

Sparse tensor approximation of parametric eigenvalue problems*

R. Andreev and Ch. Schwab

Research Report No. 2010-40
November 2010

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

*Supported by SNF grant PDFMP2-127034/1 and by ERC AdG grant
STAHPDE 247277

**SPARSE TENSOR APPROXIMATION
OF
PARAMETRIC EIGENVALUE PROBLEMS***

ROMAN ANDREEV[†] AND CHRISTOPH SCHWAB[‡]

November 12, 2010

Abstract. We design and analyze algorithms for the efficient sensitivity computation of eigenpairs of parametric elliptic self-adjoint eigenvalue problems on high-dimensional parameter spaces. We quantify the analytic dependence of eigenpairs on the parameters. For the efficient approximate evaluation of parameter sensitivities of isolated eigenpairs on the entire parameter space we propose and analyze a sparse tensor spectral collocation method on an anisotropic sparse grid in the parameter domain. The stable numerical implementation of these methods is discussed and their error analysis is given. Applications to parametric elliptic eigenvalue problems with infinitely many parameters arising from elliptic differential operators with random coefficients are presented.

Key words. Stochastic partial differential equations, uncertainty quantification, finite element method, parametric eigenvalue problem, anisotropic Smolyak collocation, multilevel approximation, sparse tensor product

AMS subject classifications. primary: 65N30, secondary: 60H35, 65C30, 65N12, 65N15

*Supported by SNF grant PDFMP2-127034/1 and by ERC AdG grant STAHPDE 247277.

[†] Seminar for Applied Mathematics, ETH Zürich, Switzerland, (andreevr@math.ethz.ch)

[‡] Seminar for Applied Mathematics, ETH Zürich, Switzerland, (schwab@math.ethz.ch)

1. Introduction. Multi-parametric eigenvalue problems (EVPs) arise in numerous applications: we mention only engineering (parametric design optimization of the spectrum of structures in solids, fluids and electromagnetics), uncertainty quantification, stability analysis of engineering systems and the like. Other applications arise in the perturbation analysis in physics. Accordingly, there is a sizable body of references devoted to eigenvalue perturbation analysis. The mathematical theory of perturbation evolved in close connection to these applications; seminal works are by Rellich and Kato, see [27, 22] and references therein.

In recent years, much attention has been devoted to the computational analysis of so-called *complex systems*; in the context of the results in the present paper, such engineering systems could be deterministic initial boundary value problems depending on possibly a large number of design parameters. Alternatively, one could consider spectral problems for partial differential equations with random field input such as, e.g. diffusion problems with random heat conductivity. Adopting parametric representations of input random fields (e.g. by Karhunen-Loève expansions) renders the EVP of interest deterministic but depending on possibly a *countable number of parameters*. We develop the numerical analysis in an abstract setting, where we consider parametric operator families $A(y) \in \mathcal{L}(V, V^*)$ depending on a vector of real parameters $y = (y_1, y_2, \dots)$, which cover differential- and integral operators acting in separable Hilbert spaces V and also parametric matrices.

2. Parametric eigenvalue problems. We present a class of abstract eigenvalue problems for parametric self-adjoint families of operators with real-analytic dependence on a vector of parameters and discuss the dependence of their eigenpairs on these parameters. This will be the foundation of the design and the analysis of sparse tensor approximation methods in subsequent sections.

We will specialize on operators of the particular form

$$A(y) = \bar{A} + y_1 B_1 + y_2 B_2 + \dots$$

with a self-adjoint “principal part” \bar{A} . In first order approximation, the affine dependence of $A(y)$ on the parameter vector $y = (y_1, y_2, \dots)$ will also appear in the case of general smooth nonlinear dependence $y \mapsto A(y)$.

Specific examples are also provided in the next section. In particular, operators depending on a countable number of parameters arise in applications such as PDEs with spatially inhomogeneous random coefficients.

2.1. Variational eigenvalue problems. Let V and H be separable Hilbert spaces over \mathbb{R} (or \mathbb{C}) with inner products $(\cdot, \cdot)_V$ and $(\cdot, \cdot)_H$ and norms $\|\cdot\|_V$ and $\|\cdot\|_H$, respectively. Unless stated otherwise, V and H are assumed to be infinite-dimensional. We assume V and H form the *Gel'fand triple* $V \subset H \cong H^* \subset V^*$ with dense and compact injections, where $H \cong H^*$ indicates identification of the “pivot space” H with its dual H^* . By $\langle \cdot, \cdot \rangle_{V \times V^*}$ we denote the duality pairing on V and V^* . Let $b : V \times V \rightarrow \mathbb{C}$ be a bilinear (or sesquilinear) form for which there exist constants $\gamma > 0$ and $C_1 > 0$ such that

$$\forall u, v \in V : |b(u, v)| \leq C_1 \|u\|_V \|v\|_V, \quad (2.1)$$

$$\inf_{0 \neq u \in V} \sup_{0 \neq v \in V} \frac{|b(u, v)|}{\|u\|_V \|v\|_V} \geq \gamma > 0 \quad (2.2)$$

and

$$\forall 0 \neq v \in V : \sup_{u \in V} |b(u, v)| > 0. \quad (2.3)$$

We denote by $A \in \mathcal{L}(V, V^*)$ the operator corresponding to the form $b(\cdot, \cdot)$ via the identification $b(u, v) = \langle v, Au \rangle_{V \times V^*}$ for all $u, v \in V$. Then (2.1)–(2.3) imply that both, A and its adjoint A^* are isomorphisms between the space V and its dual V^* .

We call $\lambda \in \mathbb{C}$ *eigenvalue* of the form $b(\cdot, \cdot)$ if there exists an *eigenvector* $0 \neq w \in V$ associated to λ , such that

$$\forall v \in V : b(w, v) = \lambda(w, v)_H, \quad (2.4)$$

in which case the pair $(\lambda, w) \in \mathbb{C} \times V$ is called *eigenpair* of $b(\cdot, \cdot)$.

By $\sigma(A)$ we denote the spectrum of an operator A [25, Ch. VI]. Conditions (2.1)–(2.3) and compactness of the embedding $V \subset H$ imply the existence of a unique compact linear operator $T \in \mathcal{L}(V, V)$ such that

$$\forall u, v \in V : b(Tu, v) = (u, v)_H. \quad (2.5)$$

The pair $(\lambda, w) \in \mathbb{C} \times V$ satisfies (2.4) if and only if $\lambda Tw = w \neq 0$, i.e., if and only if the pair (λ^{-1}, w) is an eigenpair of the compact operator T . Note that by (2.2), the eigenvalue λ is non-zero.

Let $\mu \in \sigma(T)$, $\mu \neq 0$. The number $\lambda = \mu^{-1}$ is an eigenvalue of the form $b(\cdot, \cdot)$. The smallest integer α such that $\text{Ker}((\mu - T)^\alpha) = \text{Ker}((\mu - T)^{\alpha+1})$ is called *ascent* of $\mu - T$. The dimension $m = \dim \text{Ker}((\mu - T)^\alpha)$ is finite and is called *algebraic multiplicity* of λ . Vectors in $\text{Ker}((\mu - T)^\alpha)$, *the generalized eigenspace of T corresponding to λ* are called *generalized eigenvectors of T corresponding to λ* . The *geometric multiplicity* of μ is equal to $\dim \text{Ker}(\mu - T)$. The b -adjoint of T , denoted by T_* is defined by $b(Tu, v) = b(u, T_*v)$ for all $u, v \in V$. A pair $(\lambda, v) \in \mathbb{C} \times V$ is called *adjoint eigenpair of the form $b(\cdot, \cdot)$* if and only if $(\bar{\lambda}^{-1}, v)$ is an eigenpair of T_* , i.e., $v \neq 0$ and $b(u, v) = \lambda(u, v)_H$ for all $u \in V$. In this case v is called *adjoint eigenvector corresponding to λ* . *Generalized adjoint eigenvectors of T* are exactly the generalized eigenvectors of T_* . An eigenvalue $\lambda \in \sigma(A)$ is called *isolated* if $\text{dist}(\lambda, \sigma(A) \setminus \{\lambda\}) > 0$. It is called *discrete* if it is isolated and if for self-adjoint A : it is of finite multiplicity, i.e., $\dim\{u \in V : Au = \lambda u\} < \infty$ see [25, Theorem VII.10]; for non-selfadjoint A : the spectral projection $P_\lambda = -\frac{1}{2\pi i} \oint_{|\mu-\lambda|=r} (A-\mu)^{-1} d\mu$, is finite dimensional [26, Ch. XII]. An eigenvalue $\lambda \in \sigma(A)$ is called *nondegenerate* if the respective dimension equals one.

2.2. Abstract parametric eigenvalue problems. We consider a family of *real, parametric* eigenvalue problems, and *assume until further specification that V and H are Hilbert spaces over \mathbb{R}* (rather than over \mathbb{C}). Assume we are given a family of bounded self-adjoint operators $A(y) \in \mathcal{L}(V, V^*)$ parameterized by a vector $y = (y_1, y_2, \dots)$ of real numbers, which we assume to take values in bounded intervals, after rescaling $y_m \in [-1, 1]$. In many applications, we deal with a finite, but possibly large number M of parameters, whereas for applications from elliptic PDEs with random coefficients we allow countably many parameters. Accordingly, we assume $y \in U$, where

$$U = \begin{cases} [-1, 1]^M, & M < \infty, \\ [-1, 1]^{\mathbb{N}}, & M = \infty. \end{cases} \quad (2.6)$$

For $M = \infty$ the summation “ $\sum_{m=1}^M$ ” is understood as (unconditionally) convergent infinite sum. For U we consider a Hausdorff topology, i.e., the euclidean topology if U is finite-dimensional; if U is infinite-dimensional we equip U with the topology and metric of $\ell^\infty(\mathbb{N})$. This setting fits the abstract framework of [20]. For all $y \in U$ we associate to $A(y)$ the bilinear form

$$b(y; \cdot, \cdot) : V \times V \rightarrow \mathbb{R}, \quad (u, v) \mapsto b(y; u, v) := \langle u, A(y)v \rangle_{V \times V^*} \quad (2.7)$$

We assume that $y \mapsto A(y)$ is uniformly bounded on U :

$$\forall u, v \in V : \quad \langle u, A(y)v \rangle_{V \times V^*} \leq \sup_{y \in U} \|A(y)\|_{\mathcal{L}(V, V^*)} \|u\|_V \|v\|_V, \quad (2.8)$$

with $\sup_{y \in U} \|A(y)\|_{\mathcal{L}(V, V^*)} < \infty$, and uniformly elliptic on U :

$$\exists \alpha > 0 : \quad \forall y \in U : \quad \forall v \in V : \quad \langle v, A(y)v \rangle_{V \times V^*} \geq \alpha \|v\|_V^2. \quad (2.9)$$

This implies that for every $y \in U$ the operator $A(y)$ is boundedly invertible, i.e., for its inverse holds $\|A^{-1}(y)\|_{\mathcal{L}(V^*, V)} \leq \alpha^{-1}$. The compactness of the embedding $V \hookrightarrow H$ implies that the parametric EVP: given $y \in U$, find

$$\lambda(y) \in \mathbb{R} \quad \text{and} \quad 0 \neq w(y) \in V \quad \text{s.t.} \quad A(y)w(y) = \lambda(y)w(y) \quad (2.10)$$

or, in variational form: given $y \in U$, find

$$\lambda(y) \in \mathbb{R}, \quad 0 \neq w(y) \in V : \quad \forall v \in V : \quad b(y; w(y), v) = \lambda(y) (w(y), v)_H \quad (2.11)$$

admits, for every $y \in U$, countably many *real* eigenvalues $(\lambda_j(y))_{j \geq 1} \subset \mathbb{R}$ of finite multiplicity. *Here, and in the following, we always assume the eigenvalue sequences to be numbered in increasing magnitude, counting multiplicities, i.e., an eigenvalue of multiplicity k is listed k times.* The corresponding set of eigenfunctions $\{w_j(y)\}_{j \geq 1} \subset V$ forms a countable dense set in V , and therefore, by compact and dense embedding $V \hookrightarrow H$, we assume w.l.o.g. that for every $y \in U$ the sequence $(w_j(y))_{j \geq 1}$ forms a countable orthonormal basis of H .

2.3. Analyticity. We are particularly interested in the case where the dependence of $A(y)$ on the parameters y_m is *analytic* in suitable sense and, more specifically, in the case when the dependence of $A(y)$ on each coordinate y_m is affine, possibly after linearization of $A(y)$ given smooth dependence on the parameter vector y : there exist $\bar{A}, B_m \in \mathcal{L}(V, V^*)$, $m \geq 1$, such that

$$\forall y = (y_1, y_2, \dots) \in U : \quad A(y) = \bar{A} + \sum_{m=1}^M y_m B_m \quad (2.12)$$

with convergence in $\mathcal{L}(V, V^*)$.

