SIAM J. NUMER. ANAL.
Vol. 56, No. 2, pp. 877–905

CONVERGENCE OF SPARSE COLLOCATION FOR FUNCTIONS OF COUNTABLY MANY GAUSSIAN RANDOM VARIABLES (WITH APPLICATION TO ELLIPTIC PDEs)*

OLIVER G. ERNST†, BJÖRN SPRUNGK‡, AND LORENZO TAMELLINI§

Abstract. We give a convergence proof for the approximation by sparse collocation of Hilbert-space-valued functions depending on countably many Gaussian random variables. Such functions appear as solutions of elliptic PDEs with lognormal diffusion coefficients. We outline a general L^2-convergence theory based on previous work by Bachmayr et al. [*ESAIM Math. Model. Numer. Anal.*, 51 (2017), pp. 341–363] and Chen [*ESAIM Math. Model. Numer. Anal.*, in press, 2018, https://doi.org/10.1051/m2an/2018012] and establish an algebraic convergence rate for sufficiently smooth functions assuming a mild growth bound for the univariate hierarchical surpluses of the interpolation scheme applied to Hermite polynomials. We specifically verify for Gauss–Hermite nodes that this assumption holds and also show algebraic convergence with respect to the resulting number of sparse grid points for this case. Numerical experiments illustrate the dimension-independent convergence rate.

Key words. random PDEs, parametric PDEs, lognormal diffusion coefficient, best-N-term approximation, sparse grids, stochastic collocation, high-dimensional approximation, high-dimensional interpolation, Gauss–Hermite points

AMS subject classifications. 65D05, 65D15, 65C30, 60H25

DOI. 10.1137/17M1123079

1. Introduction.

The elliptic diffusion problem

$$-\nabla\cdot(a(\omega)\,\nabla u(\omega)) = f \quad \text{in } D\subset\mathbb{R}^d, \qquad u(\omega)=0 \text{ on } \partial D, \qquad \mathbf{P}\text{-a.s.}, \tag{1}$$

with a random diffusion coefficient $a:\Omega\to L^\infty(D)$ with respect to an underlying probability space $(\Omega,\mathcal{A},\mathbf{P})$ has become a model problem for numerical methods for solving random PDEs. For modeling reasons the diffusion field is often taken to have a lognormal probability law, which complicates both the study of the well-posedness of the problem [8, 25, 20, 31] as well as the analysis of approximation methods. One of the challenges is that the most common parametrization of a Gaussian random field—the *Karhunen–Loève expansion* [2, 23]—involves a countable number of standard normal random variables

$$\log a(x,\omega) = \phi_0(x) + \sum_{m=1}^{\infty} \phi_m(x)\,\xi_m(\omega), \tag{2}$$

where $\phi_0,\phi_m\in L^\infty(D)$ and $\xi_m\sim N(0,1)$ i.i.d. for $m\in\mathbb{N}$, leading to an elliptic PDE with a countably infinite number of random parameters $\boldsymbol{\xi}=(\xi_m)_{m\in\mathbb{N}}\in\mathbb{R}^{\mathbb{N}}$. In the following we shall also denote by $a(\boldsymbol{\xi})$ the random field a parametrized by $\boldsymbol{\xi}$ and likewise for random quantities depending on a such as the solution u of (1).

*Received by the editors March 28, 2017; accepted for publication (in revised form) January 23, 2018; published electronically April 10, 2018.
http://www.siam.org/journals/sinum/56-2/M112307.html

†Department of Mathematics, TU Chemnitz, Chemnitz 09126, Germany (oernst@math.tu-chemnitz.de).

‡Institute of Mathematics, University of Mannheim, Mannheim 68131, Germany (bjoern.sprungk@uni-mannheim.de).

§Consiglio Nazionale delle Ricerche - Istituto di Matematica Applicata e Tecnologie Informatiche E. Magenes (CNR-IMATI), Via Ferrata 5/a, 27100, Pavia, Italy (tamellini@imati.cnr.it).

Beside the stochastic Galerkin method [23, 29] the most common approach for approximating the solution $u(\boldsymbol{\xi})$ of such random or parametric elliptic PDEs is *polynomial collocation* . Early work on such methods for random PDEs considered a finite (if large) number of random parameters, a setting also referred to as *finite-dimensional noise* [44, 3, 37, 36]. In this case the parametric representation of $\log a$ is typically obtained by truncating a series expansion of the random field such as (2).

The analysis of the problem involving an infinite number of random variables was first discussed by Cohen, DeVore, and Schwab in [14, 15] in a simpler setting where the diffusion field a, rather than its logarithm as in (2), is expanded in a series. This results in an affine dependence of a on the random variables ξ_m, which are, moreover, assumed to have bounded support. In this framework the convergence of the best N-term approximation of the solution of the diffusion equation by Taylor as well as Legendre series was shown to be independent of the number of random variables; this result was further refined in the recent paper [5]. Employing the theoretical concepts stated in [14, 15], Chkifa, Cohen, and Schwab analyze in [11] collocation methods based on Lagrange interpolation at Leja points for problems with diffusion coefficients depending linearly on an infinite number of bounded random variables, which are adaptive in the polynomial degree as well as the number of active dimensions or random variables, respectively. The adaptive algorithm itself has its origins in the earlier work [22]. Each interpolatory approximation gives rise to a quadrature scheme, and in [39] Schillings and Schwab consider sparse adaptive quadrature schemes in the same setting of [11] in connection with approximating expectations with respect to posterior measures in Bayesian inference. Extensions to the case where the diffusion coefficient a depends nonlinearly on an infinite number of random variables with bounded support are discussed in [12].

Returning to the original lognormal diffusion problem, i.e., with a expanded as in (2) and depending on random variables with unbounded support, Hoang and Schwab [27] have obtained convergence results on best N-term approximation by Hermite polynomials. These were recently extended by Bachmayr et al. [4] using a different analytical approach employing a weighted ℓ^2-summability of the coefficients of the Hermite expansion of the solution and their relation to partial derivatives. The theoretical tools provided in [4] enabled the convergence analysis for adaptive sparse quadrature given in [9] employing, e.g., Gauss–Hermite nodes for Banach space-valued functions of countably many Gaussian random variables.

In this paper we address the convergence of sparse polynomial collocation for functions of a countable number of Gaussian random variables, such as the solution to the lognormal diffusion problem (1). Specifically, we follow the approach of [4] and [9] to prove an algebraic convergence rate with respect to the number of grid points for sparse collocation based on Gauss–Hermite interpolation nodes in the case of countably many variables. In particular, the result applies to the solution u of (1) when a is a lognormal random field. In addition, we highlight the common ideas surrounding sparse collocation found in the works referred to above. The convergence result in terms of the number of collocation points is obtained in two steps: we first link the error to the size of the multi-index set defining the sparse collocation and then derive a bound on the number of points contained in the associated sparse grid. This procedure has also been followed in all of the above-mentioned works analyzing the convergence of sparse grid quadrature and collocation schemes. An alternative strategy, which instead links the error directly to the number of collocation points by associating so-called *profits* with each component of a sparse grid, has been discussed in [34, 26], albeit only in the case of random variables with bounded support.

We remark that, beside the classical node families such as Gauss–Hermite and Genz–Keister [21] for quadrature and interpolation on $\mathbb{R}$ with respect to a Gaussian measure, Narayan and Jakeman [32] have introduced *weighted Leja points*—a generalization of the classical Leja point construction (see, e.g., [30, 17] and references therein) to unbounded domains and arbitrary weight functions. Moreover, they have proved that these node sets possess the correct asymptotic distribution of interpolation nodes and illustrate their computational potential in numerical experiments. Note that such weighted Leja points provide a nested sequence of interpolation nodes which can be extended one node at a time. The analysis of sparse collocation based on normal Leja points, i.e., weighted Leja points for a Gaussian measure, is an interesting topic for future research.

The remainder of this paper is organized as follows. In the next section we introduce the general setting and notation and construct the sparse grid collocation operator based on univariate Lagrange interpolation. Section 3 is devoted to the convergence analysis of such operators. We first outline in section 3.1 the general approaches to proving algebraic convergence rates as they can be found in the work mentioned above. Then, in section 3.2 we follow the approach of [4, 9] and derive sufficient conditions on the underlying univariate interpolation nodes in order to obtain such rates when approximating "countably variate" functions of sufficient smoothness. Finally, in section 3.3 we verify these conditions specifically for Gauss–Hermite nodes, provide bounds for the number of nodes in the resulting sparse grids, and state a convergence result with respect to this number. Section 4 returns to our original motivation and comments on the application to random elliptic PDEs before we verify our theoretical findings in section 5 for a simple boundary value problem in one spatial dimension. We draw final conclusions in section 6.

2. Setting and sparse collocation. We consider functions f defined on a parameter domain $\Gamma \subset \mathbb{R}^{\mathbb{N}}$ taking values in a separable real Hilbert space $\mathcal{H}$ with inner product $(\cdot,\cdot)_{\mathcal{H}}$ and norm $\|\cdot\|_{\mathcal{H}}$. As our interest lies in the approximation of the dependence of $f : \Gamma \to \mathcal{H}$ on $\boldsymbol{\xi} \in \Gamma$ by multivariate polynomials based on Lagrange interpolation, a minimal requirement is that point evaluation of f at any $\boldsymbol{\xi} \in \Gamma$ be well defined. Stronger smoothness requirements on f become necessary when deriving convergence rate estimates for the approximations.

Denoting by $\mathcal{B}(\mathbb{R})$ the σ-algebra of Borel sets on $\mathbb{R}$, we introduce a probability measure μ on the measurable space $(\mathbb{R}^{\mathbb{N}}, \otimes_{m\geq 1}\mathcal{B}(\mathbb{R}))$ as the countable product measure of standard Gaussian measures on $\mathbb{R}$, i.e.,

$$\mu = \bigotimes_{m\geq 1} N(0,1), \tag{3}$$

and denote by $L^2_\mu(\Gamma;\mathcal{H})$ the space of all (equivalence classes of) functions with finite second moments with respect to μ in the sense that

$$\int_{\mathbb{R}^{\mathbb{N}}} \|f(\boldsymbol{\xi})\|^2_{\mathcal{H}}\, \mu(\mathrm{d}\boldsymbol{\xi}) < \infty.$$

$L^2_\mu(\Gamma;\mathcal{H})$ forms a Hilbert space with inner product

$$(f,g)_{L^2_\mu} = \int_{\mathbb{R}^{\mathbb{N}}} (f(\boldsymbol{\xi}), g(\boldsymbol{\xi}))_{\mathcal{H}}\, \mu(\mathrm{d}\boldsymbol{\xi}).$$

In all of the following we make the standing assumption.

ASSUMPTION 2.1. *Let $f : \Gamma \to \mathcal{H}$ where $\mu(\Gamma) = 1$. It holds (for a measurable extension of f to $\mathbb{R}^{\mathbb{N}}$) that $f \in L^2_\mu(\mathbb{R}^{\mathbb{N}}; \mathcal{H})$.*

It is shown, e.g., in [40, Theorem 2.5], that the countable tensor products of Hermite polynomials form an orthonormal basis of $L^2_\mu(\mathbb{R}^{\mathbb{N}}; \mathcal{H})$. Under Assumption 2.1 we, therefore, have

$$f(\boldsymbol{\xi}) = \sum_{\boldsymbol{\nu}\in\mathcal{F}} f_{\boldsymbol{\nu}}\, H_{\boldsymbol{\nu}}(\boldsymbol{\xi}), \qquad f_{\boldsymbol{\nu}} := \int_{\mathbb{R}^{\mathbb{N}}} f(\boldsymbol{\xi})\, H_{\boldsymbol{\nu}}(\boldsymbol{\xi})\, \mu(\mathrm{d}\boldsymbol{\xi}) \in \mathcal{H}, \tag{4}$$

where $H_{\boldsymbol{\nu}}(\boldsymbol{\xi}) = \prod_{m\geq 1} H_{\nu_m}(\xi_m)$ and H_{ν_m} denotes the univariate Hermite orthonormal polynomial of degree ν_m while the multi-index set

$$\mathcal{F} := \left\{\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} : |\boldsymbol{\nu}|_0 < \infty\right\}, \qquad |\boldsymbol{\nu}|_0 := |\{j \in \mathbb{N} : \nu_j > 0\}|$$

contains all sequences of nonnegative integers with finitely many nonzero elements.

2.1. Sparse polynomial collocation. The construction of sparse collocation operators below is based on sequences of univariate Lagrange interpolation operators U_k mapping into the set $\mathcal{P}_k$ of univariate polynomials of degree at most $k \in \mathbb{N}_0$. Thus,

$$(U_k f)(\xi) = \sum_{i=0}^{k} f(\xi_i^{(k)})\, L_i^{(k)}(\xi), \qquad f\colon \mathbb{R} \to \mathbb{R},$$

where $\{L_i^{(k)}\}_{i=0}^k$ denotes the Lagrange fundamental polynomials of degree k associated with a set of $k+1$ distinct interpolation nodes $\Xi^{(k)} := \{\xi_0^{(k)}, \xi_1^{(k)}, \ldots, \xi_k^{(k)}\}$.

Remark 2.2. It may also be of interest to consider sequences of interpolation operators U_k with a more general degree of polynomial exactness $n(k)$ where $n : \mathbb{N}_0 \to \mathbb{N}_0$ is nondecreasing and $n(0) = 0$; see, for instance, [44, 3, 37, 36, 35, 34]. However, we restrict ourselves to $n(k) = k$ for simplicity.

