 0$, then $G_{\gamma\gamma'} = \delta_{\gamma\gamma'}$ for all $\gamma \in \Gamma$. We shall therefore focus our attention on the nontrivial case when $\gamma, \gamma' \in \Gamma \setminus \{0\}$; for any such γ and γ', we have by (10.2) that

$$\begin{aligned} G_{\gamma\gamma'} &= \sum_{i,j\geq 1\,:\,\gamma_i\geq 1, \gamma'_j \geq 1} A_{ij} (D^\lambda_\gamma)^{-1} \left(\sqrt{\gamma_i/\lambda_i}\, H_{\gamma_i - 1} H^{(i)}_\gamma, \sqrt{\gamma'_j/\lambda_j}\, H_{\gamma'_j-1} H^{(j)}_{\gamma'} \right)_{\mathrm{L}^2(D,\mu)} (D^\lambda_{\gamma'})^{-1} + \delta_{\gamma\gamma'} \\ &= \sum_{i,j\geq 1\,:\,\gamma_i\geq 1, \gamma'_j \geq 1} A_{ij} (D^\lambda_\gamma)^{-1} (D^\lambda_{\gamma'})^{-1} \sqrt{\frac{\gamma_i \gamma'_j}{\lambda_i \lambda_j}} \left(H_{\gamma_i-1} H^{(i)}_\gamma, H_{\gamma'_j-1} H^{(j)}_{\gamma'} \right)_{\mathrm{L}^2(D,\mu)} + \delta_{\gamma\gamma'}. \end{aligned}$$

We now impose additional hypotheses on the matrix $\mathbf{A}$. It is illustrative (and will be used in what follows) to first consider the special case $A_{ij} = \delta_{ij}$. We denote the corresponding matrix $\mathbf{G}$ by $\mathbf{G}^{(0)}$. Inserting this choice of A_{ij} into the previous expression, we obtain from the $\mathrm{L}^2(D,\mu)$ orthonormality (4.4) of the family $\{H_\gamma : \gamma \in \Gamma\}$ that

$$\begin{aligned}
G^{(0)}_{\gamma\gamma'} &= \sum_{i,j\geq 1\,:\,\gamma_i\geq 1, \gamma'_j \geq 1} \delta_{ij} \sqrt{\frac{\gamma_i\gamma'_j}{\lambda_i\lambda_j}}\, (D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1} \left(H_{\gamma_i-1}H^{(i)}_\gamma, H_{\gamma'_j-1}H^{(j)}_{\gamma'}\right)_{\mathrm{L}^2(D,\mu)} + \delta_{\gamma\gamma'} \\
&= \sum_{i\geq 1\,:\,\gamma_i\geq 1, \gamma'_i \geq 1} \sqrt{\frac{\gamma_i\gamma'_i}{\lambda_i^2}}\, (D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1} \left(H_{\gamma_i-1}H^{(i)}_\gamma, H_{\gamma'_i-1}H^{(i)}_{\gamma'}\right)_{\mathrm{L}^2(D,\mu)} + \delta_{\gamma\gamma'} \\
&= \delta_{\gamma\gamma'}(D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1} \sum_{i\,:\,\gamma_i\geq 1, \gamma'_i\geq 1} \sqrt{\frac{\gamma_i\gamma'_i}{\lambda_i^2}} + \delta_{\gamma\gamma'} \\
&= \delta_{\gamma\gamma'}(D^\lambda_\gamma)^{-2}\langle \gamma, \lambda^{-1}\rangle + \delta_{\gamma\gamma'} = 2\delta_{\gamma\gamma'}.
\end{aligned}$$

For $A_{ij} = \delta_{ij}$ the infinite matrix $\mathbf{G} = \mathbf{G}^{(0)}$ is, therefore, ∞-computable.

We next turn to the tridiagonal matrix $\mathbf{A}$ considered in Example 5.1, whose entries are

$$A_{ij} = \delta_{ij} + \epsilon_i\delta_{i,j-1} + \epsilon_i\delta_{i,j+1}, \qquad i,j = 1,2,\dots. \tag{10.5}$$

With the diagonal term having already been discussed, by superposition we may now confine ourselves to investigating the computability of the matrix $\mathbf{G}$ when the matrix $\mathbf{A}$ has entries $A_{ij} = \epsilon_i\delta_{i,j-1}$, $i,j = 1,2,\dots$, and when $\mathbf{A}$ has entries $A_{ij} = \epsilon_i\delta_{i,j+1}$, $i,j = 1,2,\dots$. The matrices $\mathbf{G}$ corresponding to these two cases will be denoted below by $\mathbf{G}^{(-)}$ and $\mathbf{G}^{(+)}$, and we denote their respective entries by $G^{(+)}_{\gamma\gamma'}$ and $G^{(-)}_{\gamma\gamma'}$, with $\gamma,\gamma' \in \Gamma$. For the sake of excluding trivial situations we shall assume henceforth that $\epsilon_i \neq 0$, $i = 1,2,\dots$.

Given $\gamma,\gamma' \in \Gamma\setminus\{0\}$, we calculate, as before, that

$$\begin{aligned}
G^{(\pm)}_{\gamma\gamma'} &= \sum_{i,j\geq 1\,:\,\gamma_i\geq 1, \gamma'_j\geq 1} \epsilon_i\delta_{i,j\pm 1}\sqrt{\frac{\gamma_i\gamma'_j}{\lambda_i\lambda_j}}\,(D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1}\left(H_{\gamma_i-1}H^{(i)}_\gamma, H_{\gamma'_j-1}H^{(j)}_{\gamma'}\right)_{\mathrm{L}^2(D,\mu)} + \delta_{\gamma\gamma'} \\
&= \sum_{i\geq 1\,:\,\gamma_i\geq 1, \gamma'_{i\mp 1}\geq 1} \epsilon_i\sqrt{\frac{\gamma_i\gamma'_{i\mp 1}}{\lambda_i\lambda_{i\mp 1}}}\,(D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1}\left(H_{\gamma_i-1}H^{(i)}_\gamma, H_{\gamma'_{i\mp 1}-1}H^{(i\mp 1)}_{\gamma'}\right)_{\mathrm{L}^2(D,\mu)} + \delta_{\gamma\gamma'},
\end{aligned} \tag{10.6}$$