REMARK 2.1. *To ensure coercivity (2.9) of $A(y)$ in (2.12), it is sufficient that in (2.12) the “mean operator” \bar{A} and the “fluctuations” B_m satisfy*

$$\exists \bar{\alpha} > 0 \quad \text{and} \quad \exists \kappa < 1 : \quad \begin{cases} \langle v, \bar{A}v \rangle_{V \times V^*} \geq \bar{\alpha} \|v\|_V^2 \quad \forall v \in V, \\ \sum_{m=1}^M \|B_m\|_{\mathcal{L}(V, V^*)} \leq \kappa \bar{\alpha}. \end{cases} \quad (2.13)$$

Indeed, condition (2.13) implies (2.9) with $\alpha = \bar{\alpha}(1 - \kappa) > 0$: for any $y \in U$ and $v \in V$ we have

$$b(y; v, v) \geq b(0; v, v) - \left(\sum_{m=1}^M \|B_m\|_{\mathcal{L}(V, V^*)} \right) \|v\|_V^2 \geq \bar{\alpha}(1 - \kappa) \|v\|_V^2.$$

As we will show, under certain conditions the eigenpairs $(\lambda_m(y), w_m(y))_{m \geq 1}$ depend *analytically* on the parameter vector y . To make this precise, we first recall definitions and facts on Hilbert space valued analytic functions. To this end, *from now on we assume that V and H are complex separable Hilbert spaces* and extend all inner products and duality pairings sesquilinearly for complex valued arguments. We emphasize that this extension is for purposes of analysis only; the parametric eigenvalue problems under consideration here are *real and self-adjoint* (for non-selfadjoint operators some of our results require essential modifications in their statement and proof).

We recall some definitions. Assume initially that $M < \infty$. Let X be a Banach space over \mathbb{C} and let $E \subset \mathbb{C}^M$ be an open, bounded and connected domain, $M \in \mathbb{N}$. A function $x : E \rightarrow X$ is said to be

- *(strongly jointly) analytic in E* if for each $a \in E$ there is $\{c_k\}_{k \in \mathbb{N}_0^M} \subset X$ such that the Taylor series $\sum_{k \in \mathbb{N}_0^M} c_k \prod_{m=1}^M (z_m - a_m)^{k_m}$ is summable to $x(z)$ for $z \in E$ sufficiently close to a .
- *(strongly) holomorphic in E* if for each $a \in E$ each first order partial derivative $\lim_{0 \neq h \rightarrow 0} (x(a + he_m) - x(a))/h$ exists in X where $e_m \in \mathbb{C}^M$ is the m -th standard unit vector.
- *weakly analytic in E* if for each $\ell \in X^*$ the function $\ell(x(\cdot))$ is a \mathbb{C} -valued analytic function in E (equivalently, holomorphic by Hartogs' theorem [21, Theorem 2.2.8]).

For X endowed with a locally convex sequentially complete topology (e.g. Banach space) these notions coincide, see [20, Theorem 2.1.3], also [25, Theorem VI.4].

In the case $M = \infty$ we call $x : E \rightarrow X$ *(jointly) analytic* on a set $E \subset \mathbb{C}^M$ which is open w.r.t. $\ell^\infty(\mathbb{N})$, if X is locally convex sequentially complete and if the series

$$f(z) = f(a) + \sum_{\nu \in \mathbb{N}_0^{\mathbb{N}}} \frac{1}{\nu!} (D_a^\nu f)(z - a)$$

is uniformly summable for $z - a$ in any compact subset of the largest *balanced* subset of the set $E - a$ (here, $\tilde{E} \subset E - a$ is called balanced if $z \in \tilde{E} \wedge |\zeta| \leq 1 \Rightarrow \zeta z \in \tilde{E}$).

2.3.1. Case $M = 1$. The case $M = 1$ is of independent interest, and also serves as a building block for the multiparameter case. Therefore, we recapitulate the pertinent results here. For single parameter, regular¹ analytic spectral perturbation theory we refer to [22, Chapters I and II] and [26, Chap. XII]. For $M = 1$, the operator $A(y)$ in (2.12) takes the form

$$A(y) = \bar{A} + yB, \quad y \in U = [-1, 1]. \quad (2.14)$$

We can extend $A(\cdot)$ to an entire, operator-valued function by allowing $y \in \mathbb{C}$. In this case, the dependence of the eigenpairs $(\lambda_j(y), w_j(y))_{j \geq 1}$ on the parameter y is well understood. Although the dependence of $A(y)$ on y in (2.14) is analytic, the eigenpairs of $A(y)$ do not necessarily inherit this analytic dependence:

¹as opposed to ‘‘asymptotic’’, see [26], Sect. XII.2 and XII.3

EXAMPLE 2.2 ([22], Sect. II.1 or [26], Sect. XII.1). For $V = \mathbb{C}^2$ consider

$$A(y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + z \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathcal{L}(V, V^*).$$

Then $A(z)$ has eigenvalues $\lambda_{\pm}(z) = \pm\sqrt{1+z^2}$ and is real and self-adjoint for real z . Evidently, for $|z| < 1$, the eigenvalues are complex-analytic functions of z . However, even though the map $z \mapsto A(z)$ is entire analytic and $\mathbb{R} \ni z \mapsto \lambda_{\pm}(z)$ are real-analytic functions, the maps $\mathbb{C} \ni z \mapsto \lambda_{\pm}(z)$ exhibit singularities as complex-analytic functions at $z = \pm i$. Note that $A(z)$ is symmetric for $\text{Im } z = 0$, but is not Hermitean for any $\text{Im } z \neq 0$, and even not diagonalizable for $z = \pm i$.

We now consider in (2.10) the parametric operator $A(y)$ as restriction to U of the analytic, $\mathcal{L}(V, V^*)$ -valued function

$$A : \mathbb{C} \rightarrow \mathcal{L}(V, V^*), \quad z \mapsto A(z) := \bar{A} + zB. \quad (2.15)$$

Notice carefully that the extension of the self adjoint $A(y)$, $y \in \mathbb{R}$ obtained in this fashion is *not* necessarily Hermitean (cf. Example 2.2).

We will work under assumption (2.13), which in the case $M = 1$ becomes the following.

ASSUMPTION 2.3. *There exist $\bar{\alpha} > 0$ and $\kappa < 1$ s.t.*

$$\text{Re}\langle v, \bar{A}v \rangle_{V \times V^*} \geq \bar{\alpha} \|v\|_V^2 \quad \forall v \in V \quad \text{and} \quad \|B\|_{\mathcal{L}(V, V^*)} \leq \kappa \bar{\alpha}. \quad (2.16)$$

Under (2.16), the variational form of the eigenvalue problem for $A(z)$ in (2.15) satisfies the assumptions (2.1)–(2.3):

PROPOSITION 2.4. *Assume (2.16). Then, for any $0 < \delta < 1$ and all $z \in \mathbb{C}$ with $|z| \leq \kappa^{-1}\delta$ the sesquilinear form $b(z; \cdot, \cdot) : V \times V \rightarrow \mathbb{C}$, $b(z; u, v) := \langle u, A(z)v \rangle_{V \times V^*}$ satisfies (2.1)–(2.3) with*

$$\gamma = \bar{\alpha}(1 - \delta) > 0, \quad C_1 = \|\bar{A}\|_{\mathcal{L}(V, V^*)} + \kappa^{-1}\delta \|B\|_{\mathcal{L}(V, V^*)}, \quad (2.17)$$

in particular, for any $\kappa < \delta$. Moreover, for all $z \in \mathbb{C}$ with $|z| \leq \kappa^{-1}\delta$ the operator $A(z)$ is boundedly invertible with $\sup_{|z| \leq \kappa^{-1}\delta} \|A^{-1}(z)\|_{\mathcal{L}(V^, V)} \leq (\bar{\alpha}(1 - \delta))^{-1}$.*

By compactness of the embedding $V \hookrightarrow H$, for every $y \in U$, the spectrum $\sigma(A(y))$ of the self-adjoint operator $A(y)$ is discrete and consists of at most countably many real eigenvalues $\lambda_j(y)$, $j = 1, 2, \dots$ of finite multiplicity which accumulate only at infinity. As $A(y)$ is symmetric, the algebraic and geometric multiplicities of $\lambda_j(y)$ coincide.

We are interested in the relation of $\lambda(y) \in \sigma(A(y))$ to eigenvalues $\lambda(z) \in \mathbb{C}$ of the complex extension $A(z)$ in (2.15) of $A(y)$.

THEOREM 2.5. *For $z \in \mathbb{C}$, consider the family (2.15) of linear operators where \bar{A}, B satisfy (2.16). Fix $y \in U = [-1, 1]$. Let $\lambda(y) \in \sigma(A(y)) \subset \mathbb{R}$ be a discrete eigenvalue of $A(y)$ of multiplicity $m \in \mathbb{N}$. Then the following holds:*

- a) *There exist m (not necessarily distinct) complex-valued functions of z which are single-valued and analytic near $z = y$, denoted by $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(m)}$, such that $\lambda^{(j)}(y) = \lambda(y)$, $j = 1, \dots, m$ and $\lambda^{(j)}(z)$, $j = 1, \dots, m$ are discrete eigenvalues of $A(z)$ near $z = y$,*
- b) *there are no other eigenvalues of $A(z)$ near $\lambda(z)$,*
- c) *there are m complex-analytic V -valued functions $w^{(1)}, \dots, w^{(m)}$, such that $w^{(1)}(z), \dots, w^{(m)}(z)$ are corresponding eigenvectors of $A(z)$,*
- d) *the domains of analyticity contain discs $\{z \in \mathbb{C} : |z - y| < \epsilon\}$ where $\epsilon = \epsilon(\kappa) \sim \kappa^{-1}$ as $\kappa \rightarrow 0$.*

Proof. Parts *a)* and *b)* is Theorem XII.13 in [26]. Part *c)* is Problem 17 in Sect. XII of [26] applied to the projector $P(\beta)$ with $\beta = z - y$ in Theorem XII.13 of [26]. Part *d)* follows by a scaling argument. \square

The following quantitative bound on the parameter range ensuring ellipticity will be useful later for obtaining uniform bounds on convergence radii.

THEOREM 2.6. *Assume (2.16). Fix $y \in U$. Let $\lambda(y) \in \sigma(A(y))$ be isolated and nondegenerate. Define the spectral gap $\gamma = \gamma(y, \lambda(y)) = \text{dist}(\lambda(y), \sigma(A(y)) \setminus \{\lambda(y)\})$. Then there exists $z \mapsto \lambda_y(z) \in \sigma(A(z))$ with $\lambda_y(y) = \lambda(y)$ which is complex-analytic in the disc $E(y, \kappa, \gamma/\lambda(y)) = \{z \in \mathbb{C} : |z - y| < \frac{1}{2} \frac{\kappa^{-1} - |y|}{1 + \lambda(y)/\gamma}\}$. Moreover, $\lambda_y(z) \in \sigma(A(z))$ is isolated and nondegenerate for each $z \in E(y, \kappa, \gamma/\lambda(y))$.*

Proof. We have $\|Bv\|_{V^*} \leq \kappa \bar{\alpha} \|v\|_V \leq \frac{\kappa}{1 - \kappa|y|} \|A(y)v\|_{V^*}$ for all $v \in V$ by coercivity $\|A(y)v\|_{V^*} \geq \bar{\alpha}(1 - |y|\kappa) \|v\|_V$. Theorem XII.8 of [26] (Kato-Rellich) and Theorem XII.11 of [26] show the claimed analyticity inside the circle of radius

$$r = [a + \varepsilon^{-1}[b + a(\lambda(y) + \varepsilon)]]^{-1} = \frac{1}{2} \frac{\kappa^{-1} - |y|}{1 + \lambda(y)/\gamma} \quad (2.18)$$

with $a = \frac{\kappa}{1 - \kappa|y|}$, $b = 0$ and $\varepsilon = \frac{1}{2}\gamma$. \square

REMARK 2.7. *Note that*

$$\frac{1}{1 + \delta^{-1}} = \begin{cases} \delta + \mathcal{O}(\delta^2), & \text{for } \delta \rightarrow 0, \\ 1 - \delta^{-1} + \mathcal{O}(\delta^{-2}), & \text{for } \delta \rightarrow \infty. \end{cases} \quad (2.19)$$

Thus, as the ratio $\delta = \gamma/\lambda(y)$ becomes small, the size of the domain analyticity is critically restricted by γ being relatively small. If γ is large compared to $\lambda(y)$, the size of the domain of analyticity is essentially given by κ^{-1} , cf. Theorem 2.5.

An analytic continuation argument yields the “ y -uniform” version of Theorem 2.6.