We also introduce the *detail operators*

$$\Delta_k := U_k - U_{k-1}, \quad k \geq 0,$$

where we set $U_{-1} := \equiv 0$, and observe that

$$U_k = U_{k-1} + \Delta_k = \Delta_0 + \Delta_1 + \cdots + \Delta_k\ .$$

Tensorization. For any multi-index $\boldsymbol{k} = (k_m)_{m\in\mathbb{N}} \in \mathcal{F}$ the *(full) tensor product interpolation operator* $U_{\boldsymbol{k}} := \bigotimes_{m\in\mathbb{N}} U_{k_m}$ is defined by

$$(U_{\boldsymbol{k}} f)(\boldsymbol{\xi}) = \left(\bigotimes_{m\in\mathbb{N}} U_{k_m} f\right)(\boldsymbol{\xi}) = \sum_{\boldsymbol{i}\leq\boldsymbol{k}} f(\boldsymbol{\xi}_{\boldsymbol{i}}^{(\boldsymbol{k})}) L_{\boldsymbol{i}}^{(\boldsymbol{k})}(\boldsymbol{\xi}), \qquad f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}, \tag{5}$$

where $\boldsymbol{\xi}_{\boldsymbol{i}}^{(\boldsymbol{k})} \in \mathbb{R}^{\mathbb{N}}$ ranges over all points in the Cartesian product

$$\Xi^{(\boldsymbol{k})} := \bigtimes_{m\in\mathbb{N}} \Xi^{(k_m)}, \quad \text{where} \quad |\Xi^{(\boldsymbol{k})}| = \prod_{m\in\mathbb{N}} (1 + k_m), \tag{6}$$

and where

$$L_{\boldsymbol{i}}^{(\boldsymbol{k})}(\boldsymbol{\xi}) := \prod_{m\in\mathbb{N}} L_{i_m}^{(k_m)}(\xi_m) \tag{7}$$

is a multivariate polynomial of (total) degree $|\boldsymbol{k}|_1 = \sum_m k_m$. Note that $L_0^{(0)}(\xi) \equiv 1$; in particular, since $\boldsymbol{k} \in \mathcal{F}$ all but a finite number of factors in (6) and (7) are equal to one so that the corresponding products can be regarded as finite. The tensor product interpolation operator $U_{\boldsymbol{k}}$ maps into the multivariate (tensor product) polynomial space

$$\mathcal{Q}_{\boldsymbol{k}} := \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{i}} : 0 \le i_m \le k_m, m \in \mathbb{N}\}, \qquad \boldsymbol{k} \in \mathcal{F}. \tag{8}$$

Note that since both the univariate polynomial sets of Lagrange fundamental polynomials $\{L_i^{(k)}\}_{i=0}^k$ and the Hermite orthonormal polynomials $\{H_i\}_{i=0}^k$ form a basis of $\mathcal{P}_k$, equivalent characterizations are

$$\begin{aligned}\mathcal{Q}_{\boldsymbol{k}} &= \operatorname{span}\{L_{\boldsymbol{i}}^{(\boldsymbol{k})} : 0 \le i_m \le k_m, m \in \mathbb{N}\} \\ &= \operatorname{span}\{H_{\boldsymbol{i}} : 0 \le i_m \le k_m, m \in \mathbb{N}\}, \qquad \boldsymbol{k} \in \mathcal{F}.\end{aligned}$$

In order for the tensor product interpolation operator $U_{\boldsymbol{k}}$ to be applicable also to functions defined only on a subset $\Gamma \subset \mathbb{R}^{\mathbb{N}}$, we assume the interpolation nodes to all lie in Γ.

ASSUMPTION 2.3. *Let $\Gamma \subset \mathbb{R}^{\mathbb{N}}$ denote the domain from Assumption* 2.1. *For all $\boldsymbol{k} \in \mathcal{F}$ the Cartesian products of nodal sets $\Xi^{(\boldsymbol{k})}$ given in* (6) *satisfy $\Xi^{(\boldsymbol{k})} \subset \Gamma$.*

In the following we denote by $\mathbb{R}^\Gamma$ the set of all mappings from Γ to $\mathbb{R}$. Analogous to (5) we define for any multi-index $\boldsymbol{k} \in \mathcal{F}$ the tensorized detail operator

$$\Delta_{\boldsymbol{k}} := \bigotimes_{m \in \mathbb{N}} \Delta_{k_m} : \mathbb{R}^\Gamma \to \mathcal{Q}_{\boldsymbol{k}}.$$

Finally, we associate with a finite subset $\Lambda \subset \mathcal{F}$ the multivariate polynomial space

$$\mathcal{P}_\Lambda := \sum_{\boldsymbol{i} \in \Lambda} \mathcal{Q}_{\boldsymbol{i}} \tag{9}$$

and define the associated *sparse (polynomial) collocation operator* $U_\Lambda : \mathbb{R}^\Gamma \to \mathcal{P}_\Lambda$ by

$$U_\Lambda := \sum_{\boldsymbol{i} \in \Lambda} \Delta_{\boldsymbol{i}}. \tag{10}$$

We will see that U_Λ is exact on $\mathcal{P}_\Lambda$ under some natural assumptions on the multi-index set Λ, for which we first recall some basic definitions given in [13, 11, 12].

Partial orderings and monotone sets of multi-indices. We define a partial ordering on $\mathcal{F}$ by

$$\tilde{\boldsymbol{\nu}} \le \boldsymbol{\nu} \quad :\Leftrightarrow \quad \tilde{\nu}_m \le \nu_m \ \ \forall\, m \in \mathbb{N}$$

as well as

$$\tilde{\boldsymbol{\nu}} < \boldsymbol{\nu} \quad :\Leftrightarrow \quad \tilde{\boldsymbol{\nu}} \le \boldsymbol{\nu} \text{ and } \tilde{\nu}_m < \nu_m \text{ for at least one } m \in \mathbb{N}$$

and introduce the relation

$$\tilde{\boldsymbol{\nu}} \not\le \boldsymbol{\nu} \quad :\Leftrightarrow \quad \tilde{\nu}_m > \nu_m \text{ for at least one } m \in \mathbb{N}.$$

We shall call a set of multi-indices $\Lambda \subset \mathcal{F}$ *monotone* if $\boldsymbol{\nu} \in \Lambda$ and $\tilde{\boldsymbol{\nu}} \leq \boldsymbol{\nu}$ together imply that also $\tilde{\boldsymbol{\nu}} \in \Lambda$. For a multi-index $\boldsymbol{\nu} \in \mathcal{F}$ we define its *rectangular envelope* $\mathcal{R}_{\boldsymbol{\nu}}$ by

$$\mathcal{R}_{\boldsymbol{\nu}} := \{\tilde{\boldsymbol{\nu}} \in \mathcal{F} : \tilde{\boldsymbol{\nu}} \leq \boldsymbol{\nu}\}.$$

Note that $\mathcal{R}_{\boldsymbol{\nu}}$ for $\boldsymbol{\nu} \in \mathcal{F}$ is a finite (and monotone) set with cardinality

$$|\mathcal{R}_{\boldsymbol{\nu}}| = \prod_{m \in \mathbb{N}} (1 + \nu_m) < \infty. \tag{11}$$

Finally, for any $j \in \mathbb{N}$, we denote by $\boldsymbol{e}_j$ the jth standard unit sequence, i.e., the sequence whose elements are all 0 except for the jth, which is equal to 1.

2.2. Polynomial exactness of sparse collocation. The introduction of the rectangular envelope $\mathcal{R}_{\boldsymbol{\nu}}$ of a multi-index $\boldsymbol{\nu} \in \mathcal{F}$ permits a convenient characterization of monotone multi-index sets Λ and the associated polynomial spaces $\mathcal{P}_\Lambda$ introduced in (9).

PROPOSITION 2.4. *If $\Lambda \subset \mathcal{F}$ is monotone, then*

$$\Lambda = \bigcup_{\boldsymbol{\nu} \in \Lambda} \mathcal{R}_{\boldsymbol{\nu}} \quad \text{and} \quad \mathcal{P}_\Lambda = \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{\nu}} : \boldsymbol{\nu} \in \Lambda\} = \operatorname{span}\{H_{\boldsymbol{\nu}} : \boldsymbol{\nu} \in \Lambda\}.$$

Proof. Since $\boldsymbol{\nu} \in \mathcal{R}_{\boldsymbol{\nu}}$ for all $\boldsymbol{\nu} \in \Lambda$ the set on the left is obviously a subset of that on the right. Conversely, given $\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}$ for some $\boldsymbol{\nu} \in \Lambda$, the definition of $\mathcal{R}_{\boldsymbol{\nu}}$ implies $\boldsymbol{i} \leq \boldsymbol{\nu}$, which in turn implies $\boldsymbol{i} \in \Lambda$ by the monotonicity of Λ. Moreover, monotonicity also implies

$$\mathcal{P}_\Lambda = \sum_{\boldsymbol{k} \in \Lambda} \mathcal{Q}_{\boldsymbol{k}} = \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{i}} : \boldsymbol{i} \leq \boldsymbol{k}, \boldsymbol{k} \in \Lambda\} = \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{i}} : \boldsymbol{i} \in \Lambda\} = \operatorname{span}\{H_{\boldsymbol{\nu}} : \boldsymbol{\nu} \in \Lambda\},$$

where monotonicity is required for the last two equalities. ∎

In view of Proposition 2.4, $\mathcal{P}_\Lambda$ for a multi-index set $\Lambda \subset \mathcal{R}_{\boldsymbol{k}}$ represents a sparsification of $\mathcal{Q}_{\boldsymbol{k}}$, while the full tensor product polynomial space $\mathcal{Q}_{\boldsymbol{k}}$ coincides with $\mathcal{P}_\Lambda$ for $\Lambda = \mathcal{R}_{\boldsymbol{k}}$. Similarly, the full tensor approximation operator $U_{\boldsymbol{k}}$ defined in (5) can be expressed as $U_{\boldsymbol{k}} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{k}}} \Delta_{\boldsymbol{i}}$.

PROPOSITION 2.5. *Let $\Lambda \subset \mathcal{F}$ be a finite and monotone set. Then $U_\Lambda p = p$ for all $p \in \mathcal{P}_\Lambda$. In particular, for all $p \in \mathcal{P}_\Lambda$ we have $\Delta_{\boldsymbol{i}} p = 0$ for $\boldsymbol{i} \notin \Lambda$.*

Proof. Observe first that, for any $\boldsymbol{\nu}, \boldsymbol{i} \in \mathcal{F}$ such that $\boldsymbol{i} \not\leq \boldsymbol{\nu}$, we have

$$\Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}} = \prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} = \prod_{m \in \mathbb{N}} \underbrace{(U_{i_m} - U_{i_m - 1}) \xi_m^{\nu_m}}_{= \xi_m^{\nu_m} - \xi_m^{\nu_m} \equiv 0 \text{ for at least one } m} = 0.$$

It suffices to prove the assertions for all monomials $\boldsymbol{\xi}^{\boldsymbol{\nu}}$ in $\mathcal{P}_\Lambda$. For $\boldsymbol{\nu} \in \Lambda$ any $\boldsymbol{i} \in \mathcal{F} \setminus \Lambda$ must satisfy $\boldsymbol{i} \not\leq \boldsymbol{\nu}$ and, therefore, $\Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}} = 0$, proving the second assertion. We conclude that

$$U_\Lambda \boldsymbol{\xi}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{i} \in \Lambda} \Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{i} \in \Lambda \cap \mathcal{R}_{\boldsymbol{\nu}}} \Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}},$$

where the third equality follows from the fact that $\mathcal{R}_{\boldsymbol{\nu}} \subseteq \Lambda$ for all $\boldsymbol{\nu} \in \Lambda$ due to the monotonicity of Λ. The proof concludes with

$$U_\Lambda \boldsymbol{\xi}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \Delta_{\boldsymbol{i}} \boldsymbol{\xi}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \left(\prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} \right) = \prod_{m \in \mathbb{N}} \left(\sum_{i_m = 0}^{\nu_m} \Delta_{i_m} \xi_m^{\nu_m} \right) = \prod_{m \in \mathbb{N}} U_{\nu_m} \xi_m^{\nu_m}$$
$$= \prod_{m \in \mathbb{N}} \xi_m^{\nu_m} = \boldsymbol{\xi}^{\boldsymbol{\nu}}.$$

Note that the third equality is obtained by rewriting a (finite) product of sums: Since $\boldsymbol{\nu} \in \mathcal{F}$ there exists an $M \in \mathbb{N}$ such that $\nu_m = 0$ for $m > M$. For such m we have $\Delta_{i_m}^{\nu_m} \xi_m^{\nu_m} = \Delta_0 \xi_m^0 \equiv 1$ and, therefore,

$$\prod_{m \in \mathbb{N}} \left(\sum_{i_m = 0}^{\nu_m} \Delta_{i_m} \xi_m^{\nu_m} \right) = (\Delta_0 \xi_1^{\nu_1} + \cdots + \Delta_{\nu_1} \xi_1^{\nu_1}) \cdots (\Delta_0 \xi_M^{\nu_M} + \cdots + \Delta_{\nu_M} \xi_M^{\nu_M})$$
$$= \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^M \\ i_m \le \nu_m}} \Delta_{i_1} \xi_1^{\nu_1} \cdots \Delta_{i_M} \xi_M^{\nu_M} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \left(\prod_{m \in \mathbb{N}} \Delta_{i_m} \xi_m^{\nu_m} \right). \qquad \square$$

Proposition 2.5 can be seen as an extension of [6, Proposition 1] to general monotone multi-index sets as well as an extension of [13, Theorem 6.1] and [11, Theorem 2.1] to interpolation operators U_i with nonnested node sets. As mentioned in [13, p. 89], if the set Λ is not monotone, then U_Λ will not be exact on $\mathcal{P}_\Lambda$ in general. However, exactness on $\mathcal{P}_\Lambda$ is a crucial property in the subsequent convergence analysis, and we, therefore, choose to work exclusively with monotone sets Λ.

2.3. Sparse grid associated with U_Λ. The construction of $U_\Lambda f$ for $f \colon \Gamma \to \mathbb{R}$ consists of a linear combination of tensor product interpolation operators requiring the evaluation of f at certain multivariate nodes. We refer to the collection of these nodes as the *sparse grid* $\Xi_\Lambda \subset \Gamma$ associated with Λ. For a monotone and finite set $\Lambda \subset \mathcal{F}$ there holds

$$\Xi_\Lambda = \bigcup_{\boldsymbol{i} \in \Lambda} \Xi^{(\boldsymbol{i})}, \tag{12}$$

because for $\boldsymbol{i} \in \mathcal{F}$ we have

$$\Delta_{\boldsymbol{i}} f = \Big[\bigotimes_{m \ge 1} (U_{i_m} - U_{i_m - 1}) \Big] f = \sum_{\boldsymbol{i} - \boldsymbol{1} \le \boldsymbol{k} \le \boldsymbol{i}} (-1)^{|\boldsymbol{i} - \boldsymbol{k}|_1} \Big[\bigotimes_{m \ge 1} U_{k_m} \Big] f,$$

i.e., for computing $\Delta_{\boldsymbol{i}} f$ we need to evaluate f at

$$\Xi^{(\boldsymbol{i}), \Delta} := \bigcup_{\boldsymbol{i} - \boldsymbol{1} \le \boldsymbol{k} \le \boldsymbol{i}} \Xi^{(\boldsymbol{k})}.$$

Since Λ is a monotone set, the resulting sparse grid for $U_\Lambda = \sum_{\boldsymbol{i} \in \Lambda} \Delta_{\boldsymbol{i}}$ is

$$\Xi_\Lambda = \bigcup_{\boldsymbol{i} \in \Lambda} \Xi^{(\boldsymbol{i}), \Delta} = \bigcup_{\boldsymbol{i} \in \Lambda} \bigcup_{\boldsymbol{i} - \boldsymbol{1} \le \boldsymbol{k} \le \boldsymbol{i}} \Xi^{(\boldsymbol{k})} = \bigcup_{\boldsymbol{i} \in \Lambda} \Xi^{(\boldsymbol{i})}.$$

We remark that the unisolvence on $\mathcal{P}_\Lambda$ of point evaluations on Ξ_Λ is discussed in [13, Theorem 6.1].