with the notational convention $\gamma'_0 := 0$ and $\lambda_0 := \lambda_1 (> 0)$. Thus, for example, if $\gamma = 1_1 := (1,0,0,\dots)$, then $G^{(+)}_{\gamma\gamma'} = \delta_{\gamma\gamma'}$ for all $\gamma' \in \Gamma\setminus\{0\}$; $G^{(-)}_{\gamma\gamma'} = \epsilon_1$ for $\gamma' = (0,1,0,0,\dots)$, and $G^{(-)}_{\gamma\gamma'} = \delta_{\gamma\gamma'}$ for all other $\gamma' \in \Gamma\setminus\{0\}$. Also, if $\gamma = 0$, then $G^{(\pm)}_{0\gamma'} = \delta_{0,\gamma'}$ for all $\gamma' \in \Gamma$; and, analogously, if $\gamma' = 0$, then $G^{(\pm)}_{\gamma 0} = \delta_{\gamma 0}$ for all $\gamma \in \Gamma$.

It therefore remains to compute $G^{(\pm)}_{\gamma\gamma'}$ for $\gamma,\gamma' \in \Gamma\setminus\{0,1_1\}$. Hence, by defining for a fixed integer $i \geq 1$ and for $\gamma,\gamma' \in \Gamma\setminus\{0\}$ such that $\gamma_i \geq 1$ and $\gamma'_{i\mp 1} \geq 1$ the expression

$$\mathfrak{H}^{(\pm),(i)}_{\gamma\gamma'} := \left(H_{\gamma_i-1}H^{(i)}_\gamma, H_{\gamma'_{i\mp 1}-1}H^{(i\mp 1)}_{\gamma'}\right)_{\mathrm{L}^2(D,\mu)},$$

we deduce that

$$\mathfrak{H}^{(\pm),(i)}_{\gamma\gamma'} = \begin{cases} 1 & \text{if } \gamma'_i = \gamma_i - 1 \wedge \gamma'_{i\mp 1} = \gamma_{i\mp 1} + 1 \wedge \gamma'_j = \gamma_j,\ j \in \{i, i\mp 1\}^c \cap \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, for any $\gamma \in \Gamma\setminus\{0\}$ consider the finite sets $\mathcal{Z}_\pm(\gamma) \subset \Gamma\setminus\{0\}$ consisting of all $\gamma' \in \Gamma\setminus\{0\}$ for which there exists an integer $i \geq 1$ such that $\mathfrak{H}^{(\pm),(i)}_{\gamma\gamma'} = 1$. Clearly, $\mathcal{Z}_+(1_1) = \emptyset$, while $\mathcal{Z}_-(1_1)$ is equal to the singleton $\{(0,1,0,0,\dots)\}$. More generally,

(a) for each $\gamma \in \Gamma\setminus\{0,1_1\}$, the set $\mathcal{Z}_\pm(\gamma)$ is nonempty;
(b) for each $\gamma \in \Gamma\setminus\{0,1_1\}$ and $\gamma' \in \mathcal{Z}_\pm(\gamma)$, there exists a unique integer $i_\pm = i_\pm(\gamma,\gamma') \geq 1$ such that $\mathfrak{H}^{(\pm),(i_\pm)}_{\gamma\gamma'} = 1$;
(c) for each $\gamma \in \Gamma\setminus\{0,1_1\}$, the mapping $i_\pm(\gamma,\cdot) : \gamma' \in \mathcal{Z}_\pm(\gamma) \mapsto i_\pm(\gamma,\gamma') \in \mathbb{N}$ is injective;

(d) for each $\gamma \in \Gamma \setminus \{0\}$, $\gamma \notin \mathcal{Z}_\pm(\gamma)$.

Hence, for a fixed $\gamma \in \Gamma \setminus \{0, 1_1\}$, and any $\gamma' \in \mathcal{Z}_\pm(\gamma)$, we have from (10.6) (note that by (d): $\delta_{\gamma\gamma'} = 0$) that

$$G^{(\pm)}_{\gamma\gamma'} = \epsilon_{i_\pm} \sqrt{\frac{\gamma_{i_\pm}\gamma'_{i_\pm \mp 1}}{\lambda_{i_\pm}\lambda_{i_\pm\mp 1}}} \langle \gamma, \lambda^{-1}\rangle^{-\frac{1}{2}} \langle \gamma', \lambda^{-1}\rangle^{-\frac{1}{2}}, \quad \text{where} \quad \begin{cases} i_\pm = i_\pm(\gamma, \gamma') \geq 1; \\ \gamma_{i_\pm} \geq 1; \\ \gamma'_{i_\pm \mp 1} \geq 1; \end{cases} \tag{10.7}$$

otherwise, for $\gamma' \in (\Gamma \setminus \{0, 1_1\}) \setminus \mathcal{Z}_\pm(\gamma)$, we have that $G^{(\pm)}_{\gamma\gamma'} = \delta_{\gamma\gamma'}$. In any case, we deduce that in each 'row' with index $\gamma \in \Gamma \setminus \{0, 1_1\}$ the matrix $\mathbf{G}^{(+)}$ contains nonzero off-diagonal entries only in 'columns' $\gamma' \in \mathcal{Z}_+(\gamma)$; analogously, in each 'row' with index $\gamma \in \Gamma \setminus \{0\}$ the matrix $\mathbf{G}^{(-)}$ contains nonzero off-diagonal entries only in 'columns' $\gamma' \in \mathcal{Z}_-(\gamma)$. All diagonal entries of $\mathbf{G}^{(+)}$ and $\mathbf{G}^{(-)}$ are equal to 1.