COROLLARY 2.8. *Suppose that $U \ni y \mapsto \lambda(y) \in \sigma(A(y))$ is continuous and such that for each $y \in U$, $\lambda(y)$ is isolated and nondegenerate. Let $\delta > 0$ be such that*

$$\forall y \in U : \quad \text{dist}(\lambda(y), \sigma(A(y)) \setminus \{\lambda(y)\}) \geq \delta \lambda(y).$$

Then we can extend $y \mapsto \lambda(y)$ to a functions $z \mapsto \lambda(z)$ which is analytic on $E(\kappa, \delta) = \bigcup_{y \in U} E(y, \kappa, \delta)$, s.t. $\lambda(z) \in \sigma(A(z))$ is isolated and nondegenerate for all $z \in E(\kappa, \delta)$.

Proof. We need the following special case of Lemma 2.11 below: for all $y \in U$ there exists $r > 0$ s.t. for any two continuous functions $f_1, f_2 : B_r(y) \rightarrow \sigma(A(z))$ holds

$$f_1(y) = \lambda(y) = f_2(y) \quad \forall y \in B_r(y) \cap U \quad \implies \quad f_2(z) = \lambda(z) \quad \forall z \in B_r(y),$$

where $B_r(y) = \{z \in \mathbb{C} : |z - y| < r\}$.

We only sketch the rest of the proof of the theorem, omitting the technical details. Let $U \subset K_1 \subset K_2 \subset \dots \subset E(\kappa, \delta)$ be a monotonic sequence of compact sets such that $E(\kappa, \delta) = \bigcup_{n \in \mathbb{N}} K_n$. On every K_n an extension λ_n of λ can be constructed using Theorem 2.6 which is complex-analytic on any open $B \subset K_n$ and unique by Lemma 2.11. Moreover, all λ_n agree on their respective domains of definition, giving rise to a complex-analytic extension of λ to $E(\kappa, \delta)$, again unique by Lemma 2.11. \square

2.3.2. Case $M \geq 1$. In generalizing the above results to $M \geq 2$ parameters, care is necessary, as the following example due to Rellich shows, cf. [26, p. 60].

EXAMPLE 2.9. *Consider the two-parameter family of symmetric 2×2 matrices*

$$A(y_1, y_2) = y_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + y_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then A is linear, hence entire analytic in y_1 and y_2 , and symmetric for real y ; however, its eigenvalues $\lambda(y_1, y_2) = \pm\sqrt{y_1^2 + y_2^2}$ are not real-analytic with respect to y_1, y_2 in any vicinity of the origin. Notice also that $A(0, 0)$ has a double eigenvalue 0.

In this section we derive sufficient conditions which preclude this kind of singularity. For technical reasons we focus on eigenvalues which are isolated and nondegenerate. The following theorem is a local result on holomorphic dependence of an isolated and nondegenerate eigenvalue on the parameters.

THEOREM 2.10. *Suppose $z \in \mathbb{C}^M$ is such that $A(z)$ has an isolated and nondegenerate eigenvalue $\lambda(z)$. Pick $m \in \mathbb{N}$, and write $e_m = (0, \dots, 0, 1, 0, \dots)$, nonzero in m -th coordinate. Then $\exists \varepsilon_m(z) > 0$ s.t. for $E_m := \{\zeta_m \in \mathbb{C} : |\zeta_m| < \varepsilon_m(z)\}$ holds:*

1. *there exists a unique complex-analytic function*

$$E_m \ni \zeta_m \mapsto \lambda(z + e_m \zeta_m) \in \sigma(A(z + e_m \zeta_m)).$$

2. *$\lambda(z + e_m \zeta_m)$ is isolated and nondegenerate for all $\zeta_m \in E_m$.*
3. *there exists a complex-analytic V -valued function*

$$E_m \ni \zeta \mapsto w(z + e_m \zeta_m) \in V$$

such that $w(z + e_m \zeta_m)$ is a corresponding eigenvector for all $\zeta_m \in E_m$.

Proof. Note first that for any $r > 0$ the complex-analytic operator-valued function $\zeta \mapsto A(z + e_m \zeta)$ is uniformly bounded on the closed disk $\{\zeta \in \mathbb{C} : |\zeta| \leq r\}$. Thus, by [26, Exercise XII.8], the Kato-Rellich theorem [26, Theorem XII.8] applies. \square

We now prove the following lemma ensuring uniqueness of the extension.

LEMMA 2.11. *Let $B \subset U$ be connected by polygonal paths in the following sense: for any $z_a, z_b \in B$ there exist $n \in \mathbb{N}$, $z_0, \dots, z_n \in B$ with $z_0 = z_a$, $z_n = z_b$ and*

$$[z_{k-1}, z_k] := \{tz_{k-1} + (1-t)z_k : t \in [0, 1]\} \subset B \quad \text{for all } k = 1, \dots, n.$$

Take continuous functions $f, g : B \rightarrow \mathbb{C}$ with $f(z), g(z) \in \sigma(A(z))$ for all $z \in B$ such that $f(z)$ and $g(z)$ are isolated and nondegenerate for all $z \in B$. If $f(z) = g(z)$ for some $z \in B$, then $f(z) = g(z)$ for all $z \in B$.

Proof. Suppose to the contrary that for some $z_a, z_b \in B$ we have $f(z_a) = g(z_a)$ and $f(z_b) \neq g(z_b)$. Let $\phi : [0, 1] \rightarrow B$ be a continuous function with $\phi(0) = z_a$ and $\phi(1) = z_b$. Without loss of generality we can assume that

- $\phi(t) = z_a + t(z_b - z_a)$,
- $\forall \epsilon > 0$ there exists $t \in (0, \epsilon)$ with $f(\phi(t)) \neq g(\phi(t))$ (to this end notice that $\{z \in B : f(z) = g(z)\}$ is closed in B),
- the family of operators $\zeta \mapsto \tilde{A}(\zeta) = A(z_a + \zeta(z_b - z_a))$, $|\zeta| < 2$ is well-defined and is complex-analytic.

By the Kato-Rellich theorem [26, Theorem XII.8], there exists exactly one point $E(\zeta) \in \sigma(\tilde{A}(\zeta))$ close to $E(0) = f(z_a) = g(z_a)$ for ζ small enough, w.l.o.g. for $|\zeta| < 2$ by rescaling $(z_b - z_a)$. But this is a contradiction to $f(z_b) \neq g(z_b)$ if we set $\zeta = 1$. \square

In the following (Theorem 2.13) we identify the range of $z \in \mathbb{C}^M$ close to the parameter set U such that the conditions of Theorem 2.10 apply. We start with a Lemma.

LEMMA 2.12. *Fix $y \in U = [-1, 1]^M$ and $\zeta \in \mathbb{C}^M$. Assume that for some $p \in (0, 1]$*

$$B(p) := \sum_{m \geq 1} \|B_m\|_{\mathcal{L}(V, V^*)}^p \tag{2.20}$$

is finite. Assume $\kappa := \frac{1}{\alpha} B(p) \sup_{m \geq 1} |\zeta_m| \|B_m\|_{\mathcal{L}(V, V^*)}^{1-p} < \infty$, where $\alpha > 0$ is as in (2.9). Then

$$\left\| \sum_{m \geq 1} \zeta_m B_m \right\|_{\mathcal{L}(V, V^*)} \leq B(p) \sup_{m \geq 1} |\zeta_m| \|B_m\|_{\mathcal{L}(V, V^*)}^{1-p}, \quad (2.21)$$

and moreover, for $\bar{\alpha} := \alpha$, κ , $\bar{A} := A(y)$ and $B := \sum_{m \geq 1} \zeta_m B_m$ we have (2.16).

Proof. Equation (2.21) follows by the triangle and the Hölder inequalities. Thus, $\|B\|_{\mathcal{L}(V, V^*)} \leq \kappa \bar{\alpha}$ and (2.16) holds by the assumption that $\bar{\alpha} = \alpha$ and $\bar{A} = A(y)$. \square

The results of this section are now combined in the following theorem.

THEOREM 2.13. *For the family of operators $A(y)$, $y \in U$ assume that i) $A(y)$ is of the form (2.12), ii) (2.8) and (2.9) hold, iii) for a $p \in (0, 1]$ that the perturbations are p -summable as in (2.20), iv) $\lambda(y) \in \sigma(A(y))$ is an isolated and nondegenerate eigenvalue for all $y \in U$ with corresponding eigenfunction $w(y) \in V$, $\|w(y)\|_H = 1$ and v) the function $U \ni y \mapsto (\lambda(y), w(y)) \in \mathbb{R} \times V$ is continuous. Finally, assume that vi) there exists $\delta > 0$ such that $\text{dist}(\lambda(y), \sigma(A(y)) \setminus \{\lambda(y)\}) \geq \delta \lambda(y)$ for all $y \in U$.*

Let $\varepsilon \in (0, 1)$. Define $\tau_m := (1 - \varepsilon) \frac{\alpha \|B_m\|_{\mathcal{L}(V, V^*)}^{p-1}}{2B(p)(1+\delta^{-1})}$, $m \geq 1$ and

$$E(\tau) := \{z \in \mathbb{C}^M : \text{dist}(z_m, [-1, 1]) < \tau_m\}. \quad (2.22)$$

Then (λ, w) can be extended to a jointly complex-analytic function on $E(\tau)$.

Proof. For any $z \in E(\tau)$ we first identify a candidate $\lambda(z) \in \sigma(A(z))$ which is isolated and nondegenerate. Fix $z \in E(\tau)$. Take $y = y(z) \in U$ with $|z_m - y_m| < \tau_m$, $m \geq 1$. Define $\zeta = z - y$. Obviously, $|\zeta_m| < \tau_m$ and therefore

$$B(p) \sup_{m \geq 1} |\zeta_m| \|B_m\|_{\mathcal{L}(V, V^*)}^{1-p} \leq (1 - \varepsilon) \frac{\alpha}{2(1 + \delta^{-1})}.$$

Thus, for $\kappa := \frac{1}{\alpha} B(p) \sup_{m \geq 1} |\zeta_m| \|B_m\|_{\mathcal{L}(V, V^*)}^{1-p}$ we have $(1 - \varepsilon)^{-1} \leq \frac{1}{2\kappa} \frac{1}{1 + \delta^{-1}}$. Consider the complex-analytic operator-valued function $t \mapsto A(y + t\zeta) = A(y) + t \sum_{m \geq 1} \zeta_m B_m$. By Corollary 2.8 and Lemma 2.12 there exists a complex-valued function $t \mapsto \tilde{\lambda}(y + t\zeta) \in \sigma(A(y + t\zeta))$ which depends holomorphically on the parameter t in the the disk

$$\{t \in \mathbb{C} : |t| < (1 - \varepsilon)^{-1}\} \subset \{t \in \mathbb{C} : |t| < \frac{1}{2\kappa} \frac{1}{1 + \delta^{-1}}\}$$

and which is such that $\tilde{\lambda}(y + t\zeta)$ is isolated and nondegenerate eigenvalue, whenever $|t| < \frac{1}{2\kappa} \frac{1}{1 + \delta^{-1}}$, in particular for $t = 1$. Thus, $\tilde{\lambda}(y + \zeta)$ is a candidate for the holomorphic extension $\lambda(z)$ of the parametric eigenvalue. By Lemma 2.11, $\tilde{\lambda}(y + \zeta)$ is, in fact, the same eigenvalue for any choice of $y \in U$, $\zeta \in \mathbb{C}^M$ satisfying $y + \zeta = z$. Therefore, $\lambda(z) := \tilde{\lambda}(y + \zeta)$ is well-defined. Similar considerations apply to the eigenfunction $w(z)$. In the remainder of the proof, we distinguish two cases:

Case $M < \infty$. By Theorem 2.10 and Lemma 2.11, the function (λ, w) is separately complex-analytic in $E(\tau)$. The classical Hartogs' theorem [21, Theorem 2.2.8] implies joint complex-analyticity of (λ, w) on $E(\tau)$.

Case $M = \infty$. The function (λ, w) is analytic on $E(\tau)$ in the sense of Definition 2.3.1 of [20]. The claim therefore follows using [20, Prop. 3.1.2] and [20, Theorem 3.1.5]. \square

REMARK 2.14. *Notice that in (2.22) summability with $p < 1$ is required in (2.20) in order for the analyticity region $E(\tau)$ in (2.22) to increase with dimension m .*

3. Galerkin discretization of abstract variational eigenvalue problems.

To prepare the convergence analysis of the the Smolyak-Galerkin approximation of the parametric eigenvalue problems we recapitulate abstract results for the Galerkin method for elliptic eigenvalue problems from [2].