3. Convergence analysis. In this section we analyze the error

$$\|f - U_\Lambda f\|_{L^2_\mu}, \qquad f\colon \Gamma \to \mathcal{H},$$

where $\|\cdot\|_{L^2_\mu}$ denotes the norm[1] in $L^2_\mu(\mathbb{R}^\mathbb{N};\mathcal{H})$, f is assumed to satisfy Assumption 2.1, and $\Lambda \subset \mathcal{F}$ is required to be monotone and finite. Our first goal here is to establish a convergence rate $s > 0$ for the error of $U_{\Lambda_N} f$ for a nested sequence Λ_N of monotone subsets of $\mathcal{F}$ with $|\Lambda_N| = N$. Specifically, we show that

$$\|f - U_{\Lambda_N} f\|_{L^2_\mu} \le CN^{-s}, \qquad f\colon \Gamma \to \mathcal{H}, \tag{13}$$

where $C < \infty$ may depend on f as well as the univariate nodal sets. The line of proof we present here follows and builds upon the works [9, 27, 4]. We complement this convergence rate with a bound on the number of collocation points associated with a given multi-index set.

3.1. General convergence results. The subsequent error analysis for the sparse collocation operator U_Λ is based on the representation of multivariate functions $f \in L^2_\mu(\mathbb{R}^\mathbb{N};\mathcal{H})$ in the orthonormal basis of multivariate Hermite polynomials $H_{\boldsymbol{\nu}}$. We shall, therefore, examine the worst-case approximation error of any U_Λ applied to a given multivariate Hermite polynomial $H_{\boldsymbol{\nu}}$. To this end we define

$$c_{\boldsymbol{\nu}} := \sup_{\Lambda\subset\mathcal{F},|\Lambda|<\infty} \|(I - U_\Lambda)H_{\boldsymbol{\nu}}\|_{L^2_\mu}, \qquad \boldsymbol{\nu} \in \mathcal{F}. \tag{14}$$

This quantity is finite since $\Delta_{\boldsymbol{i}} H_{\boldsymbol{\nu}} = 0$ for $\boldsymbol{i} \not\le \boldsymbol{\nu}$ and hence

$$c_{\boldsymbol{\nu}} = \max_{\Lambda\subseteq\mathcal{R}_{\boldsymbol{\nu}}} \|(I - U_\Lambda)H_{\boldsymbol{\nu}}\|_{L^2_\mu},$$

where now the maximum is taken over a finite set. The quantities $c_{\boldsymbol{\nu}}$ also measure the deviation of the error of the oblique projection U_Λ from that of orthogonal projection, as these numbers would all be zero or one if U_Λ were replaced by the L^2_μ-orthogonal projection onto $\mathcal{P}_\Lambda$. Moreover, we obtain the following bound.

PROPOSITION 3.1. *For all $\boldsymbol{\nu} \in \mathcal{F}$ the quantity $c_{\boldsymbol{\nu}}$ defined in* (14) *satisfies*

$$c_{\boldsymbol{\nu}} \le \sum_{\boldsymbol{i}\in\mathcal{R}_{\boldsymbol{\nu}}} \|\Delta_{\boldsymbol{i}} H_{\boldsymbol{\nu}}\|_{L^2_\mu}.$$

In particular, if there exist $\theta \ge 0$ and $K \ge 1$ such that for the univariate Hermite polynomials there holds

$$\|\Delta_i H_\nu\|_{L^2_\mu} \le (1 + K\nu)^\theta \qquad \forall\, i \in \mathbb{N}_0, \tag{15}$$

where we have again denoted the univariate Gaussian measure by μ, then

$$c_{\boldsymbol{\nu}} \le \prod_{m\in\mathbb{N}} (1 + K\nu_m)^{\theta+1}, \qquad \boldsymbol{\nu} \in \mathcal{F}. \tag{16}$$

[1] We shall occasionally also use $\|\cdot\|_{L^2_\mu}$ to denote the norm on $L^2_\mu(\mathbb{R};\mathbb{R})$, where μ is the standard normal distribution on the real line. The meaning should be clear from the context.

Proof. In view of Proposition 2.5 we have $H_{\boldsymbol{\nu}} = U_{\boldsymbol{\nu}} H_{\boldsymbol{\nu}} = \sum_{i \in \mathcal{R}_{\boldsymbol{\nu}}} \Delta_i H_{\boldsymbol{\nu}}$ and, particularly, $\Delta_{\boldsymbol{i}} H_{\boldsymbol{\nu}} = 0$ for $\boldsymbol{i} \notin \mathcal{R}_{\boldsymbol{\nu}}$, since $H_{\boldsymbol{\nu}} \in \mathcal{P}_{\mathcal{R}_{\boldsymbol{\nu}}}$. Therefore,

$$\begin{aligned}(I - U_\Lambda) H_{\boldsymbol{\nu}} &= \sum_{i \in \mathcal{R}_{\boldsymbol{\nu}}} \Delta_i H_{\boldsymbol{\nu}} - \sum_{i \in \Lambda} \Delta_i H_{\boldsymbol{\nu}} = \sum_{i \in \mathcal{R}_{\boldsymbol{\nu}}} \Delta_i H_{\boldsymbol{\nu}} - \sum_{i \in \Lambda \cap \mathcal{R}_{\boldsymbol{\nu}}} \Delta_i H_{\boldsymbol{\nu}} \\ &= \sum_{i \in \mathcal{R}_{\boldsymbol{\nu}} \setminus \Lambda} \Delta_i H_{\boldsymbol{\nu}},\end{aligned}$$

giving

$$c_{\boldsymbol{\nu}} = \max_{\Lambda \subseteq \mathcal{R}_{\boldsymbol{\nu}}} \|(I - U_\Lambda) H_{\boldsymbol{\nu}}\|_{L^2_\mu} \le \max_{\Lambda \subseteq \mathcal{R}_{\boldsymbol{\nu}}} \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}} \setminus \Lambda} \|\Delta_i H_{\boldsymbol{\nu}}\|_{L^2_\mu} \le \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \|\Delta_i H_{\boldsymbol{\nu}}\|_{L^2_\mu}.$$

Moreover, if (15) holds, then

$$\begin{aligned}c_{\boldsymbol{\nu}} &\le \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \|\Delta_{\boldsymbol{i}} H_{\boldsymbol{\nu}}\|_{L^2_\mu} = \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \prod_{m \in \mathbb{N}} \|\Delta_{i_m} H_{\nu_m}\|_{L^2_\mu} \le \sum_{\boldsymbol{i} \in \mathcal{R}_{\boldsymbol{\nu}}} \prod_{m \in \mathbb{N}} (1 + K\nu_m)^\theta \\ &= |\mathcal{R}_{\boldsymbol{\nu}}| \prod_{m \in \mathbb{N}} (1 + K\nu_m)^\theta \le \prod_{m \in \mathbb{N}} (1 + K\nu_m)^{\theta+1},\end{aligned}$$

where we have used (11) and $K \ge 1$ in the last inequality. ∎

Remark 3.2. Bounds such as (15) can often be found in the sparse collocation or sparse quadrature literature, e.g., for quadrature operators applied to Hermite polynomials [9], norms of quadrature operators on bounded domains [39], or Lebesgue constants for Leja points [12]. Numerical estimates for the specific case of Genz–Keister points are given in [7].

The following lemma provides a natural starting point for bounding the approximation error of $U_\Lambda f$ for monotone subsets Λ. The proof follows the same line of argument as the proof of [9, Lemma 3.2].

LEMMA 3.3 (cf. [9, Lemma 3.2]). *For a finite and monotone subset $\Lambda \subset \mathcal{F}$ there holds*

$$\|f - U_\Lambda f\|_{L^2_\mu} \le \sum_{\boldsymbol{\nu} \in \mathcal{F} \setminus \Lambda} c_{\boldsymbol{\nu}} \|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}. \tag{17}$$

Proof. Due to the monotonicity of Λ we can apply Proposition 2.5 and obtain

$$\begin{aligned}\|f - U_\Lambda f\|_{L^2_\mu} &= \left\| \sum_{\boldsymbol{\nu} \in \mathcal{F}} f_{\boldsymbol{\nu}} \, (I - U_\Lambda) H_{\boldsymbol{\nu}}(\boldsymbol{\xi}) \right\|_{L^2_\mu} = \left\| \sum_{\boldsymbol{\nu} \in \mathcal{F} \setminus \Lambda} f_{\boldsymbol{\nu}} \, (I - U_\Lambda) H_{\boldsymbol{\nu}}(\boldsymbol{\xi}) \right\|_{L^2_\mu} \\ &\le \sum_{\boldsymbol{\nu} \in \mathcal{F} \setminus \Lambda} \|f_{\boldsymbol{\nu}}\|_{\mathcal{H}} \|(I - U_\Lambda) H_{\boldsymbol{\nu}}\|_{L^2_\mu} \le \sum_{\boldsymbol{\nu} \in \mathcal{F} \setminus \Lambda} c_{\boldsymbol{\nu}} \|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}.\end{aligned}$$

∎

Building on Lemma 3.3 the approximation error $\|f - U_\Lambda f\|_{L^2_\mu}$ may be further bounded given summability properties for the sequence $(c_{\boldsymbol{\nu}} \|f_{\boldsymbol{\nu}}\|_{\mathcal{H}})_{\boldsymbol{\nu} \in \mathcal{F}}$. The key result here is known as *Stechkin's lemma* which provides a decay rate for the ℓ^q-tail of a p-summable sequence for $q > p$ and is due to Stechkin [41] for $q = 2$ (cf. also [13, Lemma 3.6]).

LEMMA 3.4 (Stechkin). *Let $0 < p < q < \infty$ and let*

$$(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu} \in \mathcal{F}} \in \ell^p(\mathcal{F}) := \left\{ (b_{\boldsymbol{\nu}})_{\boldsymbol{\nu} \in \mathcal{F}} \colon \sum_{\boldsymbol{\nu} \in \mathcal{F}} |b_{\boldsymbol{\nu}}|^p < \infty \right\}$$

be a sequence of nonnegative numbers. Then with Λ_N denoting the set of multi-indices $\boldsymbol{\nu}$ corresponding to the N largest elements $a_{\boldsymbol{\nu}}$, there holds

$$\left(\sum_{\boldsymbol{\nu}\notin\Lambda_N} a_{\boldsymbol{\nu}}^q\right)^{1/q} \leq \|(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\|_{\ell^p}(N+1)^{-s}, \qquad s=\frac{1}{p}-\frac{1}{q}\,. \tag{18}$$

The index sets Λ_N in Stechkin's lemma associated with the N largest sequence elements are not necessarily monotone and, therefore, Lemmas 3.3 and 3.4 cannot be combined to bound the error without additional assumptions. An obvious way to ensure monotonicity of the sets Λ_N in Stechkin's lemma is to assume the sequence $(a_{\boldsymbol{\nu}})$ to be *nonincreasing*, i.e.,

$$\boldsymbol{\nu}\leq\tilde{\boldsymbol{\nu}} \quad \Rightarrow \quad a_{\boldsymbol{\nu}}\geq a_{\tilde{\boldsymbol{\nu}}}.$$

This leads to the following theorem.

THEOREM 3.5. *Let Assumptions* 2.1 *and* 2.3 *be satisfied and let there exist a nonincreasing sequence $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\in\ell^p(\mathcal{F})$ with $p\in(0,1)$ such that*

$$c_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}\leq\hat{c}_{\boldsymbol{\nu}} \qquad \forall\boldsymbol{\nu}\in\mathcal{F}.$$

Then there exists a nested sequence $(\Lambda_N)_{N\in\mathbb{N}}$ of finite and monotone subsets $\Lambda_N\subset\mathcal{F}$ with $|\Lambda_N|=N$ such that (13) *holds with rate $s=1/p-1$.*

We provide a proof below. The convergence analysis in [12, 39] for sparse quadrature and interpolation in case of bounded Γ proceeds along the lines of Theorem 3.5, although sometimes this is hidden in the details. There the authors employ explicit bounds on the norms of the Legendre or Taylor coefficients of $f:\Gamma\to\mathcal{H}$ to construct a dominating and nonincreasing sequence $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\in\ell^p(\mathcal{F})$, $p\in(0,1)$.

In our setting, however, it is not always possible to derive explicit bounds on the norm of the Hermite coefficients $\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}$. In [4] a technique was developed which relies on somewhat implicit bounds on $\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}$ via a weighted ℓ^2-summability property. We adapt this approach to the current setting in the following theorem.

THEOREM 3.6. *Let Assumptions* 2.1 *and* 2.3 *be satisfied and let there exist a sequence $(b_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ of positive numbers such that*

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} b_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}^2<\infty \tag{19}$$

as well as another nonincreasing sequence $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\in\ell^p(\mathcal{F})$, $p\in(0,2)$, for which

$$\frac{c_{\boldsymbol{\nu}}}{b_{\boldsymbol{\nu}}^{1/2}}\leq\hat{c}_{\boldsymbol{\nu}} \qquad \forall\boldsymbol{\nu}\in\mathcal{F}.$$

Then there exists a nested sequence $(\Lambda_N)_{N\in\mathbb{N}}$ of finite and monotone subsets $\Lambda_N\subset\mathcal{F}$ with $|\Lambda_N|=N$ such that (13) *holds with rate $s=1/p-1/2$.*

Proof of Theorems 3.5 *and* 3.6. Let Λ_N be the set of multi-indices $\boldsymbol{\nu}$ corresponding to the N largest elements of $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$. Then each Λ_N is monotone and the sequence $(\Lambda_N)_{N\in\mathbb{N}}$ can be chosen to be nested.

If the assumptions of Theorem 3.5 hold, we can apply Lemma 3.3 and Stechkin's lemma with $q=1>p$ to obtain

$$\|f-U_{\Lambda_N}f\|_{L^2_\mu}\leq\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} c_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}\leq\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N}\hat{c}_{\boldsymbol{\nu}}\leq C(N+1)^{-(1/p-1)},$$

where $C = \|(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\|_{\ell^p}$.

If the assumptions of Theorem 3.6 hold, Lemma 3.3 combined with the Cauchy–Schwarz inequality and Stechkin's lemma for $q = 2 > p$ give

$$\begin{aligned}\|f - U_{\Lambda_N} f\|_{L^2_\mu} &\le \sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} c_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}} = \sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} \left(\frac{c_{\boldsymbol{\nu}}}{b_{\boldsymbol{\nu}}^{1/2}}\right)\left(b_{\boldsymbol{\nu}}^{1/2}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}\right)\\ &\le \left(\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} b_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}^2\right)^{1/2}\cdot\left(\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N}\frac{c_{\boldsymbol{\nu}}^2}{b_{\boldsymbol{\nu}}}\right)^{1/2}\\ &\le \left(\sum_{\boldsymbol{\nu}\in\mathcal{F}} b_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}^2\right)^{1/2}\cdot\left(\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N}\hat{c}_{\boldsymbol{\nu}}^2\right)^{1/2}\\ &\le C(N+1)^{-(1/p-1/2)},\end{aligned}$$

where now $C = \|(b_{\boldsymbol{\nu}}^{1/2}\|f_{\boldsymbol{\nu}}\|)_{\boldsymbol{\nu}\in\mathcal{F}}\|_{\ell^2}\cdot\|(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\|_{\ell^p}$, respectively. ∎

Remark 3.7. Another application of the weighted ℓ^2-summability property (19) is the analysis of sparse quadrature given in [9], where the author employs the slightly different estimate

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} c_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}} \le \sup_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N} b_{\boldsymbol{\nu}}^{q-1/2}\sum_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_N}\frac{c_{\boldsymbol{\nu}}}{b_{\boldsymbol{\nu}}^{-q}}\, b_{\boldsymbol{\nu}}^{1/2}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}.$$

After showing that the series on the right is bounded and applying Stechkin's lemma to $(b_{\boldsymbol{\nu}}^{q-1/2})_{\boldsymbol{\nu}\in\mathcal{F}}$, this yields the same convergence rate as stated in Theorem 3.6.