An analogous calculation reveals that the nonzero off-diagonal entries of the transpose of $\mathbf{G}^{(\mp)}$ contained in 'row' γ (i.e. 'column' γ of $\mathbf{G}$) are given by

$$G^{(\mp)}_{\gamma'\gamma} = \epsilon_{j_\mp} \sqrt{\frac{\gamma_{j_\mp \pm 1}\gamma'_{j_\mp}}{\lambda_{j_\mp \pm 1}\lambda_{j_\mp}}} \langle \gamma, \lambda^{-1}\rangle^{-\frac{1}{2}} \langle \gamma', \lambda^{-1}\rangle^{-\frac{1}{2}}, \quad \text{where} \quad \begin{cases} j_\mp = j_\mp(\gamma, \gamma') \geq 1; \\ \gamma'_{j_\mp} \geq 1; \\ \gamma_{j_\mp \pm 1} \geq 1. \end{cases} \tag{10.8}$$

By taking $j_\mp = i_\pm \mp 1$, noting that $i_\pm \mapsto j_\mp := i_\pm \mp 1$ is a one-to-one correspondence, and comparing (10.8) with (10.7) we see that the number of nonzero off-diagonal entries in 'row' γ of $\mathbf{G}^{(\pm)}$ is equal to the number of nonzero off-diagonal entries in 'column' γ of $\mathbf{G}^{(\mp)}$. Since $\epsilon_{i_\pm}$ and $\epsilon_{i_\pm \mp 1}$ may differ, the individual nonzero off-diagonal entries in 'row' γ of $\mathbf{G}^{(\pm)}$ are not necessarily equal to the respective off-diagonal entries in 'column' γ of $\mathbf{G}^{(\mp)}$; however, that is of no relevance for the rest of our argument.

Returning to (10.5), we thus deduce by superposition that for the Fokker–Planck equation (3.23), with a tridiagonal coefficient matrix $\mathbf{A}$ as defined in Example 5.1, the infinite matrix $\mathbf{G} = \mathbf{G}^{(0)} + \mathbf{G}^{(+)} + \mathbf{G}^{(-)}$ is such that in each 'row' with index $\gamma \in \Gamma \setminus \{0\}$ the matrix $\mathbf{G}$ has nonzero off-diagonal entries in 'columns' $\gamma' \in \mathcal{Z}_+(\gamma) \cup \mathcal{Z}_-(\gamma)$ only. Thanks to the symmetry of $\mathbf{G}$ the number of nonzero off-diagonal entries in each 'row' γ of $\mathbf{G}$ is equal to the number of nonzero off-diagonal entries in 'column' γ of $\mathbf{G}$. The diagonal entry of $\mathbf{G}$ in each 'row' $\gamma \in \Gamma$ is equal to $2\delta_{\gamma\gamma} + \delta_{\gamma\gamma} + \delta_{\gamma\gamma} = 4$.

We now refer to Definition 8.8 and verify condition (8.3): i.e., we wish to show that for $\bar{s} > 0$ we have that

$$c_{\mathbf{G},\bar{s}} := \sup_{N \in \mathbb{N}} N \|\mathbf{G} - \mathbf{G}^{[N]}\|^{1/\bar{s}}_{\ell^2(\Gamma)\to\ell^2(\Gamma)},$$

where $(\mathbf{G}^{[N]})_{N=1}^\infty$ is a sequence of infinite matrices, which we shall define below.

We consider, for each $N \geq 1$, the matrix $\mathbf{G}^{[N]} := \mathbf{G}^{(0)} + \mathbf{G}^{(+),[N]} + \mathbf{G}^{(-),[N]}$, $N \geq 1$, where the matrices $\mathbf{G}^{(+),[N]}$ and $\mathbf{G}^{(-),[N]}$ are defined as follows. For $N = 1$, we choose $\mathbf{G}^{(+),[N]}$ and $\mathbf{G}^{(-),[N]}$ to be diagonal, with the same diagonal entries as $\mathbf{G}^{(+)}$ and $\mathbf{G}^{(-)}$, respectively. For $N > 1$, we again define the diagonal parts of the matrices $\mathbf{G}^{(\pm),[N]}$ to be the same as those of $\mathbf{G}^{(+)}$ and $\mathbf{G}^{(-)}$. The off-diagonal entries of $\mathbf{G}^{(+),[N]}$ and $\mathbf{G}^{(-),[N]}$ are defined, 'row'-wise, as follows.

Let $0 < p < 2$ and define $c_p := \left[2^{2p/(2-p)}\right] + 1$; for an integer $N > 1$ we shall then use the notation $N_p := c_p N$, and for a given sequence $\epsilon \in \ell^2(\mathbb{N})$ in the definition (5.3) of the infinite tridiagonal matrix $\mathbf{A}$ appearing in the bilinear form $\mathfrak{a}(\cdot,\cdot)$, we shall denote by $\epsilon^{[N_p]}$ the best N_p-term approximation of ϵ in $\ell^2(\mathbb{N})$. For $N > 1$, we define $\mathbf{G}^{(+),[N]}$ to contain, in the off-diagonal of the 'row' associated with an index $\gamma \in \Gamma \setminus \{0\}$, at most N_p nonzero elements $G^{(+)}_{\gamma\gamma'}$, whenever $\gamma' \in \mathcal{Z}_+(\gamma)$ with associated index $i_+ := i_+(\gamma, \gamma')$ is such that $i_+ \in \{j : \epsilon_j^{[N_p]} \neq 0\}$. Otherwise, the off-diagonal entry $G^{(+),[N]}_{\gamma\gamma'}$ is defined to be 0. For $N > 1$ the off-diagonal entries of the matrix $\mathbf{G}^{(-),[N]}$ are defined completely analogously to those of $\mathbf{G}^{(+),[N]}$.