Variational approximations of the abstract eigenvalue problem (2.4) are defined in terms of a one-parameter family $\{V_h\}_{h>0} \subset V$ of finite-dimensional subspaces of increasing dimension $N(h) \rightarrow \infty$ as $h \rightarrow 0$, which is dense in V , i.e., for all $u \in V$: $\lim_{h \rightarrow 0} \inf_{\chi \in V_h} \|u - \chi\|_V = 0$. We assume that

$$\inf_{0 \neq u \in V_h} \sup_{0 \neq v \in V_h} \frac{|a(u, v)|}{\|u\|_V \|v\|_V} \geq \gamma(h) > 0. \quad (3.1)$$

We further assume *stability* of $\{V_h\}_{h>0}$ in the sense that

$$\forall u \in V : \quad \lim_{h \rightarrow 0} \gamma(h)^{-1} \inf_{\chi \in V_h} \|u - \chi\|_V = 0. \quad (3.2)$$

We define P_h as the $a(\cdot, \cdot)$ projection of V onto V_h by

$$\forall u \in V, v \in V_h : \quad a(P_h u, v) = a(u, v). \quad (3.3)$$

The projection P_h is quasi-optimal in V :

$$\|u - P_h u\|_V \leq \left(1 + \frac{C_1}{\gamma(h)}\right) \inf_{\chi \in V_h} \|u - \chi\|_V. \quad (3.4)$$

Galerkin approximations (λ_h, w_h) of eigenpairs are obtained by restricting the abstract eigenvalue problem (2.4) to V_h :

$$\text{find } \lambda_h \in \mathbb{C} \quad \text{and} \quad 0 \neq w_h \in V_h \quad \text{s.t.} \quad \forall v \in V_h : \quad a(w_h, v) = \lambda_h (w_h, v)_H. \quad (3.5)$$

Equation (3.5) is a matrix eigenvalue problem of dimension $N(h) = \dim V_h$, so that there exist $N(h)$ many, in general complex eigenvalues $\lambda_j(h)$, $j = 1, \dots, N(h)$. By (3.1), the eigenvalues are non-zero. The pair (λ_h, w_h) is an eigenpair of (3.5) if and only if (λ_h^{-1}, w_h) is an eigenpair of the compact operator $T_h : V \rightarrow V_h$ defined by

$$\forall u \in V, v \in V_h : \quad a(T_h u, v) = (u, v)_H. \quad (3.6)$$

The operator T_h can be written as $P_h T$.

Let now $\lambda \in \mathbb{C}$ be an eigenvalue of (2.4) with algebraic multiplicity m and ascent $\alpha(\lambda)$, i.e., $\lambda^{-1} \in \mathbb{C}$ is an eigenvalue of the operator T with algebraic multiplicity m and $\alpha(\lambda)$ is the ascent of $\lambda^{-1} - T$. Since T is compact, $T_h = P_h T \rightarrow T$ in norm in $\mathcal{L}(V, V)$ by (3.4) and, as $h \rightarrow 0$, there are m discrete eigenvalues $\lambda_1(h), \dots, \lambda_m(h)$ of (3.5) that converge to λ . Define

$$\begin{aligned} \mathcal{M}(\lambda) &:= \{w : w \text{ is a generalized unit eigenvector of (2.4) for } \lambda\}, \\ \mathcal{M}^*(\lambda) &:= \{w : w \text{ is a generalized adjoint unit eigenvector of (2.4) for } \lambda\} \end{aligned}$$

and their approximation bounds

$$\epsilon_h(\lambda) = \sup_{u \in \mathcal{M}(\lambda)} \inf_{\chi \in V_h} \|u - \chi\|_V \quad \text{and} \quad \epsilon_h^*(\lambda) = \sup_{u \in \mathcal{M}^*(\lambda)} \inf_{\chi \in V_h} \|u - \chi\|_V. \quad (3.7)$$

There holds the following general result on the asymptotic eigenvalue error analysis.
THEOREM 3.1. *Consider the variational eigenvalue problem (2.4) and its variational approximation (3.5). For $\lambda \in \sigma(A)$, let $\alpha(\lambda)$ denote the ascent of $\lambda^{-1} - T$. Then:*

1. There exists a constant $C > 0$ such that, as $h \rightarrow 0$,

$$|\lambda - \hat{\lambda}(h)| \leq C\gamma(h)\epsilon_h(\lambda)\epsilon_h^*(\lambda) \quad \text{where} \quad \hat{\lambda}(h) = \frac{1}{m} \sum_{j=1}^m \lambda_j(h), \quad (3.8)$$

2. and

$$|\lambda - \check{\lambda}(h)| \leq C\gamma(h)^{-1}\epsilon_h(\lambda)\epsilon_h^*(\lambda) \quad \text{where} \quad \check{\lambda}(h) = \left(\frac{1}{m} \sum_{j=1}^m \frac{1}{\lambda_j(h)} \right)^{-1}, \quad (3.9)$$

3.

$$|\lambda - \lambda_j(h)| \leq C(\gamma(h)^{-1}\epsilon_h(\lambda)\epsilon_h^*(\lambda))^{1/\alpha(\lambda)}. \quad (3.10)$$

4. Let $\lambda(h)$ be an eigenvalue of (3.5) such that $\lim_{h \rightarrow 0} \lambda(h) = \lambda \in \sigma(A)$. Suppose further that for each $h > 0$ the vector $w_h \in V_h$ is a unit vector which satisfies $(\lambda(h)^{-1} - T_h)^k w_h = 0$ for some positive integer $k \leq \alpha(\lambda)$. Then, for any $\ell \in \mathbb{N}_0$ with $k \leq \ell \leq \alpha(\lambda)$ there is a vector $u_h \in V$ such that

$$(\lambda^{-1} - T)^\ell u_h = 0 \quad \text{and} \quad \|u_h - w_h\|_V \leq C(\gamma(h)^{-1}\epsilon_h(\lambda))^{(\ell-k+1)/\alpha(\lambda)}. \quad (3.11)$$

4. Parametric elliptic EVPs. Let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain. We consider the parametric diffusion operator $A(y)$ given by

$$(A(y)\zeta)(x) = -\nabla_x \cdot (a(x, y)\nabla_x \zeta(x)), \quad x \in D, \quad y \in U \quad (4.1)$$

for a family of diffusion coefficients $y \mapsto a(\cdot, y)$, which is parametrized by a vector $y = (y_1, y_2, \dots)$ belonging to the set U of admissible parameters, defined in (2.6). The parameters y_m , $m = 1, 2, \dots$ could denote design parameters of an engineering system, or states of an optimal controller; they could also denote random coefficients in a wavelet or Karhunen-Loève expansion of the random field $y \mapsto a(\cdot, y)$. We present examples for the former scenario below.

4.1. Assumptions on the data. We suppose that in addition to being bounded and Lipschitz the domain $D \subset \mathbb{R}^d$ is simply connected. We denote by (\cdot, \cdot) the $L^2(D)$ scalar product which extends uniquely by continuity to the duality pairing on $H_0^1(D) \times H^{-1}(D)$, again denoted by (\cdot, \cdot) .

The parameter dependence of the eigenvalue problem enters through the diffusion coefficient a . We assume that $a : U \rightarrow L^\infty(D)$ is uniformly positive, i.e., that there exist constants $0 < a_{\min} \leq a_{\max} < \infty$ such that

$$\forall y \in U : \quad 0 < a_{\min} \leq \operatorname{ess\,inf}_{x \in D} a(x, y) \leq \|a(\cdot, y)\|_{L^\infty(D)} \leq a_{\max} < \infty. \quad (4.2)$$

For every $y \in U$ consider the parametric Dirichlet eigenvalue problem:

$$\text{find } \lambda(y) \in \mathbb{R} \quad \text{and} \quad 0 \neq w(y) \in H_0^1(D) \quad \text{s.t.} \quad A(y)w(y) = \lambda(y)w(y). \quad (4.3)$$

Of particular interest is the case when the diffusion coefficient depends affinely on the parameter vector $y \in U$ (for background on the following assumption see [8, 11]):

ASSUMPTION 4.1. *The parametric diffusion coefficient is of the form*

$$a(x, y) = \bar{a}(x) + \sum_{m \geq 1} y_m \psi_m(x). \quad (4.4)$$

with $y_m \in [-1, 1]$, and

$$\sum_{m \geq 1} \|\psi_m\|_{L^\infty(D)} \leq \frac{\kappa}{1 + \kappa} \bar{a}_{\min} \quad (4.5)$$

with $\bar{a}_{\min} = \text{ess inf}_D \bar{a}$, $\kappa > 0$. Further, we assume

$$\sum_{m \geq 1} \|\psi_m\|_{W^{1,\infty}(D)}^p < \infty \quad (4.6)$$

with some $p \in (0, 1)$. In particular, $U \ni y \mapsto a(\cdot, y) \in W^{1,\infty}(D)$.

REMARK 4.2. Equation (4.6) implies for (4.3) that $\Delta w(y) \in L^2(D)$ for all $y \in U$. If $D \subset \mathbb{R}^d$ is convex, this in turn implies that $w(y) \in W = H^2(D) \cap H_0^1(D)$ for all $y \in U$. For polygonal domains $D \subset \mathbb{R}^2$ with straight sides, the same statement holds for W being a weighted Sobolev space, see [1]. For simplicity we assume in the following that $D \subset \mathbb{R}^d$ is convex.

Assume $z \in \mathbb{C}^M$ is such that $\text{dist}(z, [-1, 1]) < \tau_m \in (0, \infty)$, $m \geq 1$. Then (4.3) implies

$$-a(\cdot, z) \Delta w(z) = \lambda(z) w(z) + \nabla w(z) \cdot \nabla a(\cdot, z). \quad (4.7)$$

The mappings $z \mapsto a(\cdot, z)$ and $z \mapsto \nabla a(\cdot, z)$ are jointly complex-analytic with values in $L^\infty(D)$ by (4.6). Assume that the pair $(\lambda(z), w(z))$ is jointly complex-analytic in an open neighborhood E of z with values in $\mathbb{R} \times H_0^1(D)$. Then $z \mapsto w(z)$ is jointly complex-analytic on E with values in $L^2(D)$. Assume that $\sup_{z \in E} \|a(\cdot, z)\|_{W^{1,\infty}(D)} < \infty$ and that $\inf_{(x,z) \in D \times E} a(x, z) > 0$. From (4.7) we obtain $\Delta w(z) \in L^2(D)$. Moreover, $z \mapsto \Delta w(z)$ is jointly complex-analytic on E with values in $L^2(D)$, as can be easily checked using (4.7). Under (4.6), these assumptions are satisfied in the setting of Theorem 2.13, for $z \in E(\tilde{\tau})$ where $\tilde{\tau}_m = \min\{\tau_m, c \|\psi_m\|_{W^{1,\infty}(D)}^{p-1}\}$ with $c > 0$ arbitrary and τ and $E(\cdot)$ are as in (2.22). Thus we obtain the following corollary.

COROLLARY 4.3. *In the setting of Theorem 2.13 with (4.1) and (4.4), assuming in addition (4.6) and $D \subset \mathbb{R}^d$ convex, the eigenpair (λ, w) can be extended to a jointly complex-analytic function on $E(\tilde{\tau})$ with values in $\mathbb{C} \times (H^2(D) \cap H_0^1(D))$, where $\tilde{\tau}$ is given by $\tilde{\tau}_m = \min\{\tau_m, c \|\psi_m\|_{W^{1,\infty}(D)}^{p-1}\}$ with $c > 0$ being arbitrary, and τ and $E(\cdot)$ are as in (2.22).*

The weak formulation of the parametric EVP is obtained in the usual way by testing (4.3) in $H_0^1(D)$ and integrating by parts. This results in the parametric eigenvalue problem: for every $y \in U$, find $\lambda(y) \in \mathbb{R}$ and $0 \neq w(y) \in H_0^1(D)$ such that

$$(\bar{a} \nabla u(y), \nabla \eta) + \sum_{m \geq 1} y_m (\psi_m u(y), \eta) = \lambda(y) (u(y), \eta) \quad \forall \eta \in H_0^1(D). \quad (4.8)$$

By assumption (4.2) and the Lax-Milgram lemma, for every $y \in U$ the operator $A(y) \in \mathcal{L}(H_0^1(D), H^{-1}(D))$ is boundedly invertible and by the compactness of the embedding $L^2(D) \subset H^{-1}(D)$ for every $y \in U$ the inverse $T(y) = A^{-1}(y) : L^2(D) \rightarrow H_0^1(D)$ is compact. Hence, for every $y \in U$, the elliptic operator $A(y)$ admits a countable set of eigenvectors $\{w_j(y)\}_{j \geq 1} \subset H_0^1(D)$ forming an orthonormal basis of $L^2(D)$ with

corresponding eigenvalues $0 < \lambda_1(y) \leq \lambda_2(y) \leq \dots$ and $\lambda_j \rightarrow \infty$ as $j \rightarrow \infty$ which we assume to be numbered in increasing order, counting multiplicity. Due to the self-adjointness of $A(y)$, the eigenvalues' dependence on $y \in U$ is Lipschitz [22, Sect. V.4.3, Theorem 4.10]. This Lipschitz dependence (which fails to hold for general, non self-adjoint operators) allows us to speak of parametric eigenvalue families.