Remark 3.8. We mention that sparse collocation attains a smaller convergence rate than best N-term approximation in case the assumptions of Theorem 3.6 hold. Namely, under these assumptions the best N-term rate is $s = \frac{1}{p}$; see [4, Theorem 1.2]. This reduced convergence rate is not due to the additional factors $c_{\boldsymbol{\nu}}$ in the error analysis of sparse collocation. The reason for the slower rate is lack of orthogonality: in the proof of Lemma 3.3 we could not apply Parseval's identity and had to use the triangle inequality to bound the error. This led to bounds in terms of $\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}$ rather than $\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}^2$ as in the case of orthogonal projection, e.g., best N-term approximation.

We emphasize that the construction of such a nonincreasing, p-summable dominating sequence is by no means trivial. Without the first property we cannot conclude that the multi-index sets Λ_N occurring in Stechkin's lemma are monotone, which in turn is needed to apply Lemma 3.3 as the starting point of our error analysis. Of course, we could consider monotone envelopes $\Lambda_N \subset \tilde{\Lambda}_N$ of Λ_N, but their size can grow quite rapidly with N (e.g., polynomially or even faster; see the counterexample below). Moreover, it is not at all obvious that for a sequence $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^p(\mathcal{F})$ there exists a dominating and nonincreasing $(\hat{a}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^p(\mathcal{F})$. In particular, we provide the following counterexample: let $\mathcal{F} = \mathbb{N}$ and define a_n, $n \in \mathbb{N}$, by

$$a_n = \begin{cases}\frac{1}{m^2}, & n = \sum_{k=1}^m k,\\ 0 & \text{otherwise,}\end{cases}$$

i.e., $a_1 = 1, a_2 = 0, a_3 = \frac{1}{4}, a_4 = 0, a_5 = 0, a_6 = \frac{1}{9}, a_7 = 0, \ldots, a_9 = 0, a_{10} = \frac{1}{16}, a_{11} = 0, \ldots$. Then $(a_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N})$. The smallest positive nonincreasing sequence $(\hat{a}_n)_{n\in\mathbb{N}}$

dominating $(a_n)_{n\in\mathbb{N}}$ is given by $\hat{a}_n := \sup_{m\geq n} |a_m|$; see [13, section 3.8]. In our case, we get

$$\hat{a}_n = \frac{1}{m^2} \qquad \text{for each } n \text{ such that} \qquad 1 + \sum_{k=1}^{m-1} k \leq n \leq \sum_{k=1}^{m} k$$

and, thus,

$$\sum_{n=1}^{\infty} |\hat{a}_n| = \sum_{m=1}^{\infty} m \frac{1}{m^2} = \infty.$$

Although the example is somewhat pathological, it does illustrate that for $(a_{\boldsymbol{\nu}}) \in \ell^p(\mathcal{F})$ a p-summable nonincreasing dominating sequence need not exist.

3.2. Sufficient conditions for majorization and weighted summability. We now follow the strategy of Theorem 3.6 and study under which requirements the assumptions of Theorem 3.6 hold. To this end we recall a result from [4] for weighted ℓ^2-summability of Hermite coefficients $\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}$ under the following smoothness conditions on f.

ASSUMPTION 3.9. *Let f satisfy Assumption* 2.1. *There exist an integer $r \in \mathbb{N}_0$ and a sequence of positive numbers $(\tau_m^{-1})_{m\in\mathbb{N}} \in \ell^p(\mathbb{N})$, $p \in (0,2)$, such that*

(a) *for any $\boldsymbol{\alpha} \in \mathcal{F}$ with $|\boldsymbol{\alpha}|_\infty \leq r$ the (weak) partial derivative $\partial^{\boldsymbol{\alpha}} f$ exists and satisfies $\partial^{\boldsymbol{\alpha}} f \in L^2_\mu(\mathbb{R}^{\mathbb{N}}; \mathcal{H})$;*

(b) *there holds*

$$\sum_{|\boldsymbol{\alpha}|_\infty \leq r} \frac{\boldsymbol{\tau}^{2\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!} \|\partial^{\boldsymbol{\alpha}} f\|^2_{L^2_\mu} < \infty, \qquad \text{where } \boldsymbol{\tau}^{\boldsymbol{\alpha}} = \prod_{m=1}^{\infty} \tau_m^{\alpha_m} \text{ and } \boldsymbol{\alpha}! = \prod_{m=1}^{\infty} \alpha_m!. \tag{20}$$

Observe that the sum in (20) is, in general, an infinite series since there are countably many multi-indices such that $|\boldsymbol{\alpha}|_\infty \leq r$. Assumption 3.9(a) states that we require a *finite* order of partial differentiability of f, i.e., up to order r with respect to each variable ξ_m, and, maybe more importantly, Assumption 3.9(b) demands *weighted square-summability* of the L^2_μ-norms of the corresponding partial derivatives. The latter, in particular, implies bounds of the form

$$\|\partial^{\boldsymbol{\alpha}} f\|_{L^2_\mu} \leq K \sqrt{\boldsymbol{\alpha}!}\, \boldsymbol{\tau}^{-\boldsymbol{\alpha}}, \qquad |\boldsymbol{\alpha}|_\infty \leq r,$$

since otherwise the summability requirement (20) would not hold. Recalling that the sequence $(\tau_m^{-1})_{m\in\mathbb{N}}$ belongs to $\ell^p(\mathbb{N})$ this bound implies that, e.g., the L^2_μ-norm of the derivative $\partial^\alpha_{\xi_m} f$, $\alpha \leq r$, decays as $m \to \infty$.

The following result shows that the smoothness condition of Assumption 3.9 implies the first condition (19) of Theorem 3.6.

THEOREM 3.10 (cf. [4, Theorem 3.3]). *Let Assumption* 3.9 *be satisfied. Then, with the weights*

$$b_{\boldsymbol{\nu}} = b_{\boldsymbol{\nu}}(\boldsymbol{\tau}, r) = \sum_{|\boldsymbol{\alpha}|_\infty \leq r} \binom{\boldsymbol{\nu}}{\boldsymbol{\alpha}} \boldsymbol{\tau}^{2\boldsymbol{\alpha}} = \prod_{m\geq 1} \left(\sum_{l=0}^{r} \binom{\nu_m}{l} \tau_m^{2l} \right), \qquad \boldsymbol{\nu} \in \mathcal{F}, \tag{21}$$

where

$$\binom{\boldsymbol{\nu}}{\boldsymbol{\alpha}} := \prod_{m\geq 1} \binom{\nu_m}{\alpha_m} \quad \text{and} \quad \binom{\nu_m}{\alpha_m} := 0 \text{ if } \alpha_m > \nu_m,$$

there holds

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} b_{\boldsymbol{\nu}}\|f_{\boldsymbol{\nu}}\|_{\mathcal{H}}^2 = \sum_{|\boldsymbol{\alpha}|_\infty\le r} \frac{\boldsymbol{\tau}^{2\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}\|\partial^{\boldsymbol{\alpha}} f\|_{L^2_\mu}^2 < \infty. \tag{22}$$

(We mention in passing that in [4] the assertion of Theorem 3.10 was actually proven without requiring that both series in (22) be finite.) To apply Theorem 3.6 it remains to establish the existence of a nonincreasing and p-summable sequence which dominates $c_{\boldsymbol{\nu}}/b_{\boldsymbol{\nu}}^{1/2}$, $\boldsymbol{\nu}\in\mathcal{F}$.[2] Since the $b_{\boldsymbol{\nu}}$ are explicitly given in (21), this boils down to the question of how fast the projection errors $c_{\boldsymbol{\nu}}$ are allowed to grow. As it turns out, a polynomial growth w.r.t. $\boldsymbol{\nu}$ as given in (16) in Proposition 3.1 is sufficient. We therefore state the following lemma, which is strongly based on the techniques developed in the proofs of [4, Lemma 5.1] and [9, Lemma 3.4].

LEMMA 3.11. *Let there exist $\theta\ge 0$ and $K\ge 1$ such that*

$$c_{\boldsymbol{\nu}} \le \prod_{m\ge 1}^{\infty} (1+K\nu_m)^{\theta+1}, \qquad \boldsymbol{\nu}\in\mathcal{F}.$$

Then for any increasing sequence $(\tau_m)_{m\in\mathbb{N}}$ such that $\sum_{m\ge1}\tau_m^{-p}<\infty$ for a $p>0$ and for any $r>2(\theta+1)+\frac{2}{p}$ there exists a nonincreasing sequence $(\hat c_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}\in\ell^p(\mathcal{F})$ such that

$$\frac{c_{\boldsymbol{\nu}}}{b_{\boldsymbol{\nu}}^{1/2}} \le \hat c_{\boldsymbol{\nu}} \qquad \forall \boldsymbol{\nu}\in\mathcal{F},$$

with $b_{\boldsymbol{\nu}} = b_{\boldsymbol{\nu}}(\tau, r)$ as given in (21).

Proof. We begin by constructing the dominating sequence $(\hat c_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ and then show that it belongs to $\ell^p(\mathcal{F})$ and is nonincreasing. We also introduce the notation $a\wedge b := \min(a,b)$ and $a\vee b := \max(a,b)$.

Step 1: *Construction of $\hat c_{\boldsymbol{\nu}}$.* As

$$\binom{\nu_m}{\nu_m\wedge r}\tau_m^{2(\nu_m\wedge r)} \le \sum_{l=0}^{r}\binom{\nu_m}{l}\tau_m^{2l},$$

since the term on the left is part of the sum on the right, we obtain

$$\frac{c_{\boldsymbol{\nu}}^2}{b_{\boldsymbol{\nu}}} \le \prod_{m\ge1}\frac{(1+K\nu_m)^{2(\theta+1)}}{\sum_{l=0}^{r}\binom{\nu_m}{l}\tau_m^{2l}} \le \prod_{m\ge1}\frac{(1+K\nu_m)^{2\theta+2}}{\binom{\nu_m}{\nu_m\wedge r}\tau_m^{2(\nu_m\wedge r)}} = \prod_{m\ge1}\tau_m^{-2(\nu_m\wedge r)}h(\nu_m), \tag{23}$$

where we have introduced the auxiliary function

$$h(n) := \frac{(1+Kn)^{2\theta+2}}{\binom{n}{n\wedge r}}, \quad n\in\mathbb{N}.$$

We now bound $h(n)$ as well as $\tau_m^{-2(\nu_m\wedge r)}$ to construct a dominating sequence $\hat c_{\boldsymbol{\nu}}$.

For $n\le r$ we have $h(n)=(1+Kn)^{2\theta+2}$, but for $n>r$ there holds

$$h(n) = \frac{(1+Kn)^{2\theta+2}}{\binom{n}{r}} = \frac{r!\,(1+Kn)^{2\theta+2}}{(n+1)\cdots(n+r)}.$$

[2] Note that, while $b_{\boldsymbol{\nu}}^{-1/2}$ is obviously nonincreasing, the sequence $c_{\boldsymbol{\nu}}$ is generally not monotonic.

Thus, we have $h \in \mathcal{O}(n^{2\theta+2-r})$, i.e., there exists $C_h \in [1,\infty)$ such that

$$h(n) \le C_h n^{2\theta+2-r} =: \hat{h}(n) \qquad \forall n \in \mathbb{N}.$$

By setting $\hat{h}(0) := 1 = h(0)$, we obtain $h(n) \le \hat{h}(n)$ for all $n \in \mathbb{N}_0$.

Furthermore, since $(\tau_m^{-1})_{m\in\mathbb{N}} \in \ell^p(\mathbb{N})$ we have $\tau_m \to \infty$ as $m \to \infty$. Thus, there exists an $M \in \mathbb{N}$ such that $\tau_m \ge \sqrt{C_h}$ for $m \ge M$ and $\tau_m \le \sqrt{C_h}$ for $m < M$. We define

$$\hat{\tau}_m := \sqrt{C_h} \vee \tau_m, \qquad m \in \mathbb{N},$$

and note that $\hat{\tau}_m \ge 1$ as well as $(\hat{\tau}_m^{-1})_{m\in\mathbb{N}} \in \ell^p(\mathbb{N})$ by assumption. Moreover, we obtain for $m \ge M$

$$\tau_m^{2(\nu_m \wedge r)} = \hat{\tau}_m^{2(\nu_m \wedge r)} \ge \hat{\tau}_m^{2(\nu_m \wedge 1)} \qquad \forall \nu_m \in \mathbb{N}_0,$$

since $\tau_m = \hat{\tau}_m \ge \sqrt{C_h} \ge 1$ in this case. Further, we define

$$C_\tau := \min_{m\ge 1} \min_{n=0,\dots,r} \frac{\tau_m^{2n}}{C_h^{n\wedge 1}} > 0,$$

which then yields for $1 \le m < M$

$$\tau_m^{2(\nu_m\wedge r)} \ge C_\tau\, C_h^{\nu_m\wedge 1} = C_\tau\, \hat{\tau}_m^{2(\nu_m\wedge 1)} \qquad \forall \nu_m \in \mathbb{N}_0$$

since $\hat{\tau}_m = \sqrt{C_h}$ for $m < M$. We now define

$$\hat{c}_{\boldsymbol{\nu}}^2 := C_\tau^{-M} \prod_{m\ge 1} \hat{\tau}_m^{-2(\nu_m\wedge 1)} \hat{h}(\nu_m) \tag{24}$$

and note that $\hat{c}_{\boldsymbol{\nu}}^2$ dominates $\frac{c_{\boldsymbol{\nu}}^2}{b_{\boldsymbol{\nu}}}$ by (23).