Having thus defined the infinite matrices $\mathbf{G}^{(\pm),[N]}$, $N = 1, 2, \ldots$, and noting that by the triangle inequality,

$$\|\mathbf{G} - \mathbf{G}^{[N]}\|_{\ell^2(\Gamma)\to\ell^2(\Gamma)} \leq \|\mathbf{G}^{(+)} - \mathbf{G}^{(+),[N]}\|_{\ell^2(\Gamma)\to\ell^2(\Gamma)} + \|\mathbf{G}^{(-)} - \mathbf{G}^{(-),[N]}\|_{\ell^2(\Gamma)\to\ell^2(\Gamma)}, \tag{10.9}$$

it remains to bound the norms of $\mathbf{G}^{(\pm)} - \mathbf{G}^{(\pm),[N]}$. We shall use Schur's Lemma (cf. [23], p.6, ¶2 and [6], p.449, Theorem B) for this purpose. Thanks to the 'row'-wise definition of $\mathbf{G}^{(\pm),[N]}$, we already know the elements contained in each 'row' of $\mathbf{G}^{(\pm)} - \mathbf{G}^{(\pm),[N]}$; in particular, the diagonal entries of these matrices are equal to 0. Since the infinite matrices $\mathbf{G}^{(\pm)} - \mathbf{G}^{(\pm),[N]}$ are nonsymmetric, we also need to understand the structure of their 'columns' in order to be able to apply Schur's Lemma. As we already know from (10.8) what the nonzero off-diagonal entries in the columns of $\mathbf{G}^{(\pm)}$ are, it remains to consider the off-diagonal column entries of $\mathbf{G}^{(\pm),[N]}$. We shall in fact consider $\mathbf{G}^{(+),[N]}$ only, as an identical argument applies in the case of $\mathbf{G}^{(-),[N]}$.

Thanks to (10.8) the number of nonzero off-diagonal entries of $\mathbf{G}^{(+),[N]}$ in 'column' $\gamma \in \Gamma \setminus \{0\}$ is equal to the number of nonzero off-diagonal entries in 'row' γ of $\mathbf{G}^{(-),[N]}$; let us denote this number by $J_+(\gamma)$. Hence, column γ of $\mathbf{G}^{(+),[N]}$ contains at most $J_+(\gamma)$ nonzero off-diagonal entries, given by the formula (10.8) whenever $\epsilon_{j_-} = \epsilon_{i_+-1} \neq 0$ is present in the best N_p-term approximation of ϵ; every other off-diagonal entry in column γ of $\mathbf{G}^{(+),[N]}$ is equal to zero, either because the particular entry was already equal to zero in the matrix $\mathbf{G}^{(+)}$, or because the entry was nonzero in $\mathbf{G}^{(+)}$ but was set to zero while sweeping through the rows in the course of defining the off-diagonal entries of $\mathbf{G}^{(-),[N]}$ 'row'-wise.

Since, for all $N \in \mathbb{N}$ and every $0 < p < 2$,

$$\|\epsilon - \epsilon^{[N_p]}\|_{\ell^2(\mathbb{N})} \leq N_p^{-(1/p-1/2)} \|\epsilon\|_{\ell^p(\mathbb{N})} \leq \left(2^{\frac{2p}{2-p}} N\right)^{-(1/p-1/2)} \|\epsilon\|_{\ell^p(\mathbb{N})} = \tfrac{1}{2} N^{-(1/p-1/2)} \|\epsilon\|_{\ell^p(\mathbb{N})},$$

it then follows by Schur's Lemma (cf. [23], p.6, ¶2 and [6], p.449, Theorem B) that

$$\forall N \in \mathbb{N}: \quad \|\mathbf{G}^{(\pm)} - \mathbf{G}^{(\pm),[N]}\|_{\ell^2(\Gamma)\to\ell^2(\Gamma)} \leq \|\epsilon - \epsilon^{[N_p]}\|_{\ell^2(\mathbb{N})} \leq \tfrac{1}{2} N^{-(1/p-1/2)} \|\epsilon\|_{\ell^p(\mathbb{N})},$$

and, therefore, by the triangle inequality (10.9) we have that

$$\forall N \in \mathbb{N}: \quad \|\mathbf{G} - \mathbf{G}^{[N]}\|_{\ell^2(\Gamma)\to\ell^2(\Gamma)} \leq N^{-(1/p-1/2)} \|\epsilon\|_{\ell^p(\mathbb{N})},$$

from which we deduce that

$$c_{\mathbf{G},\bar{s}} = \sup_{N\in\mathbb{N}} N \|\mathbf{G} - \mathbf{G}^{[N]}\|_{\ell^2(\Gamma)}^{1/\bar{s}} \leq \sup_{N\in\mathbb{N}} N^{1-(1/p-1/2)/\bar{s}} \|\epsilon\|_{\ell^p(\mathbb{N})}^{1/\bar{s}} \leq \|\epsilon\|_{\ell^p(\mathbb{N})}^{1/\bar{s}} < \infty,$$

provided that $\epsilon \in \ell^p(\mathbb{N})$ with $0 < p < 2$ and that $\bar{s}$ is chosen as

$$0 < \bar{s}(p) := 1/p - 1/2. \tag{10.10}$$

Referring to the definition of s^*-computability (cf. Definition 8.8), we infer that $\mathbf{G}$ is s^* computable with any $0 < s^* \leq \bar{s}(p)$ if the sequence ϵ in Example 5.1 belongs to $\ell^p(\mathbb{N})$ with some $0 < p < 2$, resp. with $s^* = 1/p - 1/2$ (this encompasses the previous case of $A_{ij} = \delta_{ij}$, $i, j = 1, 2, \ldots$, if $p = 0$ is understood to indicate that ϵ is the zero sequence).