REMARK 4.4. *Enumeration of the eigenvalues $\lambda_j(y)$ can be done in two ways: first, as stated above, for every $y \in U$ in increasing eigenvalue magnitude, and second, in increasing eigenvalue magnitude with respect to a reference value of y , e.g. $y = 0 \in U$. Due to possible crossings of eigenvalues, the two eigenvalue numberings may differ.*

REMARK 4.5. *For every $y \in U$ the fundamental or spectral gap $\lambda_2(y) - \lambda_1(y)$ is strictly positive and $\lambda_1(y)$ is nondegenerate by the Krein-Rutman theorem see e.g. [19, Theorem 1.2.6]. Thus there exists $\delta > 0$ such that*

$$\text{dist}(\lambda_1(y), \sigma(A(y)) \setminus \{\lambda_1(y)\}) \geq \delta \lambda_1(y) \quad \text{for all } y \in U. \quad (4.9)$$

We are in particular interested in computing for all $y \in U$ the ‘‘ground state’’, i.e., the smallest eigenvalue $\lambda_1(y)$ and a corresponding eigenfunction $w_1(y)$ of (4.8). *This shall be understood from now on.*

4.2. Multilevel finite element spaces in D . We discretize the space $V = H_0^1(D)$ by means of a dense, nested sequence of subspaces

$$\{0\} = V_{-1} \subset V_0 \subset V_1 \subset \dots \subset V = H_0^1(D),$$

of finite dimensions $n_\ell = \dim V_\ell < \infty$. Here we assume that V_ℓ , $\ell \geq 1$ result from a uniform refinement of a given regular simplicial mesh (other subspace sequences, such as spectral or p -version finite element discretizations could be considered as well). With respect to a basis $\{\phi_i^{(\ell)}\}_{i=1}^{n_\ell}$ for V_ℓ , the abstract Galerkin discretization (3.5) becomes a family of *parametric matrix eigenvalue problems*, the Galerkin projections of (4.8) onto V_ℓ : for $y \in U$, find $\mu_\ell(y) \in \mathbb{R}$ and $0 \neq u_\ell(y) \in V_\ell$ such that

$$(\bar{a} \nabla u(y), \nabla \eta) + \sum_{m \geq 1} y_m (\psi_m \nabla u(y), \nabla \eta) = \mu_\ell(y) (u(y), \eta) \quad \forall \eta \in V_\ell, \quad (4.10)$$

where, for each $y \in U$ and $\ell = 0, 1, \dots$, $(\mu_\ell(y), u_\ell(y))$ denotes the Galerkin approximation of the first eigenpair in V_ℓ .

The algebraic structure of (4.10) is particularly convenient for computation. Indeed, denoting by $\mathbf{A}_0^{(\ell)}$, $\mathbf{A}_m^{(\ell)}$, $m \geq 1$ the stiffness matrices and by $\mathbf{M}^{(\ell)}$ the mass matrix w.r.t. the basis $\{\phi_i^{(\ell)}\}_{i=1}^{n_\ell}$, given by

$$(\mathbf{A}_0^{(\ell)})_{ij} = (\bar{a} \nabla \phi_j^{(\ell)}, \nabla \phi_i^{(\ell)}), \quad (\mathbf{A}_m^{(\ell)})_{ij} = (\psi_m \nabla \phi_j^{(\ell)}, \nabla \phi_i^{(\ell)}), \quad \mathbf{M}_{ij}^{(\ell)} = (\phi_j^{(\ell)}, \phi_i^{(\ell)}),$$

the problem (4.10) reads for any given $y \in U$: find

$$\mathbf{u}_\ell(y) \in \mathbb{R}^{n_\ell} \quad \text{s.t.} \quad (\mathbf{A}_0^{(\ell)} + \sum_{m \geq 1} y_m \mathbf{A}_m^{(\ell)}) \mathbf{u}_\ell(y) = \mu_\ell(y) \mathbf{M}^{(\ell)} \mathbf{u}_\ell(y). \quad (4.11)$$

Thus, the matrices need only be assembled once if a hierarchic basis is used, and once on each level of interest $\ell \in \mathbb{N}_0$ if standard hat functions are used. Note that the method presented below uses (4.11) only with finitely supported sequences $(y_m)_{m \geq 1}$.

REMARK 4.6. For the ground state we know that the eigenvalue is separated and that the corresponding eigenvector can be chosen to be positive, see Remark 4.5, uniformly in the choice of the parameter by Lipschitz dependence of the eigenpair on the parameter. Therefore, we normalize the approximate parametric eigenvector $\mathbf{u}_\ell(y)$ at any given $y \in U$ by imposing $\mathbf{u}_\ell(y)^\top \mathbf{M}^{(\ell)} \mathbf{u}_\ell(y) = 1$ and $\int_D \sum_{i=1}^{n_\ell} \phi_i^{(\ell)}(x) (\mathbf{u}_\ell(y))_i dx \geq 0$.

We make the following assumption on the ansatz spaces V_ℓ , which is satisfied by the usual finite element ansatz spaces (see, e.g., [10]).

ASSUMPTION 4.7. There exists a $t_0 > 0$ and $C > 0$ such that for all $t \in [0, t_0]$ holds:

$$\inf_{v_h \in V_\ell} \|v_h - v\|_{H_0^1(D)} \leq C 2^{-t\ell} \|v\|_W \quad (4.12)$$

for all $\ell \in \mathbb{N}_0$ and $v \in W$. Here, $W = H_0^1(D) \cap H^{1+t}(D)$, cf. Remark 4.2. In particular, (4.12) implies (3.2).

4.3. Parametric deterministic eigenvalue problem. In this section we discuss the numerical solution of the parametric matrix eigenvalue problems (4.11) for fixed $y \in U$. By Assumption (4.2), for each $y \in U$, $\ell \in \mathbb{N}_0$ the matrix $\mathbf{A}_0^{(\ell)} + \sum_{m \geq 1} y_m \mathbf{A}_m^{(\ell)}$ is symmetric positive definite. Setting $y_m = 0$, $m > M$, we conclude that the truncated matrix $\mathbf{A}_0^{(\ell)} + \sum_{m=1}^M y_m \mathbf{A}_m^{(\ell)}$ has the same property uniformly w.r.t. $M \in \mathbb{N}$. Hence, for any $y \in U$, $\ell \in \mathbb{N}_0$, $M \in \mathbb{N}$, the truncated problem

$$\mathbf{u}_\ell^{(M)}(y) \in \mathbb{R}^{n_\ell} : \quad (\mathbf{A}_0^{(\ell)} + \sum_{m=1}^M y_m \mathbf{A}_m^{(\ell)}) \mathbf{u}_\ell^{(M)}(y) = \mu_\ell^{(M)}(y) \mathbf{M}^{(\ell)} \mathbf{u}_\ell^{(M)}(y) \quad (4.13)$$

is a symmetric positive definite generalized matrix eigenvalue problem. By the same reasoning, the first eigenvalue of the truncated system (4.13) is simple, i.e., of single multiplicity, uniformly in $y \in U$ for all $\ell \geq 0$ sufficiently large, see Corollary 4.9. We identify $\mathbf{u}_\ell(y_1, y_2, \dots, y_M, 0, \dots) = \mathbf{u}_\ell^{(M)}(y_1, \dots, y_M)$ and denote by $u_\ell(y) \in V_\ell$ the function

$$u_\ell(y) = \sum_{i=1}^{n_\ell} \phi_i^{(\ell)}(\mathbf{u}_\ell(y))_i.$$

Any numerical method for symmetric generalized eigenvalue problems applies to (4.13), see, e.g. [28, 17, 13, 18]. For our computations we use the JDBSYM library [16], which implements a variant of the Jacobi-Davidson method.

4.4. Abstract error bounds. Based on the abstract eigenvalue approximation theory in Section 3, we obtain the following a-priori error bounds, cf. [2, Sect. 8].

PROPOSITION 4.8. Consider the eigenvalue problem (4.3) (or in variational form (2.11)) with the operator $A(y)$ being the parametric diffusion operator in (4.1). Let assumptions (4.2), (4.4) and (4.5) hold. For some $m \geq 1$ let $(\lambda_m(y), w_m(y))$ be an eigenpair of the EVP (4.3) with eigenspace of multiplicity one for all $y \in U$, i.e., eigenvalue crossings are excluded.

Then there exist constants $C > 0$, $h_0 > 0$ such that the Galerkin eigenvalue approximation from the finite element space V_ℓ is quasi-optimal uniformly in $y \in U$: for every $y \in U$ and every $0 < h < h_0$ holds

$$\|w_m(y) - w_m^h(y)\|_{H_0^1(D)} \leq C \epsilon_h(\lambda_m(y)) = C \inf_{v_h \in V_\ell} \|w_m(y) - v_h\|_{H_0^1(D)}. \quad (4.14)$$

For the Galerkin eigenvalue approximations holds

$$|\lambda_m(y) - \lambda_m^h(y)| \leq C(\epsilon_h(\lambda_m(y)))^2 \leq C \inf_{v_h \in V_h} \|w_m(y) - v_h\|_{H_0^1(D)}^2. \quad (4.15)$$

Using Remark 4.5, Assumption 4.7 and $\sup_{y \in U} \|w_m(y)\|_W < \infty$ we obtain the following corollary.

COROLLARY 4.9. *For $h > 0$ small enough, the spectral gap of the discretized problem, $\lambda_2^h(y) - \lambda_1^h(y)$ is strictly positive uniformly in $y \in U$. Moreover, for $0 \leq t \leq t_0$ as in Assumption 4.1, we have*

$$\sup_{y \in U} \|w_1(y) - w_1^h(y)\|_{H_0^1(D)} \leq Ch^t \quad \text{and} \quad \sup_{y \in U} |\lambda_1(y) - \lambda_1^h(y)| \leq Ch^{2t} \quad (4.16)$$

where $C < \infty$ is independent of h for sufficiently small $h > 0$.

5. Sparse composite collocation method. In this section we introduce a generalization of the sparse grid collocation operator introduced in [23, 6]. We start by collecting the necessary definitions.

Let L_n denote the univariate Lagrange polynomial of degree $n \in \mathbb{N}_0$ scaled such that $\int_{-1}^1 |L_n(t)|^2 \frac{dt}{2} = 1$. For $n \in \mathbb{N}_0$ let i_n denote the Lagrange interpolation operator of degree n which maps a continuous function $v \in C^0([-1, 1])$ to a polynomial $i_n v$ of degree n with $(i_n v)(z) = v(z)$ in the Gauss-Legendre points $z \in \{-1 < z_0^n < \dots < z_n^n < 1 : L_{n+1}(z_k^n) = 0\}$. Denote by $\{w_k^n\}_{k=0}^n \subset (0, \infty)$ the corresponding Gauss-Legendre quadrature weights [15]. Define $j_n = i_n - i_{n-1}$, $n \geq 0$ where, by convention, $i_{-1} := 0$. Let $\mathcal{F} \subset \mathbb{N}_0^{\mathbb{N}}$ denote the collection of all multi-indices $\nu \in \mathbb{N}_0^{\mathbb{N}}$ such that for each $\nu \in \mathcal{F}$ the *support* $\text{supp } \nu = \{m \in \mathbb{N} : \nu_m \neq 0\}$ is finite. For a multiindex $\nu \in \mathcal{F}$ we denote by L_ν the tensorized Legendre polynomial $L_\nu = L_{\nu_1} \otimes L_{\nu_2} \otimes \dots$, that is $L_\nu(y) = \prod_{m \geq 1} L_{\nu_m}(y_m)$ for all $y \in U$. For a multiindex set $\Lambda \subset \mathcal{F}$ we denote by \mathbb{P}_Λ the span of $\{L_\nu : \nu \in \Lambda\}$. Note that $\mathbb{P}_\mathcal{F}$ is dense in $L_\pi^2(U)$, where π is the uniform probability measure. Below we frequently use the notation $1_A = 1$ if A is true and $1_A = 0$ otherwise, as well as $\tilde{\nu} \leq \nu$ (or $\nu \geq \tilde{\nu}$) iff $\tilde{\nu}_m \leq \nu_m$ for all $m \geq 1$. For any finite multiindex set $\Lambda \subset \mathcal{F}$ we define the *sparse collocation operator*

$$I_\Lambda = \sum_{\nu \in \Lambda} \bigotimes_{m \geq 1} j_{\nu_m} \quad (5.1)$$

with the convention $I_\emptyset := 0$. Due to Λ being finite, for each of the finitely many $\nu \in \Lambda$ the tensorized operator $\bigotimes_{m \geq 1} j_{\nu_m}$ has only finitely many non-trivial increment factors j_{ν_m} , $m \in \text{supp } \nu$, while $j_{\nu_m} = j_0 = i_0$ for $m \notin \text{supp } \nu$. This implies that I_Λ is well defined on bounded continuous functions on U . The operator I_Λ has appeared in various variants, i.e., for specific choices of Λ in the literature, see [6, 3, 7, 4, 24, 23] and references therein. In particular, the generic definition (5.1) accomodates all formulas of [5].

DEFINITION 5.1. *A multiindex set $\Lambda \subseteq \mathcal{F}$ is called monotone if the following implication holds:*

$$\forall \nu \in \Lambda : \quad \tilde{\nu} \in \mathcal{F} \wedge \tilde{\nu} \leq \nu \Rightarrow \tilde{\nu} \in \Lambda.$$

In particular, $\{\}, \{0\}$ and \mathcal{F} are monotone. Monotone multiindex sets Λ for I_Λ defined by (5.1) are of particular interest, as the following lemma shows.