Step 2: *Show that* $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^p(\mathcal{F})$. As for the p-summability, there holds

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\mathcal{F}} \hat{c}_{\boldsymbol{\nu}}^p &= C_\tau^{-pM/2} \sum_{\boldsymbol{\nu}\in\mathcal{F}} \prod_{m\ge 1} \hat{\tau}_m^{-p(\nu_m\wedge 1)} \hat{h}^{p/2}(\nu_m) \\ &= C_\tau^{-pM/2} \prod_{m\ge 1} \sum_{n\ge 0} \hat{\tau}_m^{-p(n\wedge 1)} \hat{h}^{p/2}(n).\end{aligned}$$

We get

$$\sum_{n\ge 0} \hat{\tau}_m^{-p(n\wedge 1)} \hat{h}^{p/2}(n) = 1 + C_h^{p/2} \hat{\tau}_m^{-p} \underbrace{\sum_{n\ge 1} n^{-p(r-2\theta-2)/2}}_{=:S},$$

where the sum S is finite due to the assumption $\frac{p}{2}(r-2\theta-2) = \frac{p}{2}(r-2\theta-2) > 1$. The rest follows by using $\log(1+x) \le x$ for x positive in order to get

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} \hat{c}_{\boldsymbol{\nu}}^p = C_\tau^{-pM/2} \prod_{m\ge 1} (1 + C_h^{p/2} S \hat{\tau}_m^{-p}) \le C_\tau^{-pM/2} \exp\left(C_h^{p/2} S \sum_{m\ge 1} \hat{\tau}_m^{-p} \right) < \infty$$

since $(\hat{\tau}_m^{-1})_{m\in\mathbb{N}}$ is in $\ell^p(\mathbb{N})$ by construction.

Step 3: *Show that* $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ *is nonincreasing.* Let $\boldsymbol{\nu} \in \mathcal{F}$ be arbitrary. If $m \in \operatorname{supp}\boldsymbol{\nu} = \{m \in \mathbb{N} : \nu_m > 0\}$, then

$$\hat{c}^2_{\boldsymbol{\nu}+e_m} = \hat{c}^2_{\boldsymbol{\nu}} \cdot \frac{\hat{h}(\nu_m+1)}{\hat{h}(\nu_m)} \le \hat{c}^2_{\boldsymbol{\nu}},$$

since $\hat{h}(n)$ is nonincreasing for $n \ge 1$. Now let $m \notin \operatorname{supp}\boldsymbol{\nu}$. Then

$$\hat{c}^2_{\boldsymbol{\nu}+e_m} = \hat{c}^2_{\boldsymbol{\nu}} \cdot \hat{\tau}_m^{-2} \cdot \hat{h}(1) = \hat{c}^2_{\boldsymbol{\nu}} \cdot C_h \hat{\tau}_m^{-2} \le \hat{c}^2_{\boldsymbol{\nu}} \cdot C_h (\sqrt{C_h})^{-2} \le \hat{c}^2_{\boldsymbol{\nu}}.$$

In summary, we obtain

$$\hat{c}_{\boldsymbol{\nu}+e_m} \le \hat{c}_{\boldsymbol{\nu}} \qquad \forall m \in \mathbb{N};$$

hence, $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is nonincreasing. ∎

Observe that the proof of Lemma 3.11 is actually constructive and provides the basis for a computational technique for generating suitable multi-index sets to be introduced in section 5.

We can now state our main convergence result for sparse collocation.

THEOREM 3.12 (convergence of sparse collocation). *Assume that for* $\theta \ge 0$ *and* $K \ge 1$ *there holds*

$$\|\Delta_i H_\nu\|_{L^2_\mu} \le (1+K\nu)^\theta, \qquad i \in \mathbb{N}_0. \tag{25}$$

Then for any function f *satisfying Assumption* 3.9 *with* $r > 2(\theta+1) + \frac{2}{p}$, $p \in (0,2)$, *and Assumption* 2.3, *there exists a nested sequence of monotone finite subsets* $\Lambda_N \subset \mathcal{F}$ *with* $|\Lambda_N| = N$ *such that for the sparse collocation error there holds*

$$\|f - U_{\Lambda_N} f\|_{L^2_\mu} \le C(1+N)^{-\left(\frac{1}{p}-\frac{1}{2}\right)}.$$

Proof. We prove the assertion by verifying the assumptions of Theorem 3.6. Since f satisfies Assumption 3.9 with $r > 2(\theta+1)+\frac{2}{p}$, condition (19) of Theorem 3.6 holds due to Theorem 3.10. Moreover, we can apply Lemma 3.11 to verify the remaining assumption of Theorem 3.6 about a nonincreasing dominating sequence $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^p(\mathcal{F})$, $p \in (0,2)$: due to Proposition 3.1 the bound (25) implies

$$c_{\boldsymbol{\nu}} \le \prod_{m\ge 1}^{\infty} (1+K\nu_m)^{\theta+1}, \qquad \boldsymbol{\nu} \in \mathcal{F},$$

and the sequence $(\tau_m)_{m\in\mathbb{N}}$ appearing in Assumption 3.9 can w.l.o.g. be assumed to be increasing (otherwise one can permute the dimensions accordingly). ∎

3.3. Convergence of sparse collocation at Gauss–Hermite nodes. In the following, we verify the assumption (25) in Theorem 3.12 for the interpolation operators U_i based on Gauss–Hermite nodes. Moreover, we bound the number of sparse grid points $|\Xi_{\Lambda_N}|$ associated with a multi-index set Λ_N allowing us to relate the convergence rate previously derived to a quantity which reflects the computational effort of the collocation approximation. For nested univariate node sets, i.e., when $\Xi_{i+i} = \Xi_i \cup \{\xi_{i+1}^{(i+1)}\}$, we have $|\Xi_{\Lambda_N}| = |\Lambda_N|$. This simple relation, however, fails to hold for nonnested interpolation sequences such as those based on Gauss–Hermite nodes.

LEMMA 3.13. *For U_i being the interpolation operator based on the zeros of the $(i+1)$th Hermite polynomial, we have for each $\nu \in \mathbb{N}$ that*

$$\|U_i H_\nu\|_{L^2_\mu}^2 \leq c^2 \,\mathrm{e}\, \sqrt{2\nu - 1} \qquad \forall i \in \mathbb{N}_0,$$

where $c = 1.086435$ is the constant appearing in Cramér's inequality *for Hermite functions. In particular, there holds*

$$\|\Delta_i H_\nu\|_{L^2_\mu} \leq (1 + K\nu)$$

with $K = 2c\sqrt{\mathrm{e}} > 1$.

Proof. We start by recalling the L^2_μ-orthogonality (μ refers here to the univariate standard Gaussian measure $N(0,1)$) of Lagrange basis polynomials $L_k^{(i)}$ constructed from the zeros $\{\xi_k^{(i)}\}_{k=0}^i$ of the Hermite polynomial of degree $i+1$ (see [42, Theorem 14.2.1]). This orthogonality yields

$$\begin{aligned}\|U_i H_\nu\|_{L^2_\mu}^2 &= \int_{\mathbb{R}} \left(\sum_{k=0}^{i} H_\nu(\xi_k^{(i)}) L_k^{(i)}(\xi) \right)^2 \mu(\mathrm{d}\xi) = \sum_{k=0}^{i} H_\nu^2(\xi_k^{(i)}) \int_{\mathbb{R}} \left(L_k^{(i)}(\xi) \right)^2 \mu(\mathrm{d}\xi) \\ &= \sum_{k=0}^{i} H_\nu^2(\xi_k^{(i)}) w_k^{(i)},\end{aligned}$$

where $\{w_k^{(i)}\}_{k=0}^i$ denotes the weights of the Gauss quadrature formula based on the zeros of the $(i+1)$th Hermite polynomial; see also [42, Theorem 14.2.1].

Next, we recall Cramér's inequality for the physicists' Hermite polynomials $\tilde{H}_\nu$ taken w.r.t. the weight function $\tilde{\rho}(\xi) = \exp(-\xi^2)$, i.e.,

$$|\tilde{H}_n(\xi)| \leq c\pi^{-1/4} \exp(\xi^2/2);$$

see, e.g., [1, Chapter 22, p. 787]. From the relation $\tilde{H}_n(\xi) = \pi^{-1/4} H_n(\xi\sqrt{2})$ [1, Chapter 22, p. 778], we conclude that

$$|H_n(\xi)| \leq c \exp(\xi^2/4)$$

and thus

$$\|U_i H_\nu\|_{L^2_\mu}^2 \leq c^2 \sum_{k=0}^{i} \exp(\xi_{ki}^2/2) w_{ki},$$

where we have switched the notation to $\xi_{ki} := \xi_k^{(i)}$ and $w_{ki} := w_k^{(i)}$ for convenience. Furthermore, we use a consequence of [33, Lemma 4]. The latter states, for $\tilde{\xi}_{kn}$ denoting the zeros of $\tilde{H}_n$ and $\tilde{w}_{kn}$ the Christoffel numbers of the corresponding Gauss–Hermite quadrature (i.e., Gauss–Hermite weights for $\tilde{\rho}$), that

$$\sum_{k=1}^{n} \tilde{w}_{kn} \exp(\tilde{\xi}_{kn}^2) \leq \mathrm{e}\, \sqrt{\pi(2n+1)}.$$

It can be easily verified that

$$\xi_{kn} = \sqrt{2}\tilde{\xi}_{kn} \quad \text{and} \quad w_{kn} = \pi^{-1/2} \tilde{w}_{kn}.$$

Hence, we get

$$\sum_{k=0}^{i} \exp(\xi_{ki}^2/2) w_{ki} \leq \mathrm{e}\sqrt{2(i+1)+1},$$

and by noticing that for $i \geq \nu$ we have $U_i H_\nu = H_\nu$ and, thus, $\|U_i H_\nu\|_{L^2_\mu}^2 = 1$, and for $i = \nu - 1$ we get $U_i H_\nu \equiv 0$ and the first assertion is shown.

For the second statement we observe

$$\|U_i H_\nu\|_{L^2_\mu}^2 \leq c^2 \,\mathrm{e}\, \nu \qquad \forall i \in \mathbb{N}_0,\ \forall \nu \geq 1$$

since $\nu \geq \sqrt{2\nu - 1}$ for $\nu \geq 1$. And, because of $\Delta_i H_0 \equiv 0$ for $i \geq 1$ and $\Delta_0 H_0 \equiv H_0$, we get

$$\|\Delta_i H_\nu\|_{L^2_\mu} \leq 1 + K\nu \qquad \forall i, \nu \in \mathbb{N}_0. \qquad \square$$

Hence, interpolation at Gauss–Hermite points satisfies the assumptions of Theorem 3.12 with $\theta = 1$ and we obtain the following theorem.

THEOREM 3.14 (convergence of sparse collocation, Gauss–Hermite nodes). *For any function f satisfying Assumption 3.9 with $r > 4 + \frac{2}{p}$, $p \in (0, 2)$, and Assumption 2.3 there exists a nested sequence of monotone finite subsets $\Lambda_N \subset \mathcal{F}$ with $|\Lambda_N| = N$ such that for the error of the sparse collocation operator U_{Λ_N} based on Gauss–Hermite nodes there holds*

$$\|f - U_{\Lambda_N} f\|_{L^2_\mu} \leq C(1+N)^{-\left(\frac{1}{p} - \frac{1}{2}\right)}.$$

Remark 3.15. In numerical experiments we have actually observed for $\nu = 0, \ldots, 39$ that

$$\|U_i H_\nu\|_{L^2_\mu} \leq 1 \qquad \forall i \in \mathbb{N}_0;$$

see Figure 1. This would imply

$$\|\Delta_i H_\nu\|_{L^2_\mu} \leq \begin{cases} 1 & \text{if } \nu = 0, \\ 2 & \text{otherwise} \end{cases} \qquad \forall i, \nu \in \mathbb{N}_0.$$

Again, we even observed a smaller bound numerically; see the right plot in Figure 1. However, we have not been able to prove $\|U_i H_\nu\|_{L^2_\mu} \leq 1$, and the improvement in the statement of Theorem 3.14 would have been minor, i.e., the assertion would also hold with the same rate for functions $f : \Gamma \to \mathcal{H}$ satisfying Assumption 3.9 with $r > 2 + \frac{2}{p}$. Note that similar numerical evidence was presented in [9] for quadrature operators applied to Hermite polynomials. See also [7] for analogous numerical bounds in the case of Genz–Keister points.

3.4. Convergence rate with respect to the number of collocation nodes. We now derive bounds for the number of nodes in the sparse grid Ξ_Λ associated with U_Λ. Consider first the following simple monotone index set of cardinality N: $\Lambda_N = \{0\boldsymbol{e}_j, \ldots, (N-1)\boldsymbol{e}_j\}$ for some $j \in \mathbb{N}$. Then due to $|\Xi_{\{k\boldsymbol{e}_j\}}| = (k+1)$ we get for this Λ_N that

$$|\Xi_{\Lambda_N}| \leq \sum_{k=0}^{N-1} (k+1) = \frac{N(N+1)}{2} \in \mathcal{O}(N^2).$$

The quadratic complexity is essentially sharp, since 0 is the only reoccurring Gauss–Hermite node. We show in the subsequent proposition that this complexity holds also for arbitrary monotone multi-index sets.

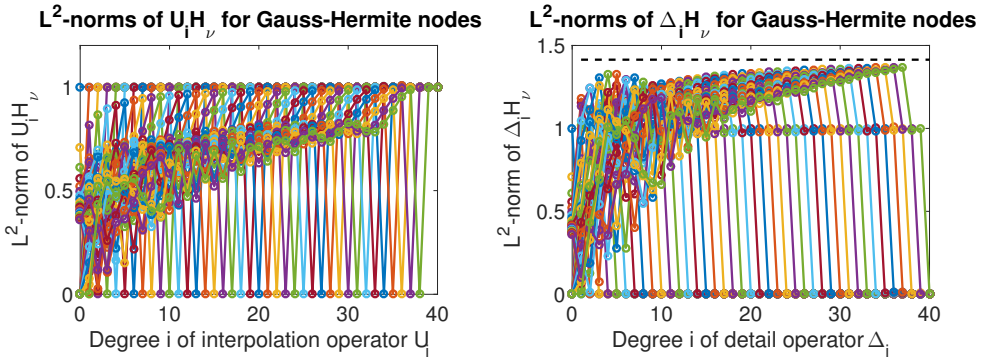

FIG. 1. *Computed values of* $\|U_i H_\nu\|_{L^2_\mu}$ *(left) and* $\|\Delta_i H_\nu\|_{L^2_\mu}$ *(right) for Gauss–Hermite nodes. Each color represents the trend of* $\|U_i H_\nu\|_{L^2_\mu}$ *(resp.,* $\|\Delta_i H_\nu\|_{L^2_\mu}$*) for a fixed polynomial* H_ν *as* i *increases. The dashed, black line in the right plot indicates the value* $\sqrt{2}$.