11. Fokker–Planck equations with drift

11.1. **Bounded invertibility of B and B^* for nonsymmetric $\mathfrak{a}(\cdot,\cdot)$.** So far, we have assumed that the "spatial" differential operator is defined by the symmetric bilinear form $\mathfrak{a}(\cdot,\cdot)$. This is always possible when the vector function appearing in the coefficient of the drift term in the Fokker–Planck equation is the gradient of a spring potential, by introducing Maxwellian-weighted spaces $\mathrm{L}^2_M(D)$ and $\mathrm{H}^1_M(D)$ and by rescaling the probability density function by the Maxwellian; cf. (3.26). In cases when the drift term cannot be removed by transformation to Maxwellian-weighted operators, however, nonsymmetric bilinear forms $\mathfrak{a}(\cdot,\cdot)$ must be considered. The bounded invertibility of the operator B corresponding to the bilinear form $\mathfrak{B}^*(\cdot,\cdot)$ in (5.6) must then be considered separately. We shall now prove existence of weak solutions for the Fokker-Planck equation with an Ornstein–Uhlenbeck type drift term. To this end, we verify the conditions for the abstract existence result Theorem 2.5. To this end, we consider in detail the structure of the drift term. Let μ be the countable product of Gaussian measures μ_k, $k \geq 1$, on $\mathbb{R}^d$ with trace class covariance operator Q. The new contribution added to the symmetric bilinear form $\mathfrak{a}(\cdot,\cdot)$ that was under consideration in previous sections is assumed, for a sequence $\sigma = (\sigma_k)_{k=1}^{\infty} \in [\ell^\infty(\mathbb{N})]^{d\times d}$, to be the bilinear form $\mathfrak{d}(\cdot,\cdot)$ defined by

$$\mathfrak{d}(\widehat{\psi},\widehat{\varphi}) = -\sum_{k\geq 1}\Big(\sigma_k q_k \widehat{\psi}, \mathrm{D}_k\widehat{\varphi}\Big)_{\mathrm{L}^2(H,\mu)}.$$

Proposition 11.1. *Assume that $\sigma \in [\ell^\infty(\mathbb{N})]^{d\times d}$ and that the covariance operator Q of the Gaussian measure μ on H is trace-class. Then, $\mathfrak{d}(\cdot,\cdot) : \mathrm{W}^{1,2}(H,\mu)\times \mathrm{W}^{1,2}(H,\mu) \to \mathbb{R}$ is continuous and*

$$|\mathfrak{d}(\widehat{\psi},\widehat{\varphi})| \leq \|\sigma\|_{[\ell^\infty(\mathbb{N})]^{d\times d}}\left(2\operatorname{Tr} Q\int_H |\widehat{\psi}(q)|^2\mu(\,\mathrm{d}q) + 4\|Q\|^2\int_H |\mathrm{D}\widehat{\psi}(q)|^2\mu(\,\mathrm{d}q)\right)^{1/2}\|\mathrm{D}\widehat{\varphi}\|_{\mathrm{L}^2(H,\mu)}. \tag{11.11}$$

Proof. We write

$$\begin{aligned}|\mathfrak{d}(\widehat{\psi},\widehat{\varphi})| &\leq \|\sigma\|_{[\ell^\infty(\mathbb{N})]^{d\times d}}\left(\sum_{k\geq 1}\Big\||q_k|\,\widehat{\psi}\Big\|^2_{\mathrm{L}^2(H,\mu)}\right)^{1/2}\left(\sum_{k'\geq 1}\|\mathrm{D}_{k'}\widehat{\varphi}\|^2_{\mathrm{L}^2(H,\mu)}\right)^{1/2}\\ &= \|\sigma\|_{[\ell^\infty(\mathbb{N})]^{d\times d}}\left(\sum_{k\geq 1}\Big\||q_k|\,\widehat{\psi}\Big\|^2_{\mathrm{L}^2(H,\mu)}\right)^{1/2}\|\mathrm{D}\widehat{\varphi}\|_{\mathrm{L}^2(H,\mu)}.\end{aligned}$$

Using [12, Prop. 9.2.10] and the assumption that $\widehat{\psi}\in \mathrm{W}^{1,2}(H,\mu)$, we deduce that, for a Gaussian measure μ with trace-class covariance Q,

$$\int_H \|q\|^2|\widehat{\psi}(q)|^2\mu(\mathrm{d}q) \leq 2\operatorname{Tr} Q\int_H |\widehat{\psi}(q)|^2\mu(\mathrm{d}q) + 4\|Q\|^2\int_H |\mathrm{D}\widehat{\psi}(q)|^2\mu(\mathrm{d}q).$$

This then yields the desired inequality (11.11). □

The bound (11.11) implies a Garding inequality for the Fokker–Planck operator with drift.

Proposition 11.2. *Assume that the covariance operator Q of the Gaussian measure μ is trace-class, and that the infinite coefficient matrix A satisfies* (5.1), (5.2). *Then, the bilinear form*

$$\mathfrak{a}(\cdot,\cdot) + \mathfrak{d}(\cdot,\cdot) : \mathcal{V}\times\mathcal{V}\to\mathbb{R}$$

is continuous (i.e. (2.21) *holds) and satisfies the Garding inequality* (2.22) *in the triple $\mathcal{V}\subset\mathcal{H}\simeq\mathcal{H}^*\subset\mathcal{V}^*$. In particular, the space-time variational formulation* (5.5) *with the spatial bilinear form $\mathfrak{a}(\cdot,\cdot)+\mathfrak{d}(\cdot,\cdot)$ in place of $\mathfrak{a}(\cdot,\cdot)$ in* (5.6) *is well-posed and its bilinear form $\mathfrak{B}^*(\cdot,\cdot)$ induces a boundedly invertible operator $B^*\in\mathcal{L}(\mathcal{Y},(\mathcal{X}_{0,\{T\}})^*)$.*

Proof. The continuity of the bilinear form $\mathfrak{a}(\cdot,\cdot)+\mathfrak{d}(\cdot,\cdot)$ is evident from the previous proposition and the continuity of $\mathfrak{a}(\cdot,\cdot)$. The Garding inequality (2.22) follows from the coercivity of $\mathfrak{a}(\cdot,\cdot)$ on $\mathcal{V}\times\mathcal{V}$ and from the continuity estimate (11.11) using a Cauchy inequality. The bounded invertibility of $B^*\in\mathcal{L}(\mathcal{Y},(\mathcal{X}_{0,\{T\}})^*)$ follows from Theorem 2.5. □