LEMMA 5.2. *Let $\Lambda \subset \mathcal{F}$ be finite and monotone. Then $I_\Lambda p = p$ for all $p \in \mathbb{P}_\Lambda$. Moreover, I_Λ is unisolvent on \mathbb{P}_Λ , i.e., $\forall p \in \mathbb{P}_\Lambda: I_\Lambda p = 0$ implies $p = 0$. Further, defining \mathbb{Q}_Λ as the span of the monomials $\{U \ni y \mapsto y^\nu : \nu \in \Lambda\}$, we have $\dim \mathbb{P}_\Lambda = \#\Lambda = \dim \mathbb{Q}_\Lambda < \infty$ and moreover $\mathbb{Q}_\Lambda = \mathbb{P}_\Lambda$.*

Proof. Observe that $I_{\{\nu'\}} L_\nu = \bigotimes_{m \geq 1} j_{\nu'_m} L_{\nu_m} = 0$ if $\nu'_m > \nu_m$ for at least one $m \in \mathbb{N}$. This allows to write for $\nu \in \Lambda$

$$I_\Lambda L_\nu = \sum_{\nu' \leq \nu} \bigotimes_{m \geq 1} j_{\nu'_m} L_{\nu_m} = \bigotimes_{m \geq 1} \sum_{\nu'_m=0}^{\nu_m} j_{\nu'_m} L_{\nu_m} = \bigotimes_{m \geq 1} L_{\nu_m} = L_\nu, \quad (5.2)$$

which shows the first assertion. Let now $p \in \mathbb{P}_\Lambda = \text{span}\{L_\nu : \nu \in \Lambda\}$ s.t. $I_\Lambda p = 0$. By (5.2), linearity of I_Λ and linear independence of the set $\{L_\nu\}_{\nu \in \Lambda}$ imply $p = 0$. The statement $\dim \mathbb{P}_\Lambda = \#\Lambda = \dim \mathbb{Q}_\Lambda < \infty$ is obvious. Since Λ is monotone, we have $L_\nu \in \mathbb{Q}_\Lambda \forall \nu \in \Lambda$, hence $\mathbb{P}_\Lambda \subseteq \mathbb{Q}_\Lambda$. Now $\dim \mathbb{P}_\Lambda = \dim \mathbb{Q}_\Lambda < \infty$ implies $\mathbb{P}_\Lambda = \mathbb{Q}_\Lambda$. \square

Note that in general, however, I_Λ is not interpolatory, as the cardinality of the set of collocation nodes in (5.1) may be larger than $\dim \mathbb{P}_\Lambda$, as the following example shows.

EXAMPLE 5.3. *Consider $\Lambda = \{\nu_0 = 0, \nu_1 = (1, 0, 0, \dots), \nu_2 = (0, 1, 0, 0, \dots)\}$ consisting of three multiindices. The corresponding sparse collocation operator I_Λ is then based on the following five collocation nodes: $z_0^{\nu_0} = 0$, $z_-^{\nu_1} = (-\frac{1}{\sqrt{3}}, 0, \dots)$, $z_+^{\nu_1} = (+\frac{1}{\sqrt{3}}, 0, \dots)$, $z_-^{\nu_2} = (0, -\frac{1}{\sqrt{3}}, 0, \dots)$ and $z_+^{\nu_2} = (0, +\frac{1}{\sqrt{3}}, 0, \dots)$.*

The following observation will be useful: for any monotone $\Lambda \subset \mathcal{F}$ we have

$$\forall \nu \in \Lambda : \prod_{m \geq 1} (\nu_m + 1) = \#\{\tilde{\nu} \in \Lambda : \tilde{\nu} \leq \nu\}. \quad (5.3)$$

LEMMA 5.4 (Number of collocation points). *Let $\Lambda \subset \mathcal{F}$ be monotone and finite. Then the number collocation points in I_Λ is at most $(\#\Lambda)^2$.*

Proof. We can write I_Λ as $I_\Lambda = \sum_{\nu \in \Lambda} c_\nu^\Lambda \bigotimes_{m \geq 1} i_{\nu_m}$, where each term requires $\prod_{m \geq 1} (\nu_m + 1)$ collocation points by definition of i_n , $n \geq 0$. The number of collocation points in I_Λ can therefore be estimated as

$$\sum_{\nu \in \Lambda} \prod_{m \geq 1} (\nu_m + 1) \stackrel{(5.3)}{\leq} \sum_{\nu \in \Lambda} \#\{\tilde{\nu} \in \Lambda : \tilde{\nu} \leq \nu\} \leq \sum_{\nu \in \Lambda} \#\Lambda = (\#\Lambda)^2.$$

\square

REMARK 5.5. *Note that the upper bound $(\#\Lambda)^2$ is independent of the “effective dimension” $\max_{\nu \in \Lambda} \#\text{supp } \nu$. Moreover, the exponent 2 in the upper bound $(\#\Lambda)^2$ on the number of collocation points in I_Λ is sharp, as can be seen from the sequence*

$$\Lambda_\ell = \{(\nu_1, \dots, \nu_M, 0 \dots) \in \mathcal{F} : \|\nu\|_{\ell^1(\mathbb{N})} \leq \ell\}, \quad \ell = 0, 1, \dots$$

where $M \geq 1$ is fixed. Indeed, by [6, Sect. 6.2], we have $\#\Lambda_\ell = \binom{M+\ell}{M}$ while the number of collocation points in I_Λ is given by $N_{\ell, M} = \binom{2M+\ell}{2M}$. This yields

$$N_{\ell, M} \geq \left(\frac{2M+\ell}{2M}\right)^{2M} \gtrsim (2M+\ell)^{2M} \geq (M+\ell)^{2M} \geq \left(\frac{(M+\ell)^M}{M!}\right)^2 \geq (\#\Lambda_\ell)^2$$

with the implied constant independent of $\ell \geq 0$. Thus, $N_{\ell, M} \sim (\#\Lambda_\ell)^2$ as $\ell \rightarrow \infty$.

REMARK 5.6. For $\ell \geq 0$ let $n_\ell^* := 1_{(\ell \geq 1)} 2^\ell$. For a monotone and finite $\Lambda \subset \mathcal{F}$ define $\Lambda^* := \bigcup_{\nu \in \Lambda} \{\tilde{\nu} \in \mathcal{F} : \forall m \geq 1 : \tilde{\nu}_m \leq \min\{n_\ell^* \geq \nu_m : \ell \geq 0\}\}$. For each $\ell \geq 0$, $n_\ell^* + 1$ is the number of nodes in the Clenshaw-Curtis quadrature rule of order $n_\ell^* + 1$. The nodes of these Clenshaw-Curtis quadrature rules are nested, cf. [23]. Thus, if we defined i_n to be the Lagrange interpolation operator based on the nodes of the Clenshaw-Curtis quadrature rule of order $\min\{n_\ell^* \geq n : \ell \geq 0\} + 1$, for monotone and finite Λ the number of collocation points in I_{Λ^*} would be exactly $\#\Lambda^* = \dim \mathbb{P}_{\Lambda^*}$, as opposed to the possibly quadratic growth in Remark 5.5.

The following is a characterization of I_Λ for monotone multiindex sets $\Lambda \subset \mathcal{F}$ by an expansion into tensorized multivariate Legendre polynomials.

LEMMA 5.7. Let $\Lambda \subset \mathcal{F}$ be monotone and finite, and assume $c_\nu^\Lambda \in \mathbb{R}$, $\nu \in \Lambda$ are such that $I_\Lambda = \sum_{\nu \in \Lambda} c_\nu^\Lambda \bigotimes_{m \geq 1} i_{\nu_m}$. Then for $f \in C^0(\bar{U}; \mathbb{R})$ we have $I_\Lambda f = \sum_{\nu' \in \Lambda} d_{\nu'}^\Lambda(f) L_{\nu'}$, where $d_{\nu'}^\Lambda(f)$, $\nu' \in \Lambda$ is defined by

$$d_{\nu'}^\Lambda(f) = \sum_{\substack{\nu \in \Lambda \\ \nu \geq \nu'}} c_\nu^\Lambda \sum_{\eta \leq \nu} w_\eta^\nu L_{\nu'}(z_\eta^\nu) f(z_\eta^\nu),$$

where $w_\eta^\nu = \prod_{m \geq 1} w_{\eta_m}^{\nu_m} \in \mathbb{R}$ and $z_\eta^\nu = (z_{\eta_1}^{\nu_1}, z_{\eta_2}^{\nu_2}, \dots) \in U$, $\eta \leq \nu \in \Lambda$.

Proof. We have the univariate formula $i_n v = \sum_{n' \leq n} d_{n'}(v) L_{n'}$ with coefficients $d_{n'}(v) = \sum_{k \leq n'} w_k^n L_{n'}(z_k^n) v(z_k^n)$ as can be immediately checked using the fact that the Gauss-Legendre quadrature formula $p \mapsto \sum_{k \leq n} w_k^n p(z_k^n)$ integrates polynomials $p = L_{n'} L_{n''}$ of degree at most $2n + 1$ exactly [15, (1.4.7) and (1.4.14)]. Thus we obtain

$$I_\Lambda f = \sum_{\nu \in \Lambda} c_\nu^\Lambda \left(\bigotimes_{m \geq 1} i_{\nu_m} \right) f = \sum_{\nu \in \Lambda} c_\nu^\Lambda \sum_{\nu' \leq \nu} L_{\nu'} \sum_{\eta \leq \nu} w_\eta^\nu L_{\nu'}(z_\eta^\nu) f(z_\eta^\nu),$$

which, after exchanging the summation, yields the claim. \square

In what follows, we specialize our considerations to multiindex sets Λ of a particular structure, motivated by a-priori estimates obtained in [12] for coefficients $a(x, y)$ satisfying Assumption 4.1. To describe the structure of these sets, let c_0 denote the collection of non-increasing sequences of reals less than one and tending to zero:

$$c_0 = \{\mu = (\mu_1, \mu_2, \dots) \in [0, 1]^\mathbb{N} : 1 > \mu_1 \geq \mu_2 \geq \dots \text{ and } \lim_{m \rightarrow \infty} \mu_m = 0\}.$$

For $\mu \in c_0$, $\nu \in \mathcal{F}$ we write $\mu^\nu = \prod_{m \in \mathbb{N}} \mu_m^{\nu_m}$ (with $0^0 := 1$) and for $\varepsilon > 0$ we define

$$\Lambda(\mu, \varepsilon) = \{\nu \in \mathcal{F} : \mu^\nu \geq \varepsilon\}. \quad (5.4)$$

Clearly, for any $\mu \in c_0$ and any $\varepsilon > 0$ the multiindex set $\Lambda(\mu, \varepsilon)$ is finite and monotone. The multiindex sets $\Lambda(\mu, \varepsilon)$ were introduced in [9] and investigated in the context of the stochastic Galerkin method for elliptic stochastic PDEs in [8]. The next result give precise asymptotics for the cardinality of the sets $\Lambda(\mu, \varepsilon)$ defined in (5.4) for a sequence $\mu \in c_0$ which models algebraic decay of terms in the expansions (2.12), (4.4).

LEMMA 5.8. For $\mu \in c_0$ given by $\mu_m = (1 + m)^{-\sigma}$ with $\sigma > 1$,

1. as $\varepsilon \rightarrow 0$, the cardinality of the set $\Lambda(\mu, \varepsilon)$ in (5.4) equals

$$\#\Lambda(\mu, \varepsilon) = F(\varepsilon^{-1/\sigma}) \quad \text{where} \quad F(x) = x \frac{e^{2\sqrt{\log x}}}{2\sqrt{\pi}(\log x)^{3/4}} (1 + \mathcal{O}(1/\log x))$$

2. for $p > 1/\sigma$ we have $\sup_{\varepsilon > 0} \sum_{\nu \in \Lambda(\mu, \varepsilon)} (\mu^\nu)^p < \infty$ while for $0 < p \leq 1/\sigma$ we have $\sum_{m \geq 1} \mu_m^p = \infty$.

Here and throughout, the function \log always denotes the natural logarithm.