PROPOSITION 3.16. *Let* $\Lambda \subset \mathcal{F}$ *be finite and monotone. Then there holds*

$$|\Xi_\Lambda| \leq |\Lambda|^2. \tag{26}$$

Proof. We obtain due to $|\Xi^{(i)}| = \prod_{m\geq 1}(1+i_m) = |\mathcal{R}_{\boldsymbol{i}}|$ that[3]

$$|\Xi_\Lambda| \leq \sum_{\boldsymbol{i}\in\Lambda} |\Xi^{(\boldsymbol{i})}| = \sum_{\boldsymbol{i}\in\Lambda}\prod_{m\in\mathbb{N}}(1+i_m) = \sum_{\boldsymbol{i}\in\Lambda}|\mathcal{R}_{\boldsymbol{i}}| \leq |\Lambda| \max_{\boldsymbol{i}\in\Lambda}|\mathcal{R}_{\boldsymbol{i}}| \leq |\Lambda|^2. \qquad \square$$

Thus, employing nonnested points such as Gauss–Hermite nodes yields at most a quadratic growth of the number of sparse grid points

$$|\Xi_\Lambda| \in \mathcal{O}(|\Lambda|^2),$$

whereas in the nested case one has $|\Xi_\Lambda| = |\Lambda|$.

Remark 3.17. We provide some numerical validation of the bound (26). More precisely, we consider the following two families of multi-index sets Λ (cf. [6]):

- *Total Degree (TD):*
$$\Lambda = \Lambda(w, M) = \left\{\boldsymbol{\nu} \in \mathcal{F} : \sum_{m=1}^{M} \nu_m \leq w,\ \nu_m = 0 \text{ for } m > M\right\}.$$
- *Hyperbolic Cross (HC):*
$$\Lambda = \Lambda(w, M) = \left\{\boldsymbol{\nu} \in \mathcal{F} : \prod_{m=1}^{M} (\nu_m + 1) \leq w,\ \nu_m = 0 \text{ for } m > M\right\}.$$

In Figure 2 we fix the number of (active) dimensions M and display the cardinality of $\Xi_{\Lambda(w,M)}$ for both choices of $\Lambda(w, M)$ and increasing values of $w \in \mathbb{N}$. The plot shows that estimate (26) is valid but slightly pessimistic for the two specific examples considered.

We finally arrive at the resulting error-cost theorem.

[3] We remark that the sharper bound $|\Xi_\Lambda| \leq \frac{1}{2}|\Lambda|^2 + \frac{1}{2}|\Lambda|$ can also be shown with a little more effort; see [18, section 3.4].

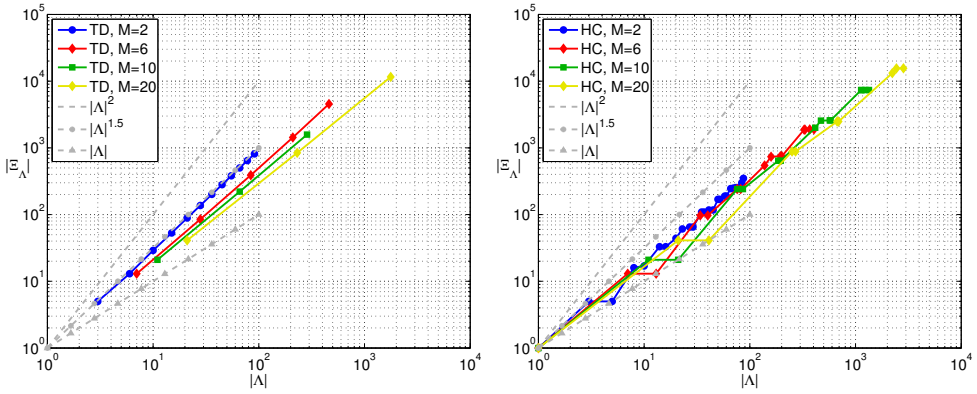

FIG. 2. *Numerical verification of estimate* (26) *for "Total Degree" sparse grids (left) and "Hyperbolic Cross" sparse grids (right).*

THEOREM 3.18 (convergence rate of Gauss–Hermite sparse grid collocation in terms of nodes). *For any function f satisfying Assumption* 3.9 *with $r > 4 + \frac{2}{p}$, $p \in (0, 2)$, and Assumption* 2.3 *there exists a nested sequence of monotone finite subsets $\Lambda_N \subset \mathcal{F}$ with $|\Lambda_N| = N$ such that for the error of the sparse collocation operator U_{Λ_N} based on Gauss–Hermite nodes there holds*

$$\|f - U_{\Lambda_N} f\|_{L^2_\mu} \leq C |\Xi_{\Lambda_N}|^{-\left(\frac{1}{2p} - \frac{1}{4}\right)},$$

where C depends on f.

To achieve an approximation error $\|f - U_{\Lambda_N} f\|_{L^2_\mu} \leq \varepsilon$, Theorem 3.18 states that this corresponds to

$$\text{cost}(\varepsilon) \in \mathcal{O}\left(\varepsilon^{\frac{1}{2p} - \frac{1}{4}}\right) \tag{27}$$

number of function evaluations of f. The cost complexity (27) does not account for the computational work necessary to construct the associated multi-index sets Λ_N. This is a very important issue. Typically, these are constructed employing adaptive algorithms; see [11, 39, 35] and also section 5. Our result makes no statement about the actual computational work these involve.

Remark 3.19 (on sparse collocation employing weighted Leja points). As mentioned in the introduction, weighted Leja points [32] seem to be a promising node family for interpolation and sparse collocation. So far, we have, however, been unable to prove bounds like (25) for these. Possibly a more suitable approach for analyzing convergence in the case of weighted Leja nodes is to measure the approximation error in the L^∞_μ-norm instead of the L^2_μ-norm and to estimate the corresponding Lebesgue constant. See [28] for first results on the latter—which does not yet imply an analogous estimate to (25)—and [12, 11] for the convergence analysis of sparse collocation using Leja points on $[-1, 1]$ via estimates of the associated Lebesgue constant [10].

4. Application to elliptic PDEs. We recall our motivation from the introduction: Approximating the weak solution u of an elliptic boundary value problem with lognormal diffusion coefficient as in (1) where $f \in L^2(D)$ and $a(\boldsymbol{\xi}) \in L^\infty(D)$ is given as in (2). We will discuss now under which conditions the mapping $\boldsymbol{\xi} \mapsto u(\boldsymbol{\xi}) \in H^1_0(D)$

satisfies Assumptions 2.1, 2.3, and 3.9 and can therefore be approximated by sparse grid collocation methods based on Gauss–Hermite nodes as outlined in the previous section. We mainly cite results from [4] but try to emphasize those details which are sometimes omitted in the literature.

Verifying Assumption 2.1. We first investigate the domain Γ of the mapping $\boldsymbol{\xi} \mapsto u(\boldsymbol{\xi})$. Note that $\Gamma \neq \mathbb{R}^{\mathbb{N}}$ since for arbitrary $\boldsymbol{\xi} \in \mathbb{R}^{\mathbb{N}}$ the expansion (2) need not converge. A natural domain Γ for the mapping $\boldsymbol{\xi} \mapsto u(\boldsymbol{\xi})$ is

$$\Gamma := \left\{ \boldsymbol{\xi} \in \mathbb{R}^{\mathbb{N}} : \left\| \sum_{m=1}^{\infty} \phi_m \xi_m \right\|_{L^\infty(D)} < \infty \right\}. \tag{28}$$

Further, a natural condition on the decay of the ϕ_m is

$$\sum_{m=1}^{\infty} \|\phi_m\|_{L^\infty(D)} < \infty \tag{29}$$

since (29) implies that the series (2) converges $\mathbf{P}$-a.s. in $L^\infty(D)$; see [40, Lemma 2.28]. Thus, if (29) holds, then $\mu(\Gamma) = 1$. It remains to state conditions under which we can ensure that $\boldsymbol{\xi} \mapsto u(\boldsymbol{\xi})$ belongs to $L^2_\mu(\Gamma; H^1_0(D))$. Measurability follows from the continuous dependence of the weak solution $u \in H^1_0(D)$ on $\exp(a) \in L^\infty(D)$; see [24]. Moreover, if we can ensure that for $\underline{a}(\boldsymbol{\xi}) := \operatorname{essinf}_{x\in D} \exp(a(x, \boldsymbol{\xi}))$ we have $\underline{a}^{-1} \in L^2_\mu(\Gamma; \mathbb{R})$ (e.g., by Fernique's lemma, as shown in [8]), then the $\boldsymbol{\xi}$-pointwise application of the Lax–Milgram lemma [24] yields for the random solution $u \in L^2_\mu(\Gamma; H^1_0(D))$. The latter can be guaranteed by an even weaker assumption than (29).

ASSUMPTION 4.1 (see [4, Assumption A]). *There exists a sequence $(\tau_m)_{m\in\mathbb{N}}$ of strictly positive numbers such that*

$$\sup_{x\in D} \sum_{m=1}^{\infty} \tau_m |\phi_m(x)| < \infty, \qquad \sum_{m=1}^{\infty} \exp(-\tau_m^2) < \infty.$$

Under Assumption 4.1, it is shown in [4, Corollary 2.1] that $u \in L^2_\mu(\Gamma; H^1_0(D))$ with $\mu(\Gamma) = 1$; hence $u : \Gamma \to H^1_0(D)$ satisfies Assumption 2.1.

Verifying Assumption 2.3. It is obvious that for Gauss–Hermite nodes there holds $\Xi^{(\boldsymbol{i})} \subset \Gamma$, $\boldsymbol{i} \in \mathcal{F}$, with Γ as in (28), because due to $\boldsymbol{i} \in \mathcal{F}$ there exists an $M \in \mathbb{N}$ such that for $\boldsymbol{\xi} \in \Xi^{(\boldsymbol{i})}$ we have $\xi_m = \xi_0^{(0)}$ for any $m \geq M$ and $\xi_0^{(0)} = 0$. Actually, by Assumption 4.1 there holds for any $\boldsymbol{\xi} \in \ell^\infty(\mathbb{N})$ that $\boldsymbol{\xi} \in \Gamma$:

$$\left\| \sum_{m=1}^{\infty} \phi_m \xi_m \right\|_{L^\infty(D)} \leq \|\boldsymbol{\xi}\|_{\ell^\infty} \sup_{x\in D} \sum_{m=1}^{\infty} |\phi_m(x)| \leq \frac{\|\boldsymbol{\xi}\|_{\ell^\infty}}{\min_m \tau_m} \sup_{x\in D} \sum_{m=1}^{\infty} \tau_m \, |\phi_m(x)| < \infty,$$

where $\min_m \tau_m > 0$, because Assumption 4.1 implies $\tau_m \to \infty$ as $m \to \infty$.

Verifying Assumption 3.9. Again, we refer to results from [4], namely, [4, Theorem 4.2], where the authors show that the (weak) solution u of (1) satisfies Assumption 3.9 for any $r \in \mathbb{N}_0$ given the following assumption holds.

ASSUMPTION 4.2. *There exists a strictly positive sequence $(\tau_m^{-1})_{m\in\mathbb{N}} \in \ell^p(\mathbb{N})$, $p \in (0, 2)$, such that*

$$\sup_{x\in D} \sum_{m=1}^{\infty} \tau_m |\phi_m(x)| < \infty.$$

Note that Assumption 4.2 implies Assumption 4.1; see [4, Remark 2.2]. Hence, we obtain the following theorem.

THEOREM 4.3. *Let a be given as in* (2) *and satisfy Assumption* 4.2. *Then there exists a nested sequence of monotone finite subsets $\Lambda_N \subset \mathcal{F}$ with $|\Lambda_N| = N$ such that for the sparse collocation operator U_{Λ_N} based on Gauss–Hermite nodes applied to the solution u of* (1) *there holds*

$$\|u - U_{\Lambda_N} u\|_{L^2_\mu} \;\le\; C_1 N^{-\left(\frac{1}{p}-\frac{1}{2}\right)} \;\le\; C_2 |\Xi_{\Lambda_N}|^{-\left(\frac{1}{2p}-\frac{1}{4}\right)}.$$

5. Numerical experiments. We apply the sparse collocation outlined and analyzed in the previous sections to approximate the solution u of a simple boundary value problem taken from [4, section 7]. In particular, we verify numerically the statement of Theorem 4.3 and provide some comments on algorithms for constructing sparse grid approximations.

5.1. Problem setting. We consider the following boundary value problem on the unit interval $D = [0, 1]$:

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(a(x,\boldsymbol{\xi})\,\frac{\mathrm{d}}{\mathrm{d}x}u(x,\boldsymbol{\xi})\right) = f(x), \quad u(0,\boldsymbol{\xi}) = u(1,\boldsymbol{\xi}) = 0, \qquad \mu\text{-a.e.}, \tag{30}$$

where we choose $f(x) = 0.03\sin(2\pi x)$ and employ for $\log a$ the following expansion:

$$\log a(x,\boldsymbol{\xi}) = 0.1\sum_{m=1}^{\infty} \frac{\sqrt{2}}{(\pi m)^q}\,\sin(m\pi x)\,\xi_m, \qquad \xi_m \sim N(0,1) \text{ i.i.d.}, \; q \ge 1. \tag{31}$$

For $q = 1$ the random field $\log a$ is a Brownian bridge (cf. [4, section 7]), and for $q > 1$ it represents a smoother random field. In particular, for $\phi_m(x) := \frac{\sqrt{2}}{(\pi m)^q}\sin(m\pi x)$ and $k = q - 1 - \varepsilon$ with $\varepsilon > 0$ we obtain the bound

$$\sup_{x\in D}\sum_{m\ge 1} m^k |\phi_m(x)| \le \frac{\sqrt{2}}{\pi^q}\sum_{m\ge 1} m^{-(q-k)} \propto \sum_{m\ge 1} m^{-(1+\varepsilon)} < \infty.$$

Thus, given $q > 1$ the expansion (31) satisfies Assumption 4.2 with $\tau_m = m^{q-1}$ for all $p > \frac{1}{q-1}$ and by Theorem 4.3 there exists for $q > 1.5$ a nested sequence of monotone finite subsets $\Lambda_N \subset \mathcal{F}$, $|\Lambda_N| = N$, such that for the sparse collocation operator U_{Λ_N} based on Gauss–Hermite nodes there holds

$$\|u - U_{\Lambda_N} u\|_{L^2_\mu(\mathbb{R}^{\mathbb{N}};H^1_0(D))} \;\le\; C N^{-(q-1.5)} \;\le\; C|\Xi_{\Lambda_N}|^{-\left(\frac{q-1.5}{2}\right)}. \tag{32}$$

In the following we will verify these rates numerically for various values of q.

5.2. Numerical algorithms. The multi-index sets Λ_N appearing in Theorem 4.3 and (32) correspond to the largest entries in a p-summable decreasing sequence $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ which dominates $\left(c_{\boldsymbol{\nu}}/b_{\boldsymbol{\nu}}^{1/2}\right)_{\boldsymbol{\nu}\in\mathcal{F}}$. Such a dominating sequence was constructed in Lemma 3.11 and can be employed in the subsequent a priori algorithm in order to obtain the multi-index sets Λ_N numerically. However, in practice we employ a slightly modified dominating sequence $(\hat{d}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ in the a priori algorithm which shows a better performance, as well as an a posteriori algorithm for adaptively constructing the sets Λ_N.