11.2. **s^*-computability of $\int_0^T \mathfrak{d}(\Theta\otimes[\Upsilon]_{\mathcal{V}},\Theta\otimes[\Upsilon]_{\mathcal{V}})\,\mathrm{d}t$ and of its adjoint.** As in Section 10.4, thanks to the independence of the sequence $\sigma\in[\ell^\infty(\mathbb{N})]^{d\times d}$ of t, we have that

$$\int_0^T \mathfrak{d}(\Theta\otimes[\Upsilon]_{\mathcal{V}},\Theta\otimes[\Upsilon]_{\mathcal{V}})\,\mathrm{d}t = (\Theta,\Theta)_{\mathrm{L}^2(0,T)}\otimes\mathfrak{d}([\Upsilon]_{\mathcal{V}},[\Upsilon]_{\mathcal{V}}).$$

As in the discussion of (10.4), the sparsity of the factor $(\Theta,\Theta)_{\mathrm{L}^2(0,T)}$ is considered in [24, Sec. 8.5]. Sparsity and s^*-computability of $\mathfrak{d}(\cdot,\cdot)$ are therefore determined by that of the infinite *drift matrix* $\mathbf{D}=\mathbf{D}[\sigma]$ defined via $\mathfrak{d}([\Upsilon]_{\mathcal{V}},[\Upsilon]_{\mathcal{V}}) = (D_{\gamma,\gamma'})_{\gamma,\gamma\in\Gamma}$. With the Hermite polynomial Riesz basis $[\Upsilon]_{\mathcal{V}}$ of $\mathrm{L}^2(H,\mu)$, the matrix entries $D_{\gamma\gamma'} = D_{\gamma\gamma'}[\sigma]$ are (assuming, once again, for the sake of simplicity of the exposition that $d=1$; abbreviating D_{q_k} as D_k to simplify the notation; and recalling the definition (10.1) of D^λ_γ):

$$D_{\gamma\gamma'}[\sigma] = \sum_{k\geq 1}\sigma_k(D^\lambda_\gamma)^{-1}(D^\lambda_{\gamma'})^{-1}\left(\mathrm{D}_k(q_kH_\gamma),H_{\gamma'}\right)_{\mathrm{L}^2(H,\mu)}. \tag{11.12}$$

For the ensuing calculations, we define

$$\theta_{\gamma\gamma'}(k) := (\mathrm{D}_k(q_k H_\gamma), H_{\gamma'})_{\mathrm{L}^2(H,\mu)}$$

and observe that

$$\theta_{\gamma\gamma'}(k) = (H_\gamma, H_{\gamma'})_{\mathrm{L}^2(H,\mu)} + (q_k \mathrm{D}_k H_\gamma, H_{\gamma'})_{\mathrm{L}^2(H,\mu)} = \delta_{\gamma\gamma'} + (q_k \mathrm{D}_k H_\gamma, H_{\gamma'})_{\mathrm{L}^2(H,\mu)}. \qquad (11.13)$$

In order to calculate the second term on the right-hand side of (11.13), we note that, by (10.2),

$$\mathrm{D}_k H_\gamma(q) = \begin{cases} \sqrt{\frac{\gamma_k}{\lambda_k}}\, H_{\gamma - e_k}(q) & \text{if} \quad \gamma_k \geq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $e_k \in \Gamma$ denotes the multi-index with entry 1 in position k and with zero entries in all other positions. Also, thanks to the three-term recurrence relation

$$q_k H_{\gamma_k}(q_k) = \sqrt{\gamma_k + 1}\, H_{\gamma_k+1}(q_k) + \sqrt{\gamma_k}\, H_{\gamma_k - 1}(q_k), \quad \gamma_k \in \mathbb{N} \cup \{0\},$$

for the univariate Hermite polynomials (4.3), with the notational convention that $H_m(q_k) \equiv 0$ when $m < 0$, we have that

$$q_k \mathrm{D}_k \left[H_{\gamma_k}\left(\frac{q_k}{\sqrt{\lambda_k}}\right)\right] = q_k \sqrt{\frac{\gamma_k}{\lambda_k}} H_{\gamma_k - 1}\left(\frac{q_k}{\sqrt{\lambda_k}}\right) = \gamma_k\, H_{\gamma_k}\left(\frac{q_k}{\sqrt{\lambda_k}}\right) + \sqrt{\gamma_k(\gamma_k - 1)}\, H_{\gamma_k - 2}\left(\frac{q_k}{\sqrt{\lambda_k}}\right)$$

for any $\gamma_k \geq 0$. Consequently, upon multiplying both sides of the last identity with $H_\gamma^{(k)}$, we deduce that

$$q_k \mathrm{D}_k H_\gamma = \gamma_k\, H_\gamma + \sqrt{\gamma_k(\gamma_k - 1)}\, H_{\gamma - 2e_k},$$

which, upon substitution into (11.13) and then inserting the resulting expression into (11.12) yields, for $\gamma, \gamma' \in \Gamma$, that

$$D_{\gamma\gamma'}[\sigma] \quad = \quad \sum_{k \geq 1} \sigma_k \left((1 + \gamma_k)\delta_{\gamma\gamma'} + \sqrt{\gamma_k(\gamma_k - 1)}\, \delta_{\gamma - 2e_k, \gamma'} \right) \langle \gamma, \lambda \rangle^{-1/2} \langle \gamma', \lambda \rangle^{-1/2}.$$

We observe that $\mathbf{D}[\sigma]$ depends linearly on the sequence σ. The verification of the s^*-computability now proceeds analogously to that of $\mathbf{A}[\epsilon]$.