Proof. See [8, Prop. 4.5] for the proof of the first claim. The second is a special case of [11, Lemma 7.1]. We provide the argument for completeness here. Let $0 < \delta < p\sigma - 1$, $c(\delta) > 0$ be such that $F(x) \leq c(\delta)x^{1+\delta}$ for $x > 1$. Take $\gamma = \sigma/(p\sigma - 1 - \delta)$ and let I_n be the interval $I_n = [2^{-\gamma n}, 2^{-\gamma(n-1)})$. Note that

$$\#\{\nu \in \mathcal{F} : \mu^\nu \in I_n\} \leq \#\Lambda(\mu, 2^{-\gamma n}) = F(2^{\gamma n/\sigma}) \leq c(\delta)2^{\gamma n/\sigma(1+\delta)}.$$

We compute

$$\begin{aligned} \sup_{\varepsilon > 0} \sum_{\nu \in \Lambda(\mu, \varepsilon)} (\mu^\nu)^p &= \lim_{\varepsilon \rightarrow 0} \sum_{\nu \in \Lambda(\mu, \varepsilon)} (\mu^\nu)^p = \sum_{n \geq 1} \sum_{\substack{\nu \in \mathcal{F} \\ \mu^\nu \in I_n}} (\mu^\nu)^p \\ &\leq \sum_{n \geq 1} c(\delta)2^{\gamma n/\sigma(1+\delta)} 2^{-p\gamma(n-1)} \\ &= c(\delta)2^{p\gamma} \sum_{n \geq 1} 2^{-\gamma n(p-(1+\delta)/\sigma)} = c(\delta)2^{p\gamma}, \end{aligned}$$

which is finite. For $0 < p \leq 1/\sigma$, $\sum_{m \geq 1} \mu_m^p \geq \sum_{m \geq 1} (1+m)^{-1}$ diverges. \square

In Example 5.3 we showed that, in general, $\#\Lambda > \dim \mathbb{P}_\Lambda$, i.e., the number of deterministic problems to be solved for the determination of I_Λ is larger than the number of monomials in \mathbb{P}_Λ which determine the precision of I_Λ , cf. (5.6). To facilitate comparison with Monte-Carlo methods, we quantify the convergence of I_Λ in terms of the number of deterministic problems to be solved. To this end, we bound $\dim \mathbb{P}_\Lambda = \#\Lambda$ for several classes of monotone index sets Λ .

DEFINITION 5.9. For $\mu \in c_0$ and $0 < \varepsilon \leq 1$, define

$$\mathcal{B}(\mu, \varepsilon) := \max\{m \geq 1 : \mu_m \geq \varepsilon\} \sum_{\nu \in \Lambda(\mu, \varepsilon)} 4^{\#\text{supp } \nu} \prod_{m \in \text{supp } \nu} \frac{1 + \mu_m}{1 - \mu_m} \quad (5.5)$$

where $\Lambda(\mu, \varepsilon)$ is as in (5.4) and

$$\varkappa^*(\mu) := \inf\{\varkappa > 0 : \sup_{0 < \varepsilon \leq 1} \varepsilon^\varkappa \mathcal{B}(\mu, \varepsilon) < \infty\},$$

which may be infinite. We refer to $\varkappa^*(\mu)$ as asymptotic overhead order of $\mu \in c_0$.

The class of sequences $\mu \in c_0$ which have finite asymptotic overhead order $\varkappa^*(\mu) < \infty$ includes some important families, as we show in the following.

LEMMA 5.10 (Asymptotic overhead order for algebraic decay). *For the model sequence $\mu_m = (1+m)^{-\sigma}$ with algebraic decay with fixed order $\sigma > 1$ the asymptotic overhead order $\varkappa^*(\mu)$ is bounded by $\varkappa^*(\mu) \leq 2(1 + \log 4)/\sigma$.*

Proof. Let $0 < \varepsilon \leq \mu_1$. Clearly $\max\{m \geq 1 : \mu_m \geq \varepsilon\} \leq \varepsilon^{-1/\sigma}$. Using [8, Lemma 4.8] we therefore have $\#\text{supp } \nu \leq 2 \log(\varepsilon^{-1/\sigma})$ if μ^ν is sufficiently small for $\nu \in \Lambda(\mu, \varepsilon)$ to hold. Thus, for $\nu \in \Lambda(\mu, \varepsilon)$ and any fixed $\delta > 0$ we obtain

$$4^{\#\text{supp } \nu} \prod_{m \in \text{supp } \nu} \frac{1 + \mu_m}{1 - \mu_m} \lesssim (4 + \delta)^{-2/\sigma \log \varepsilon} = \varepsilon^{-2 \log(4+\delta)/\sigma}$$

μ^ν is small enough and trivially the same bound otherwise. Using $\#\Lambda(\mu, \varepsilon) \lesssim \varepsilon^{-(1+\delta)/\sigma}$ (Lemma 5.8) we obtain for $\varepsilon \rightarrow 0$

$$\#\Lambda(\mu, \varepsilon) \max\{m \geq 1 : \mu_m \geq \varepsilon\} \max_{\nu \in \Lambda(\mu, \varepsilon)} 4^{\#\text{supp } \nu} \prod_{m \in \text{supp } \nu} \frac{1 + \mu_m}{1 - \mu_m} \lesssim \varepsilon^{-\varkappa_\delta}$$

with $\varkappa_\delta = (1 + \delta)/\sigma + 1/\sigma + 2 \log(4 + \delta)/\sigma$. Thus $\varkappa^*(\mu) \leq \varkappa_\delta$. Since $\delta > 0$ was arbitrary, the claim follows. \square

LEMMA 5.11 (Isotropic sparse tensor product asymptotic overhead order). *Assume $\mu \in c_0$ is of the form $\mu_1 = \mu_2 = \dots = \mu_M = \mu_0 > 0$, $\mu_m = 0$ for $m > M$, where $M \in \mathbb{N}$ is fixed. Then the asymptotic overhead order of μ is $\varkappa^*(\mu) = 0$.*

Proof. Take any $\varkappa > 0$. With $L = \log \varepsilon / \log \mu_0$ (for simplicity an integer) we have [29]

$$\#\Lambda(\mu, \varepsilon) = \binom{M+L}{M} \leq \frac{(M+L)^M}{M!} \lesssim (-\log \varepsilon)^M \lesssim \varepsilon^{-\varkappa} \quad \text{for } \varepsilon \rightarrow 0$$

and the other terms in (5.5) are bounded, independently of $\varepsilon > 0$. \square

Lemma 5.10 and Lemma 5.11 give rise to an even larger family of sequences μ for which the asymptotic overhead order of μ can be estimated.

PROPOSITION 5.12. *Let $\mu^{(1)} \in c_0$ for which the asymptotic overhead order of $\mu^{(1)}$ is finite, i.e., $\varkappa := \varkappa^*(\mu^{(1)}) < \infty$. Let $M \geq 1$ and $\mu \in c_0$ be a sequence with $\mu_m^{(1)} \geq \mu_m$ for $m > M$. Then we have $\varkappa^*(\mu) \leq \varkappa^*(\mu^{(1)})$.*

Proof. We assume w.l.o.g. that $0 < \varepsilon \leq \min\{\mu_1, \mu_1^{(1)}\}$. Define $\mu^{(0)} \in c_0$ by $\mu_m^{(0)} = 1_{(m \leq M)} \mu_1$, $m \geq 1$. Let H_ε be the map $H_\varepsilon : \Lambda(\mu^{(0)}, \varepsilon) \times \Lambda(\mu^{(1)}, \varepsilon) \rightarrow \Lambda(\mu, \varepsilon)$ given by

$$h = (\nu^{(0)}, \nu^{(1)}) \mapsto H_\varepsilon(h) = 1_{(\mu^\nu \geq \varepsilon)} \nu \quad \text{with} \quad \nu_m = \begin{cases} \nu_m^{(0)} & m \leq M \\ \nu_m^{(1)} & m > M. \end{cases}$$

Observe that H_ε is well-defined and surjective, and thus

$$\#\Lambda(\mu, \varepsilon) \leq \#\Lambda(\mu^{(0)}, \varepsilon) \#\Lambda(\mu^{(1)}, \varepsilon)$$

due to $\varepsilon < 1$. Noticing further that

$$\max\{m \geq 1 : \mu_m \geq \varepsilon\} \leq \max\{m \geq 1 : \mu_m^{(0)} \geq \varepsilon\} \max\{m \geq 1 : \mu_m^{(1)} \geq \varepsilon\}$$

and for all $h = (\nu^{(0)}, \nu^{(1)}) \in H_\varepsilon^{-1}(\Lambda(\mu, \varepsilon))$ we have

$$\text{supp } H_\varepsilon(h) \subseteq \text{supp } \nu^{(0)} \cup (\text{supp } \nu^{(1)} \cap \{m \in \mathbb{N} : m > M\})$$

and thus also

$$\begin{aligned} \prod_{m \in \text{supp } H_\varepsilon(h)} \frac{1 + \mu_m}{1 - \mu_m} &\leq \left(\prod_{m \in \text{supp } \nu^{(0)}} \frac{1 + \mu_m}{1 - \mu_m} \right) \left(\prod_{\substack{m \in \text{supp } \nu^{(1)} \\ m > M}} \frac{1 + \mu_m}{1 - \mu_m} \right) \\ &\leq \left(\prod_{m \in \text{supp } \nu^{(0)}} \frac{1 + \mu_m^{(0)}}{1 - \mu_m^{(0)}} \right) \left(\prod_{m \in \text{supp } \nu^{(1)}} \frac{1 + \mu_m^{(1)}}{1 - \mu_m^{(1)}} \right), \end{aligned}$$

we conclude from Lemma 5.11 that $\mathcal{B}(\mu) \leq \mathcal{B}(\mu^{(0)})\mathcal{B}(\mu^{(1)}) \lesssim \varepsilon^{-\delta} \varepsilon^{-\varkappa}$ as $\varepsilon \rightarrow 0$ for any $\delta > 0$. This shows the claim. \square

For complex-analytic functions on product domains we obtain the following approximation property of I_Λ for multiindex sets Λ of type (5.4).

THEOREM 5.13. *Let $\{\rho_m\}_{m \geq 1}$ be a sequence with $\rho_m > 1$ and $\rho_m \rightarrow \infty$. For each $m \geq 1$ let \mathcal{E}_{ρ_m} denote the ellipse in \mathbb{C} with sum of semiaxes ρ_m (as in [14, p. 312]). Let $v : \mathcal{E} := \prod_{m \geq 1} \mathcal{E}_{\rho_m} \rightarrow \mathbb{C}$ be a jointly complex-analytic function. Define $\mu \in c_0$ by*

$\mu_m = \sup_{m' \geq m} \frac{1}{\rho_{m'}}$, $m \in \mathbb{N}$. Assume $\sup_{m \geq 1} \mu_m m^\sigma < \infty$ for some $\sigma > 0$ and that $\varkappa^*(\mu) < \infty$. Take $\varkappa > \varkappa^*(\mu)$. Then there exists $C > 0$ such that

$$\|I_{\Lambda(\mu, \varepsilon)} v - v\|_{L^2_\pi(U; \mathbb{C})} \leq C \left(\varepsilon^{1-\varkappa} + \varepsilon^{1-1/\sigma} \right) \|v\|_{C^0(\overline{\mathcal{E}}; \mathbb{C})} \quad (5.6)$$

for all $0 < \varepsilon \leq \mu_1$, where π is the uniform probability measure on U .

Proof. As in [6, 24] we use the following approximation property of the univariate interpolation operator i_n for complex-analytic functions f on \mathcal{E}_{ρ_m} :

$$\|f - i_n f\|_{L^2((-1,1); \frac{d\mu}{2})} \leq C(\rho_m) \mu_m^n \|f\|_{C^0(\overline{\mathcal{E}_{\rho_m}})}, \quad n = 0, 1, \dots$$

where $C(\rho_m) = \frac{4}{\rho_m - 1} \lesssim \mu_m \rightarrow 0$ as $m \rightarrow \infty$. In particular, for $n = 1, 2, \dots$

$$\|j_n f\|_{L^2((-1,1); \frac{d\mu}{2})} \leq C(\rho_m) \mu_m^n \left(1 + \frac{1}{\mu_m} \right) \|f\|_{C^0(\overline{\mathcal{E}_{\rho_m}})} \quad (5.7)$$

and there exists $C < \infty$, independent of v , with $\prod_{m \in \text{supp } \nu} C(\rho_m) \leq C$ for any $\nu \in \mathcal{F}$. Clearly,

$$\|j_0 f\|_{L^2((-1,1); \frac{d\mu}{2})} \leq \|f\|_{C^0([-1,1])}. \quad (5.8)$$

Another observation used in the following is this: for $0 < \varepsilon^* \leq \varepsilon \leq 1$ setting $n_0 = \max\{n \geq 0 : \mu_m^{n_0} \geq \varepsilon^*\}$ yields $\mu_m^{n_0+1} < \varepsilon^*$, and thus for $m \geq 1$

$$\|f - \sum_{n \geq 0} 1_{(\mu_m^n \geq \varepsilon^*)} j_n f\|_{L^2((-1,1); \frac{d\mu}{2})} \leq C(\rho_m) \varepsilon^* \mu_m^{-1} \|f\|_{C^0(\overline{\mathcal{E}_{\rho_m}})}. \quad (5.9)$$