A priori algorithm. The following greedy algorithm is based on [22] and appears in a similar form in the recent work [9]. It successively adds to the set of multi-indices Λ a new multi-index $\boldsymbol{\nu}$ from the set of neighbors $\mathcal{N}(\Lambda)$ which maximizes $\hat{c}_{\boldsymbol{\nu}}$ with $\hat{c}_{\boldsymbol{\nu}}$ as given in (24). A constraint m_{buffer} restricts the index of dimensions considered for admissible neighbors:

1. Initialize $N = 1$ and $\tilde{\Lambda}_N := \{\mathbf{0}\}$; choose $m_{\text{buffer}} \in \mathbb{N}$ and $N_{\max} \in \mathbb{N}$.
2. For $N = 2, \dots, N_{\max}$ set

$$\tilde{\Lambda}_N := \tilde{\Lambda}_{N-1} \cup \{\boldsymbol{\nu}_N^*\}, \qquad \boldsymbol{\nu}_N^* := \operatorname*{argmax}_{\boldsymbol{\nu} \in \mathcal{N}(\tilde{\Lambda}_{N-1})} \hat{c}_{\boldsymbol{\nu}}, \tag{33}$$

where, with $\operatorname{supp}(\boldsymbol{\nu}) := \{m \in \mathbb{N} : \nu_m > 0\}$ and $\operatorname{supp}(\Lambda) := \bigcup_{\boldsymbol{\nu}\in\Lambda} \operatorname{supp}(\boldsymbol{\nu})$, we define

$$\mathcal{N}(\Lambda) := \{\boldsymbol{\nu} \in \mathcal{F} \setminus \Lambda : \boldsymbol{\nu} - \boldsymbol{e}_m \in \Lambda \ \forall m \in \operatorname{supp}(\boldsymbol{\nu}) \text{ and } \nu_m = 0 \text{ for } m > \max(\operatorname{supp}(\Lambda)) + m_{\text{buffer}}\}.$$

The set of admissible neighbors $\mathcal{N}(\Lambda)$ of Λ is defined such that adding any $\boldsymbol{\nu} \in \mathcal{N}(\Lambda)$ to Λ maintains monotonicity. The restriction in the definition of $\mathcal{N}(\Lambda)$ above is that we do not allow the *activation* of any dimension $m \in \mathbb{N}$, i.e., including $\boldsymbol{\nu} = \boldsymbol{e}_m$ for arbitrarily (large) $m \in \mathbb{N}$, but restrict the selection to the "next" m_{buffer} higher dimensions. If the assumptions of Lemma 3.11 are satisfied, in particular that τ_m is increasing, a buffer of $m_{\text{buffer}} = 1$ is sufficient to construct the multi-index sets Λ_N consisting of the largest entries of $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$.

However, for our numerical simulations, we choose a slightly different sequence

$$\hat{d}_{\boldsymbol{\nu}} := \prod_{m\geq 1} (\nu_m)^{2\theta+2-r} \tau_m^{-2(1\wedge\nu_m)}$$

with $\theta = 1$ and a suitable value[4] for $r > 2(\theta+1) + \frac{2}{p}$, and modify (33) by $\boldsymbol{\nu}_N^* := \operatorname{argmax}_{\boldsymbol{\nu}\in\mathcal{N}(\tilde{\Lambda}_{N-1})} \hat{d}_{\boldsymbol{\nu}}$. Besides omitted multiplicative constants the only difference between $\hat{c}_\nu$ and $\hat{d}_\nu$ is that $\hat{c}_\nu$ employs $\hat{\tau}_m = \sqrt{C_h} \vee \tau_m$ instead of[5] τ_m—the former can yield a longer preasymptotic behavior of the resulting numerical algorithm.

A posteriori algorithm. In addition to this a priori construction which requires little overhead to construct the monotone multi-index sets Λ_N, we also apply a more costly a posteriori algorithm for generating monotone multi-index sets Λ_N. Such an algorithm already appeared in [39, 11, 12, 9, 35] and is motivated by using a posteriori heuristics for estimating the improvement of including $\Delta_{\boldsymbol{\nu}} u$ in the sparse collocation approximation. In particular, the a posteriori algorithm works exactly as the a priori algorithm except for substituting the choice (33) by

$$\boldsymbol{\nu}_n^* := \operatorname*{argmax}_{\boldsymbol{\nu}\in\mathcal{N}(\tilde{\Lambda}_{n-1})} \frac{\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}}{|\Xi^{(\boldsymbol{\nu})}|}, \tag{34}$$

where $\|f\|_{L^\infty_\mu(\Gamma;H^1_0(D))} = \sup_{\boldsymbol{\xi}\in\Gamma} \|\rho(\boldsymbol{\xi})\, f(\boldsymbol{\xi})\|_{H^1_0(D)}$ and $\rho(\boldsymbol{\xi}) = \exp(-\frac{1}{2}\sum_{m\geq 1}\xi_m^2)$ represents the (unnormalized) product density function of $\mu = \bigotimes_{m\geq 1} N(0,1)$; see, e.g., [35]. The ratio $\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}/|\Xi^{(\boldsymbol{\nu})}|$ represents the *profitability* or *profit* of the

[4] We used $r = 2\,(2(\theta+1) + 2/p + 1) = 10 + 4(q-1)$ in the numerical simulations.

[5] The choice of $\hat{\tau}_m$ ensures that $(\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is nonincreasing, but since the a priori algorithm always constructs monotone multi-index sets, we can ignore this property.

multi-index $\boldsymbol{\nu} \in \mathcal{F}$, i.e., the associated gain in approximation $\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}$ relative to the associated computational cost $|\Xi^{(\boldsymbol{\nu})}|$. By choosing the most profitable multi-index in the neighborhood of $\tilde{\Lambda}_{n-1}$ we may obtain a better sparse collocation approximation than when applying the a priori construction (33), although the theory developed above does not apply to the multi-indices generated in this way. Observe that the idea of "profit-driven" adaptive sparse grids is not bound to the current specific choice of norms, and other alternatives might be envisaged, as is indeed the case in [39, 11, 12, 9, 35]. Here we follow, in particular, [35] and choose to use $\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}$ instead of $\|\Delta_{\boldsymbol{\nu}} u\|_{L^2_\mu(\Gamma;H^1_0(D))}$ employed in the convergence theorems stated earlier, due to its convenient implementation and in analogy with the algorithms for the uniform random variables case. Note also that $\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}$ is a stronger norm than $\|\Delta_{\boldsymbol{\nu}} u\|_{L^2_\mu(\Gamma;H^1_0(D))}$. Again, we remark at this juncture that all possible choices of norms here are motivated just by heuristic arguments, and no claim is made on the convergence of the resulting algorithm. Specifically, we have estimated $\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))}$ as in [35] by

$$\|\Delta_{\boldsymbol{\nu}} u\|_{L^\infty_\mu(\Gamma;H^1_0(D))} \approx \max_{\boldsymbol{\xi_k} \in \Xi^{(\boldsymbol{\nu})}} \|\rho(\boldsymbol{\xi_k})\, [\Delta_{\boldsymbol{\nu}} u]\,(\boldsymbol{\xi_k})\|_{H^1_0(D)}.$$

Thus, for the a posteriori algorithm we have to evaluate u on a much larger grid than just $\Xi_{\tilde{\Lambda}_N}$, namely, $\Xi_{\tilde{\Lambda}_N} \cup \Xi_{\mathcal{N}(\tilde{\Lambda}_{N-1})}$, $\Xi_{\mathcal{N}(\tilde{\Lambda}_{N-1})} := \bigcup_{\boldsymbol{\nu} \in \mathcal{N}(\tilde{\Lambda}_{N-1})} \Xi^{(\boldsymbol{\nu})}$. We will refer to $\Xi_{\tilde{\Lambda}_N}$ as the a posteriori grid (associated with Λ_N) and to $\Xi_{\tilde{\Lambda}_N} \cup \Xi_{\mathcal{N}(\tilde{\Lambda}_{N-1})}$ as the extended grid (associated with Λ_N). The latter represents the "true" computational cost of the sparse collocation approximation generated by the a posteriori algorithm.

Remark 5.1. For our numerical simulations we choose a maximal number of parameter dimensions M, which may be arbitrarily large,[6] to construct the reference solution. Then, for a given $\boldsymbol{\xi} \in \mathbb{R}^M$ we approximate the solution $u(x, \boldsymbol{\xi})$ to (30) by evaluating its exact representation

$$u(x, \boldsymbol{\xi}) = \int_0^x \frac{K(\boldsymbol{\xi}) - F(y)}{a(y, \boldsymbol{\xi})}\, \mathrm{d}y, \qquad F(x) := \int_0^x f(y)\, \mathrm{d}y, \quad K(\boldsymbol{\xi}) := \frac{\int_0^1 \frac{F(y)}{a(y,\boldsymbol{\xi})}\, \mathrm{d}y}{\int_0^1 \frac{1}{a(y,\boldsymbol{\xi})}\, \mathrm{d}y},$$

by numerical quadrature, particularly the trapezoidal rule based on an equidistant spatial grid with spacing $\Delta x = 2^{-10}$.

5.3. Results. We now present the details and results of numerical experiments. The tests are divided into two parts: in the first set of experiments we aim at validating the sharpness of our analysis, i.e., whether we can actually observe numerically the rate predicted by Theorem 4.3 for the case of countably many random variables; in the second set of experiments, we will instead gradually increase the number of random variables and see if the observed rate of convergence is actually dimension-independent. Concerning the first set of experiments, we recall that the convergence results in Theorem 4.3 strictly apply only to the sparse collocation constructed by the a priori index selection algorithm. However, we will assess whether the set of indices generated by the a posteriori construction, i.e., the a posteriori grid, achieves the same rate and also examine the convergence rate with respect to the number of points in the extended grid.

[6] This is because the exact solution to the random PDE is known and inexpensively evaluated.

TABLE 1
Statistics for numerical results in Tests—Part I. *See* (32) *for the theoretical rates.*

q	% of total variance	Rate w.r.t. $\lvert\Lambda_N\rvert$			Rate w.r.t. $\lvert\Xi_{\Lambda_N}\rvert$		
		theory	a post.	a priori	theory	a post.	a priori
1	99.91%	N.A.	0.5	0.5	N.A.	0.5	0.4
1.5	99.9999%	0	0.9	0.8	0	0.8	0.7
2	99.9999999%	0.5	1.2	1.1	0.25	1.1	1.0
3	100%	1.5	2	2	0.75	1.7	1.7

Tests—Part I. In this section we compare the numerical convergence rate of both the a priori and the a posteriori versions of the proposed algorithm against the theoretical convergence rate for $q = 1, 1.5, 2, 3$ to verify the sharpness of our theoretical analysis. For each tested value of q, the errors will be computed against a reference solution u_{ref} based on the first 640 random variables which captures more than 99% of the log-diffusion variance for every value of q (see Table 1 for the precise value). The error is computed with a Monte Carlo sampling over $N_{MC} = 1000$ random samples:

$$\begin{aligned}\|u - U_{\Lambda_N} u\|_{L^2_\mu(\mathbb{R}^{\mathbb{N}};H^1_0(D))} &\approx \|u_{ref} - U_{\Lambda_N} u\|_{L^2_\mu(\mathbb{R}^{\mathbb{N}};H^1_0(D))} \\ &\approx \frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} \|u_{ref}(\boldsymbol{\xi}_k) - U_{\Lambda_N} u(\boldsymbol{\xi}_k)\|_{H^1_0(D)}, \end{aligned} \tag{35}$$

where $\boldsymbol{\xi}_k$ are samples drawn from $\bigotimes_{m=1}^{640} N(0,1)$. We remark that we have verified that N_{MC} is large enough for our purposes[7] and that the numerical algorithms never activate all 640 random variables during execution.

We begin by reporting in Figure 3 the convergence of the error measure (35) with respect to the number of collocation points needed to construct the sparse grid approximation for each value of q. The convergence plots in Figure 3 show a monotone, well-established decreasing trend for the error for all the variations of the sparse grid considered. As expected, the errors get larger in size and the convergence rate gets worse as q decreases for all the reported sparse grids (a posteriori grid, extended grid, a priori). In particular, the convergence rate appears to be similar for the a priori and the a posteriori algorithms, with the rate of the latter being actually slightly larger, thus validating the a posteriori construction. On top of this, the error of the a posteriori algorithm appears to be smaller in size than the a priori construction. We also remark that the rate that we measure numerically is better than the one predicted by our theory; cf. Table 1. The quite significant difference between the rate of convergence of the a posteriori grid and the extended grid is also to be expected. These results are consistent with those detailed in [9], although there the a priori construction is a bit different from the one we propose. At this junction, two factors can explain the suboptimality of our theoretical result: A conservative estimate of the growth of the number of points in the sparse grid with respect to the number of indices in the set Λ_N and a conservative link between the summability of the log-diffusion field representation and the convergence of the sparse grid. As will be clearer later, both issues turn out to actually affect our analysis. The numerical results we show were obtained with $m_{\text{buffer}} = 5$.[8]

[7] That is, repeating the same analysis with $N_{MC} = 5000$ produced identical results.

[8] We report (not shown) that we have also run the same simulations with a larger buffer $m_{\text{buffer}} = 20$ and the results were identical (i.e., same a posteriori grid and same number of activated random variables).

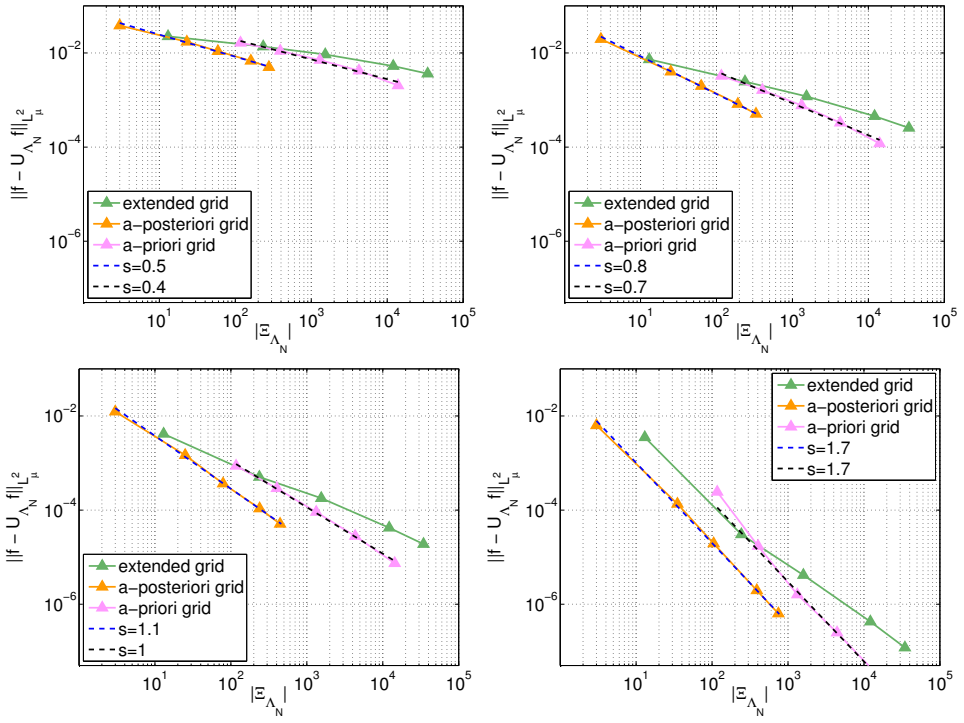

FIG. 3. *From top-left to bottom-right: Convergence with respect to the number of points in the sparse grid for* $q = 1, 1.5, 2, 3$.