We denote by $\sigma^{[N]}$ an N-term approximation of the sequence σ and observe that $\mathbf{D}^{[N]}[\sigma] := \mathbf{D}[\sigma^{[N]}]$ has at most $2N + 1$ nonzero entries in each 'row' with index $\gamma \in \Gamma$. To verify s^*-computability, by Definition 8.8 we must estimate

$$c_{\mathbf{D},\bar{s}} = \sup_{N \in \mathbb{N}} N \left\| \mathbf{D}[\sigma] - \mathbf{D}^{[N]}[\sigma] \right\|^{1/\bar{s}}_{\ell^2(\Gamma) \to \ell^2(\Gamma)} \leq \sup_{N \in \mathbb{N}} (2N + 1) \left\| \mathbf{D}[\sigma - \sigma^{[N]}] \right\|^{1/\bar{s}}_{\ell^2(\Gamma) \to \ell^2(\Gamma)}. \qquad (11.14)$$

For a given fixed $\gamma \in \Gamma$ and for any $\gamma' \in \Gamma$ and any $k \in \mathbb{N}$, we have that

$$\theta_{\gamma\gamma'}(k) := (1 + \gamma_k)\delta_{\gamma\gamma'} + \sqrt{\gamma_k(\gamma_k - 1)}\, \delta_{\gamma - 2e_k, \gamma'} = \begin{cases} 1 + \gamma_k & \text{if } \gamma' = \gamma, \\ \sqrt{\gamma_k(\gamma_k - 1)} & \text{if } \gamma' = \gamma - 2e_k, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, we can estimate

$$\begin{aligned} \|\mathbf{D}[\sigma] - \mathbf{D}^{[N]}[\sigma]\|^2_{\ell^2(\Gamma) \to \ell^2(\Gamma)} \quad &\leq \quad \sup_{\gamma \in \Gamma} \sum_{\gamma' \in \Gamma} \left| D_{\gamma\gamma'}[\sigma - \sigma^{[N]}] \right|^2 \\ &= \quad \sup_{\gamma \in \Gamma} \frac{1}{\langle \gamma, \lambda^{-1} \rangle} \sum_{\gamma' \in \Gamma} \frac{1}{\langle \gamma', \lambda^{-1} \rangle} \left| \sum_{k \geq 1} (\sigma_k - \sigma_k^{[N]})\, \theta_{\gamma\gamma'}(k) \right|^2. \end{aligned}$$

Fixing $\gamma \in \Gamma$ for now, we deduce that

$$\begin{aligned} \mathrm{S}_\gamma \quad &:= \quad \sum_{\gamma' \in \Gamma} \frac{1}{\langle \gamma', \lambda^{-1} \rangle} \left| \sum_{k \geq 1} (\sigma_k - \sigma_k^{[N]})\, \theta_{\gamma\gamma'}(k) \right|^2 \\ &\leq \quad \|\sigma - \sigma^{[N]}\|^2_{\ell^1(\mathbb{N})} \left\{ \sup_k \frac{(1 + \gamma_k)^2}{\langle \gamma, \lambda^{-1} \rangle} + \sup_k \frac{\gamma_k(\gamma_k - 1)}{\langle \gamma - 2e_k, \lambda^{-1} \rangle} \right\}. \end{aligned}$$

This implies that

$$\begin{aligned}\|\mathbf{D}-\mathbf{D}^{[N]}\|^2_{\ell^2(\Gamma)\to\ell^2(\Gamma)} &\le \sup_{\gamma\in\Gamma}\frac{\mathrm{S}_\gamma}{\langle\gamma,\lambda^{-1}\rangle}\\ &\le \|\sigma-\sigma^{[N]}\|^2_{\ell^1(\mathbb{N})}\sup_{\gamma\in\Gamma}\frac{1}{\langle\gamma,\lambda^{-1}\rangle}\left\{\sup_k\frac{(1+\gamma_k)^2}{\langle\gamma,\lambda^{-1}\rangle}+\sup_k\frac{\gamma_k(\gamma_k-1)}{\langle\gamma-2e_k,\lambda^{-1}\rangle}\right\}\\ &\le C\|Q\|^2\,\|\sigma-\sigma^{[N]}\|^2_{\ell^1(\mathbb{N})}.\end{aligned}$$

Inserting this into (11.14), we infer that for $\sigma\in\ell^p(\mathbb{N})$ with some $0<p<1$ the infinite drift matrix $\mathbf{D}[\sigma]$ is s^*-computable for any $0<s^*\le 1/p-1$.

12. Optimality

Based on Proposition 10.1 and on the definition (9.2) of the infinite matrix $\mathbf{B}^*$, we deduce from these observations our main result.

Theorem 12.1. *Consider the space-time variational formulation* (5.5) *of the infinite-dimensional Fokker–Planck equation with bilinear form* $\mathfrak{a}(\cdot,\cdot)$ *defined in* (5.8) *corresponding to the tridiagonal matrix* $\mathbf{A}$ *as in Example 5.1 based on the sequences* $\epsilon,\sigma\in\ell^p(\mathbb{N})$ *of elements* ϵ_i, σ_i *with* $0<p<1$.

Consider its representation $\mathbf{B}^*\mathbf{u}=\mathbf{f}^*$ *using a temporal wavelet basis* Θ *as in Section 10.1 and a spatial Hermite polynomial chaos basis* $\Upsilon=\{H_\gamma(q):\gamma\in\Gamma,\quad q\in D\}$. *Then, for any* $\varepsilon>0$, *the adaptive wavelet methods from* [10] *or* [9] *(and* [15]*) applied to the normal equation* (8.4) *with* $\mathbf{f}^*$ *and* $\mathbf{B}^*$ *as in (9.1) and (9.2), respectively, of the infinite matrix representation of the Fokker–Planck equation* (5.5) *in countably many dimensions produce an approximation* $\mathbf{u}_\varepsilon$ *with*

$$\|\widehat{\psi}-\mathbf{u}_\varepsilon^\top[\Theta\otimes\Upsilon]\|_{\mathcal{Y}}\eqsim\|\mathbf{u}-\mathbf{u}_\varepsilon\|\le\varepsilon.$$

Suppose that for some $0<s<\min\{d_t-1,\bar{s}(p)\}$ *we have that* $\mathbf{u}\in\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))$*; then,*

$$\operatorname{supp}\mathbf{u}_\varepsilon\lesssim\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))}.$$

The number of arithmetic operations and storage locations required by one call of the space-time adaptive solver with tolerance ε *is bounded by some multiple of* $\varepsilon^{-1/s}\|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(\ell^2(\nabla_{\mathcal{Y}}))}+1$.