We then report in Figure 4 the convergence of the error (35) with respect to the number of indices in the set Λ_N. In this figure, we show the convergence of both the a priori and the a posteriori algorithms, as well as an estimate of the convergence of the best N-term approximation of u (we detail below how this approximation was computed). Also in this case, the convergence plots show a monotone, well-established decreasing trend for the error. The results are similar to the previous case: (a) the convergence rate of the sparse grid gets worse as q decreases; (b) the convergence rate seems to be identical for both the a priori and the a posteriori constructions, and again quite larger than the theoretical estimate; cf. Table 1; (c) the error of the a posteriori algorithm is substantially smaller than the one of the a priori algorithm. It is also relevant to note that the measured convergence rate here is essentially identical to the one observed with respect to the number of sparse grid points. This is in agreement with the results in [9] and implies that for the sparse grids constructed here the growth of the number of points w.r.t. the number of indices is essentially linear, and therefore, our Lemma 3.16 is quite conservative.

The numbers in the plot show the number of activated random variables in the a posteriori grid and in the a priori grid, i.e., in how many random variables these grids allocate at least one nontrivial point (observe that by construction the numbers for the extended grid are the ones of the a posteriori grid plus the buffer m_{buffer}). It can be seen that this number steadily increases.

We now turn our attention to the best N-term approximation presented in Figure 4. To compute this approximation, we follow [19, 38, 43] and convert the extended grid first into its *combination technique* form, i.e., as a linear combination of Lagrange

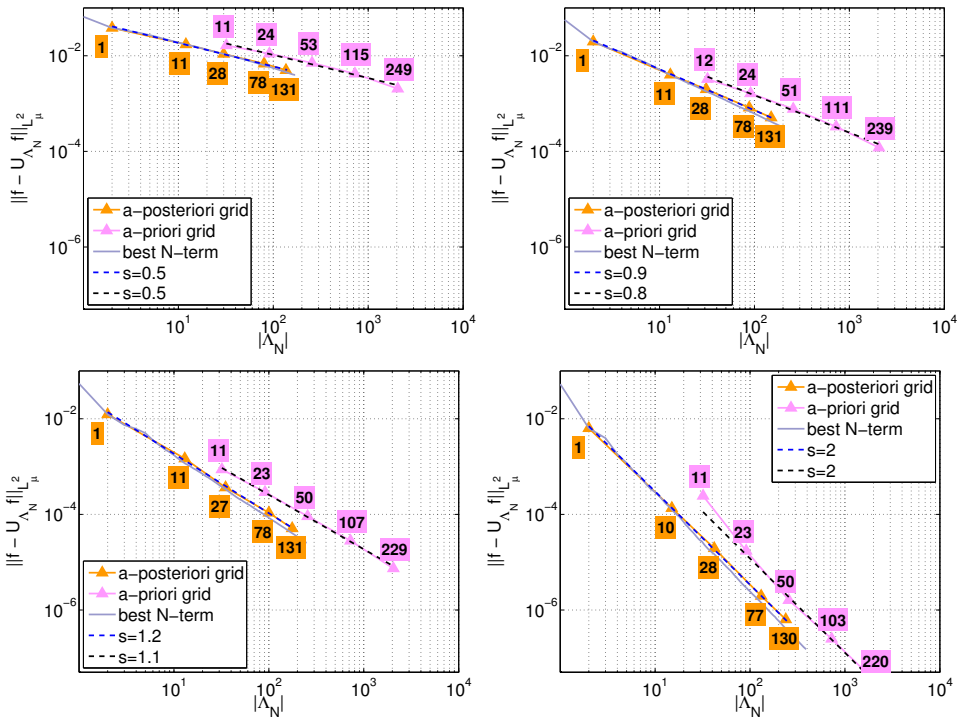

FIG. 4. *From top-left to bottom-right: Convergence with respect to the number of indices in the set* Λ_N *for* $q = 1, 1.5, 2, 3$.

polynomials, and then we further convert this expression into the equivalent linear combination of Hermite polynomials; see also [16]. By sorting in decreasing order the coefficients of the Hermite expansion thus computed and picking them one at a time, we obtain an approximation of the sequence of best N-term approximations.[9] The comparison of the best N-term and the a posteriori grid in Figure 4 reveals that the two approximations are actually very close for every value of q, which suggests that the a posteriori algorithm is producing an excellent approximation.

Tests—Part II. In this set of experiments, we fix $q = 2$ and we consider log-diffusion coefficients with $M = 10, 20, 40, 80, 120, 160$ random variables. For each M, the reference solution uses M random variables as well, contrary to the previous experiment, where the reference solution was based on 640 random variables. In this way, we aim at assessing the behavior of the convergence rate as M increases: indeed, the previous experiment was only intended to verify that we obtain a nonzero limiting rate for $M \to \infty$. We report our results in Figure 5, where we display the convergence with respect to the cardinality of the index set Λ_N. It is clearly visible that the convergence curves coincide initially and then depart from each other: The point of departure is actually that where all M variables have been activated. The result seems to suggest that the convergence rate with respect to the cardinality of Λ_N for finite M actually depends on M and decreases as M increases, until reaching

[9] Of course, this approximation is as good as the original extended grid; however, we found the results to be stable as the number of points in the extended grid grows, and therefore, we deemed this approximation to be sufficient for our purposes.

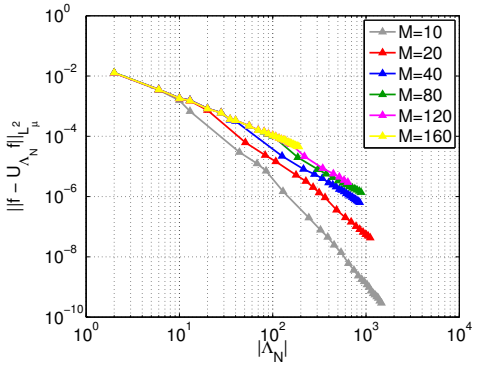

FIG. 5. *Convergence of the sparse grid approximation with an increasingly larger number of dimensions: The asymptotic rate is not constant with respect to M.*

the asymptotic rate for $M \to \infty$.

6. Conclusions. We have presented a general convergence analysis of sparse grid collocation based on Lagrange interpolation for functions of countably many Gaussian variables. In particular, we have stated sufficient conditions on the underlying univariate interpolation nodes such that for functions of a certain smoothness we obtain an algebraic rate of convergence for the sparse collocation approximation with respect to the number of multi-indices. Moreover, we verified these assumptions for the classical Gauss–Hermite nodes and were able to also state a convergence result in terms of the resulting number of collocation points. We finally discussed in detail that these methods can be applied to weak solutions of lognormal diffusion problems and illustrated our theory with numerical tests, which show that the convergence rate achieved by a priori sparse grid constructions is actually higher than predicted, both with respect to the number of multi-indices and the number of collocation points. The classical adaptive a posteriori sparse grid construction is also seen to achieve such rates, although it is not covered by our theory.

Acknowledgments. The authors are grateful to the anonymous reviewers for many valuable suggestions which improved our work and, in particular, for pointing out the simple proof of Proposition 3.16. The authors also thank Hans-Jörg Starkloff for providing the original reference to Stechkin's lemma.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, 10th ed., Natl. Bureau Standards Appl. Math. Ser. 55, U.S. Government Printing Office, Washington, DC, 1972.

[2] R. J. ADLER, *The Geometry of Random Fields*, John Wiley & Sons, New York, 1981.

[3] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Rev., 52 (2010), pp. 317–355, https://doi.org/10.1137/100786356.

[4] M. BACHMAYR, A. COHEN, R. DEVORE, AND G. MIGLIORATI, *Sparse polynomial approximation of parametric elliptic PDEs. Part* II*: Lognormal coefficients*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 341–363.

[5] M. BACHMAYR, A. COHEN, AND G. MIGLIORATI, *Sparse polynomial approximation of parametric elliptic PDEs. Part* I*: Affine coefficients*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 321–339.

[6] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: A numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations, Lect. Notes Comput. Sci. Eng. 76, Springer, Heidelberg, 2011, pp. 43–62.

[7] J. Beck, F. Nobile, L. Tamellini, and R. Tempone, *A quasi-optimal sparse grids procedure for groundwater flows*, in Spectral and High Order Methods for Partial Differential Equations - ICOSAHOM 2012, Lect. Notes Comput. Sci. Eng. 95, Springer, Cham, 2014, pp. 1–16.

[8] J. Charrier, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 216–246, https://doi.org/10.1137/100800531.

[9] P. Chen, *Sparse quadrature for high-dimensional integration with Gaussian measure*, ESAIM Math. Model. Numer. Anal., in press, 2018, https://doi.org/10.1051/m2an/2018012.

[10] A. Chkifa, *On the Lebesgue constant of Leja sequences for the complex unit disk and of their real projection*, J. Approx. Theory, 166 (2013), pp. 176–200.

[11] A. Chkifa, A. Cohen, and C. Schwab, *High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs*, Found. Comput. Math., 14 (2014), pp. 601–633.

[12] A. Chkifa, A. Cohen, and C. Schwab, *Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs*, J. Math. Pures Appl. (9), 103 (2015), pp. 400–428.

[13] A. Cohen and R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numer., 24 (2015), pp. 1–159.

[14] A. Cohen, R. DeVore, and C. Schwab, *Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs*, Found. Comput. Math., 10 (2010), pp. 615–646.

[15] A. Cohen, R. DeVore, and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, Anal. Appl. (Singap.), 9 (2011), pp. 11–47.

[16] P. Constantine, M. S. Eldred, and E. T. Phipps, *Sparse pseudospectral approximation method*, Comput. Methods Appl. Mech. Engrg., 229/232 (2012), pp. 1–12.

[17] S. De Marchi, *On Leja sequences: Some results and applications*, Appl. Math. Comput., 152 (2004), pp. 621–647.

[18] O. G. Ernst, B. Sprungk, and L. Tamellini, *Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Elliptic PDEs)*, preprint, https://arxiv.org/abs/1611.07239, 2017.

[19] L. Formaggia, A. Guadagnini, I. Imperiali, V. Lever, G. Porta, M. Riva, A. Scotti, and L. Tamellini, *Global sensitivity analysis through polynomial chaos expansion of a basin-scale geochemical compaction model*, Comput. Geosci., 17 (2013), pp. 25–42.

[20] J. Galvis and M. Sarkis, *Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity*, SIAM J. Numer. Anal., 47 (2009), pp. 3624–3651, https://doi.org/10.1137/080717924.

[21] A. Genz and B. D. Keister, *Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight*, J. Comput. Appl. Math., 71 (1996), pp. 299–309.

[22] T. Gerstner and M. Griebel, *Dimension-adaptive tensor-product quadrature*, Computing, 71 (2003), pp. 65–87.

[23] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.

[24] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, Heidelberg, 2001.

[25] C. J. Gittelson, *Stochastic Galerkin discretization of the log-normal isotropic diffusion problem*, Math. Models Methods Appl. Sci., 20 (2010), pp. 237–263.

[26] A.-L. Haji-Ali, F. Nobile, L. Tamellini, and R. Tempone, *Multi-index stochastic collocation convergence rates for random PDEs with parametric regularity*, Found. Comput. Math., 16 (2016), pp. 1555–1605.

[27] V. H. Hoang and C. Schwab, *N-term Wiener chaos approximation rates for elliptic PDEs with lognormal Gaussian random inputs*, Math. Models Methods Appl. Sci., 24 (2014), pp. 797–826.

[28] P. Jantsch, C. G. Webster, and G. Zhang, *On the Lebesgue constant of weighted Leja points for Lagrange interpolation on unbounded domains*, IMA J. Numer. Anal., in press, https://doi.org/10.1093/imanum/dry002, 2017.

[29] O. P. Le Maitre and O. M. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Scientific Computation, Springer, New York, 2010.

[30] F. Leja, *Une méthode élémentaire de résolution du problème de Dirichlet dans le plan*, Ann. Soc. Math. Polon., 23 (1950), pp. 230–245.
[31] A. Mugler and H.-J. Starkloff, *On the convergence of the stochastic Galerkin method for random elliptic partial differential equations*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 1237–1263.
[32] A. Narayan and J. D. Jakeman, *Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation*, SIAM J. Sci. Comput., 36 (2014), pp. A2952–A2983, https://doi.org/10.1137/140966368.
[33] P. G. Nevai, *Mean convergence of Lagrange interpolation*, II, J. Approx. Theory, 30 (1980), pp. 263–276.
[34] F. Nobile, L. Tamellini, and R. Tempone, *Convergence of quasi-optimal sparse-grid approximation of Hilbert-space-valued functions: Application to random elliptic PDEs*, Numer. Math., 134 (2016), pp. 343–388.
[35] F. Nobile, L. Tamellini, F. Tesei, and R. Tempone, *An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient*, in Sparse Grids and Applications – Stuttgart 2014, Lect. Notes Comput. Sci. Eng. 109, Springer, Cham, 2016, pp. 191–220.
[36] F. Nobile, R. Tempone, and C. G. Webster, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442, https://doi.org/10.1137/070680540.
[37] F. Nobile, R. Tempone, and C. G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345, https://doi.org/10.1137/060663660.
[38] G. Porta, L. Tamellini, V. Lever, and M. Riva, *Inverse modeling of geochemical and mechanical compaction in sedimentary basins through polynomial chaos expansion*, Water Resour. Res., 50 (2014), pp. 9414–9431.
[39] C. Schillings and C. Schwab, *Sparse, adaptive Smolyak quadratures for Bayesian inverse problems*, Inverse Problems, 29 (2013), 065011.
[40] C. Schwab and C. J. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer., 20 (2011), pp. 291–467.
[41] S. B. Stechkin, *On the absolute convergence of orthogonal series*, Doklady Akademii Nauk SSSR, 102 (1955), pp. 37–40.
[42] G. Szegő, *Orthogonal Polynomials*, 4th ed., AMS, Providence, RI, 1975.
[43] L. Tamellini, *Polynomial Approximation of PDEs with Stochastic Coefficients*, Ph.D. thesis, Politecnico di Milano, Milano, Italy, 2012.
[44] D. Xiu and J. S. Hesthaven, *High-order collocation methods differential equations with random inputs*, SIAM J. Sci. Comput., 37 (2005), pp. 1118–1139, https://doi.org/10.1137/040615201.