The above assertions remain valid when in (3.23) *the dimension* K *is finite,* with the computability and admissibility constants $c_{\mathbf{B}^*,s}$ *and* $a_{\mathbf{B}^*,s}$ *independent of the dimension* K.

We remark in closing that for matrices $\mathbf{A}$ as in Example 5.2 this result remains valid, however now with admissibility constants depending on the block $\mathbf{A}_{11}$ in an unspecific way.

We also remark that the assumptions on the sequence σ could be slightly weakened, as the limit $1/p-1$ of s^*-computability of $\mathbf{D}[\sigma]$ is larger than the value $\bar{s}(p)$ found in (10.10).

References

1. J.W. Barrett and E. Süli, *Existence of global weak solutions to some regularized kinetic models of dilute polymers*, Multiscale Model. Simul. **6** (2007), 506–546.
2. ———, *Existence and equilibration of global weak solutions to kinetic models for dilute polymers I: Finitely extensible nonlinear bead-spring chains*, Math. Models Methods Appl. Sci. **21** (2011), no. 6, 1211–1289, Extended version available from `http://arxiv.org/abs/1004.1432`.
3. ———, *Existence and equilibration of global weak solutions to kinetic models for dilute polymers II: Hookean-type bead-spring chains*, Math. Models Methods Appl. Sci. **22** (2012), no. 5, 1–82, Extended version avialable from `http://arxiv.org/abs/1008.3052`.
4. W. Beckner, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc. **105** (1989), no. 2, 397–400. MR 954373 (89m:42027)
5. V. Bogachev, G. Da Prato, and M. Röckner, *Existence and uniqueness of solutions for Fokker-Planck equations on Hilbert spaces*, J. Evol. Equ. **10** (2010), no. 3, 487–509. MR 2674056 (2011g:60107)
6. D. Borwein and X. Gao, *Matrix operators on ℓ_p to ℓ_q*, Canad. Math. Bull. **37** (1994), no. 4, 448–446.
7. S. Cerrai, *Second Order PDE's in Finite and Infinite Dimension*, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001, A probabilistic approach. MR 1840644 (2002j:35327)
8. N. Chegini and R. Stevenson, *Adaptive wavelet schemes for parabolic problems: sparse matrices and numerical results*, SIAM J. Numer. Anal. **49** (2011), no. 1, 182–212. MR 2783222

9. A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations – Convergence rates*, Math. Comp **70** (2001), 27–75.
10. ———, *Adaptive wavelet methods II - Beyond the elliptic case*, Found. Comput. Math. **2** (2002), no. 3, 203–245.
11. G. Da Prato, P. Malliavin, and D. Nualart, *Compact families of Wiener functionals*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), no. 12, 1287–1291. MR 1194537 (94a:60085)
12. G. Da Prato and J. Zabczyk, *Second Order Partial Differential Equations in Hilbert Spaces*, London Mathematical Society Lecture Note Series, vol. 293, Cambridge University Press, Cambridge, 2002. MR 1985790 (2004e:47058)
13. T.J. Dijkema, C. Schwab, and R. Stevenson, *An adaptive wavelet method for solving high-dimensional elliptic PDEs*, Constr. Approx. **30** (2009), no. 3, 423–455. MR 2558688 (2010m:65270)
14. L. Figueroa and E. Süli, *Greedy approximation of high-dimensional Ornstein–Uhlenbeck operators with unbounded drift*, Tech. report, arXiv:1103.0726v1 [math.NA], `http://arxiv.org/abs/1103.0726` (Submitted), 2011.
15. T. Gantumur, H. Harbrecht, and R. Stevenson, *An optimal adaptive wavelet method without coarsening of the iterands*, Tech. Report 1325, Utrecht University, March 2005, To appear in *Math. Comp.*
16. M. Griebel and P. Oswald, *Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems*, Adv. Comput. Math. **4** (1995), no. 1–2, 171–206.
17. R. Guberović, C. Schwab, and R. Stevenson, *Space-time variational saddle point formulations of Stokes and Navier–Stokes equations*, Tech. report, Seminar for Applied Mathematics, ETH Zürich, `http://www.sam.math.ethz.ch/reports/2011` (Submitted), 2011.
18. David Knezevic and Endre Süli, *Spectral Galerkin approximation of Fokker–Planck equations with unbounded drift*, M2AN Math. Model. Numer. Anal. **43** (2009), no. 3, 445–485.
19. A. Kufner, O. John, and S. Fučík, *Function spaces*, Noordhoff International Publishing, Leyden, 1977, Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis. MR MR0482102 (58 #2189)
20. P.-A. Nitsche, *Best N term approximation spaces for tensor product wavelet bases*, Constr. Approx. **24** (2006), no. 1, 49–70. MR 2217525 (2007a:41032)
21. S. Peszat, *On a Sobolev space of functions of infinite number of variables*, Bull. Polish Acad. Sci. Math. **41** (1993), no. 1, 55–60. MR 1401886 (97m:46059)
22. A. Pietsch, *Nuclear Locally Convex Spaces*, Springer-Verlag, New York, 1972, Translated from the second German edition by William H. Ruckle, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 66. MR 0350360 (50 #2853)
23. I. Schur, *Bemerkungen zur Theorie der Beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. reine angew. Math. **140** (1911), 1–28.
24. C. Schwab and R. Stevenson, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp. **78** (2009), no. 267, 1293–1318. MR 2501051 (2011a:65347)