Similarly to [29, 6] we now write using $\mu^* \in c_0$, $\mu_m^* = \mu_{m+1}$, $m \geq 1$ for any $0 < \varepsilon \leq 1$

$$I_{\Lambda(\mu, \varepsilon)} = \sum_{\nu^* \in \Lambda(\mu^*, \varepsilon)} \left(\sum_{n \geq 0} 1_{(\mu_1^n \geq \varepsilon / (\mu^*)^{\nu^*})} j_n \right) \otimes j_{\mu_1^*} \otimes \dots$$

and use the decomposition

$$I_{\Lambda(\mu, \varepsilon)} - \text{Id} = \text{Id}^1 \otimes (I_{\Lambda(\mu^*, \varepsilon)} - \text{Id}^{1^*}) - J_{\Lambda(\mu, \varepsilon)}^* \quad (5.10)$$

where Id^m is the identity on m -variate functions, $\text{Id} = \text{Id}^m \otimes \text{Id}^{m^*}$ and

$$J_{\Lambda(\mu, \varepsilon)}^* = \sum_{\nu^* \in \Lambda(\mu^*, \varepsilon)} \left(\text{Id}^1 - \sum_{n \geq 0} 1_{(\mu_1^n \geq \varepsilon / (\mu^*)^{\nu^*})} j_n \right) \otimes j_{\mu_1^*} \otimes \dots \quad (5.11)$$

In the remaining part of the proof we assume without loss of generality $\|v\|_{C^0(\overline{\mathcal{E}}; \mathbb{C})} \leq 1$. In particular, fixing $z_j \in \mathcal{E}_{\rho_j}$, $j \neq m$ we have $\|z_m \mapsto v(z_1, z_2, \dots)\|_{C^0(\overline{\mathcal{E}_{\rho_m}}; \mathbb{C})} \leq 1$, and by stability (5.8) we discard the effect of the countably many trivial increment operators j_0 in (5.11) in the following estimate. For v as in the statement of the claim and any $\varepsilon > 0$ we now estimate using (5.9) and (5.7)

$$\begin{aligned} \|J_{\Lambda(\mu, \varepsilon)}^* v\|_{L^2_\pi(U; \mathbb{C})} &\leq \sum_{\nu^* \in \Lambda(\mu^*, \varepsilon)} C(\rho_1) \frac{\varepsilon}{(\mu^*)^{\nu^*}} \mu_1^{-1} \prod_{m \in \text{supp } \nu^*} C(\rho_{m+1}) (\mu_m^*)^{\nu_m^*} \left(1 + \frac{1}{\mu_m^*} \right) \\ &\leq \varepsilon C(\rho_1) \mu_1^{-1} \sum_{\nu^* \in \Lambda(\mu^*, \varepsilon)} \prod_{m \in \text{supp } \nu^*} C(\rho_{m+1}) \left(1 + \frac{1}{\mu_{m+1}} \right) \\ &\leq \varepsilon \sum_{\nu \in \Lambda(\mu, \varepsilon)} \prod_{m \in \text{supp } \nu} 4 \frac{1 + \mu_m}{1 - \mu_m} \end{aligned}$$

Repeating this argument using the decomposition (5.10) we obtain (5.6). To this end observe also that by the elementary inequality [12, (3.13)] for $0 < p \leq q \leq \infty$

$$\left(\sum_{m>M} |\mu_m|^q \right)^{1/q} \leq M^{\frac{1}{q} - \frac{1}{p}} \left(\sum_{m \geq 1} |\mu_m|^p \right)^{1/p}, \quad (|\mu_m|)_{m \geq 1} \in \ell^p(\mathbb{N}) \quad \text{nonincreasing}$$

for all $M \geq 1$ we have, setting $q = 1$ and $p = 1/\sigma$,

$$\|\text{Id}^M \otimes (I_{\{0\}} - \text{Id}^{M^*})v\|_{L^2_\pi(U; \mathbb{C})} \leq \sum_{m>M} C(\rho_m) \leq C \sum_{m>M} \mu_m \leq C \|\mu\|_{\ell^{1/\sigma}(\mathbb{N})} M^{1-\sigma}$$

for all $M \geq 1$, in particular for $M = \max\{m \geq 0 : \mu_m \geq \varepsilon\} \lesssim \varepsilon^{-1/\sigma}$. \square

REMARK 5.14. *The preceding result remains true for Banach space valued complex-analytic functions $v : \mathcal{E} \rightarrow B$.*

The bound (5.6) implies convergence $I_{\Lambda(\mu, \varepsilon)}v \rightarrow v$ in $L^2_\pi(U)$ as $\varepsilon \rightarrow 0$ if $\mu \in c_0$ has asymptotic overhead order $\varkappa^*(\mu) < 1$. Combining Theorem 5.13, Lemma 5.10 and Proposition 5.12 we have the following statement: provided $\rho_m \sim m^\sigma$ with a σ larger than some fixed $\sigma_0 > 0$, the operator defined in (5.1) based on index sets $\Lambda((\rho_m^{-1})_{m \geq 1}, \cdot)$ approximates in mean-square sense a complex-valued function v which is complex-analytic on $\prod_{m \geq 1} \mathcal{E}_{\rho_m}$ with rate at least $(\sigma - \sigma_0)$ with respect to $\#\Lambda(\mu, \varepsilon)$ as $\varepsilon \rightarrow 0$ (strong tractability in the sense of [29]). *In particular, in this case the so-called curse of dimension does not appear.* The ‘‘effective dimension’’ $\max\{m \geq 1 : \mu_m \geq \varepsilon\}$, however, appears in the computation cost of each sample.

REMARK 5.15. *For a particular choice of $\mu \in c_0$ (as described below) and $\varepsilon > 0$ the multiindex set $\Lambda(\mu, \varepsilon)$ can be identified with the index set suggested in [23],*

$$X_\alpha(w, N) = \{i \in \mathbb{N}_+^N, i \geq 1 : \sum_{n=1}^N (i_n - 1)\alpha_n \leq w \min_{1 \leq n \leq N} \alpha_n\}$$

where $w \in \mathbb{R}$, $N \in \mathbb{N}$, $\alpha \in \mathbb{R}_+^N$. *Indeed, assuming without loss of generality that α is increasing and setting $\mu_m = e^{-c\alpha_m}$, $\varepsilon = \mu_M$ for $m = 1, \dots, M$, $\mu_m = 0$ for $m > M$ for a suitable $c > 0$ and $w = \log \varepsilon / \log \mu_1$ it is easy to check that $\nu \in \Lambda(\mu, \varepsilon)$ iff $i \in X_\alpha(w, M)$, where we identify $\nu_m = i_m - 1$, $m = 1, \dots, M$.*

We now compose the multilevel finite element discretization of the eigenvalue problem from Sect. 4.2 with the sparse collocation operator (5.1) based on multiindex sets of the form (5.4). For a given $\mu \in c_0$ let $(\varepsilon_j)_{j \geq 0} \in c_0$ be a sequence of *thresholds*. With $(\varepsilon_j)_{j \geq 0}$ we associate a sequence of *nested* multiindex sets $\Lambda_j = \Lambda(\mu, \varepsilon_j)$, $\Lambda_{-1} := \emptyset$. For $\nu \in \mathcal{F}$ we define $k(\nu) := \inf\{k \geq 0 : \nu \in \Lambda_k\}$, which may be infinite.

Now we consider the following *sparse composite collocation operator*, proposed for isotropic collocation in [6] (cf. also [6, Remark 6.2.5]):

$$\lambda \mapsto \hat{\lambda}_L := \sum_{0 \leq k+\ell \leq L} (I_{\Lambda_k} - I_{\Lambda_{k-1}})(\lambda_\ell - \lambda_{\ell-1}) \quad (5.12)$$

and

$$u \mapsto \hat{u}_L := \sum_{0 \leq k+\ell \leq L} (I_{\Lambda_k} - I_{\Lambda_{k-1}})(u_\ell - u_{\ell-1}). \quad (5.13)$$

LEMMA 5.16. *We have for $L \geq 0$*

$$\hat{\lambda}_L = \sum_{\nu \in \Lambda_L} I_{\{\nu\}} \lambda_{L-k(\nu)} \quad \text{and} \quad \hat{u}_L = \sum_{\nu \in \Lambda_L} I_{\{\nu\}} u_{L-k(\nu)}.$$

Proof. For $\hat{\lambda}_L$, $L \geq 0$ we compute

$$\begin{aligned} \hat{\lambda}_L &= \sum_{\ell=0}^L I_{\Lambda_{L-\ell}} (\lambda_\ell - \lambda_{\ell-1}) = \sum_{k=0}^L I_{\Lambda_k} (\lambda_{L-k} - \lambda_{L-(k+1)}) \\ &= \sum_{k=0}^L \sum_{\nu \in \Lambda_k} I_{\{\nu\}} (\lambda_{L-k} - \lambda_{L-(k+1)}) = \sum_{k=0}^L \sum_{\nu \in \Lambda_L} 1_{(k(\nu) \leq k)} I_{\{\nu\}} (\lambda_{L-k} - \lambda_{L-(k+1)}), \end{aligned}$$

and thus

$$\begin{aligned} \hat{\lambda}_L &= \sum_{\nu \in \Lambda_L} I_{\{\nu\}} \sum_{k=0}^L 1_{k(\nu) \leq k} (\lambda_{L-k} - \lambda_{L-(k+1)}) \\ &= \sum_{\nu \in \Lambda_L} I_{\{\nu\}} \sum_{k=k(\nu)}^L (\lambda_{L-k} - \lambda_{L-(k+1)}) = \sum_{\nu \in \Lambda_L} I_{\{\nu\}} \lambda_{L-k(\nu)}, \end{aligned}$$

and similarly for \hat{u}_L . \square

Note that the non-composite case corresponds to setting $k(\tilde{\nu}) = 0$ in the above.

From Theorem 5.13 we now deduce the following result.

THEOREM 5.17. *Let $W \subset H_0^1(D)$ and $t \geq 0$ be as in Assumption 4.7. Let $D \subset \mathbb{R}^d$ be convex according to Remark 4.2. Let $\{\rho_m\}_{m \geq 1}$ be a sequence with $\rho_m > 1$ and $\rho_m \rightarrow \infty$ as $m \rightarrow \infty$. Assume that $u : \mathcal{E} = \prod_{m \geq 1} \mathcal{E}_{\rho_m} \rightarrow W$ is jointly complex-analytic. Define μ_m , $m \geq 1$ as in Theorem 5.13. Assume that $\mu_m m^\sigma \rightarrow 0$ as $m \rightarrow \infty$ for some $\sigma > 0$. If $\sigma > 0$ is sufficiently large there exist $\varkappa < 1$ and $C > 0$ such that for $L \geq 0$*

$$\|u - \hat{u}_L\|_{L_\pi^2(U; H_0^1(D))} \leq C \left(\sum_{k=0}^L \varepsilon_k^{1-\varkappa} 2^{-t(L-k)} + \varepsilon_L^{1-\varkappa} \right) \|u\|_{C^0(\bar{\mathcal{E}}; W)}. \quad (5.14)$$

Proof. If $\sigma > 0$ is sufficiently large, by Prop. 5.12 and Lemma 5.10 the asymptotic overhead order of μ satisfies $\varkappa^*(\mu) < 1$. The rest follows as in [7, Sect. 6.3] using Theorem 5.13 and Corollary 4.9. \square

COROLLARY 5.18. *Assume $\varkappa^*(\mu) < \varkappa < 1$. Set $\varepsilon_k := 2^{-tk/(1-\varkappa)}$, $k \in \mathbb{N}$. Then*

$$\|u - \hat{u}_L\|_{L_\pi^2(U; H_0^1(D))} \lesssim L 2^{-tL} \|u\|_{C^0(\bar{\mathcal{E}}; W)}. \quad (5.15)$$

Next, we estimate the computational effort of the approximation $u \mapsto \hat{u}_L$. As in [29, 23, 6] for computational purposes we rewrite (5.12) in terms of the interpolation operators $\bigotimes_{m \geq 1} i_{\nu_m}$. Using Lemma 5.16 we have

$$\begin{aligned} \hat{\lambda}_L &= \sum_{\nu \in \Lambda_L} \bigotimes_{m \geq 1} (i_{\nu_m} - i_{\nu_m-1}) \lambda_{L-k(\nu)} \\ &= \sum_{\nu \in \Lambda_L} \sum_{\substack{\eta \in \{0,1\}^{\mathbb{N}} \\ \text{supp } \eta \subseteq \text{supp } \nu}} (-1)^{|\eta|} \bigotimes_{m \geq 1} i_{\nu_m - \eta_m} \lambda_{L-k(\nu)} \\ &= \sum_{\nu \in \Lambda_L} \sum_{\tilde{\nu} \in \Lambda_L} 1_{(\tilde{\nu} \leq \nu)} 1_{(|\tilde{\nu} - \nu|_{\ell^\infty(\mathbb{N})} \leq 1)} (-1)^{|\tilde{\nu} - \nu|} \bigotimes_{m \geq 1} i_{\tilde{\nu}_m} \lambda_{L-k(\nu)}. \end{aligned}$$

Finally, exchanging the sums and renaming the multiindices we obtain

$$\hat{\lambda}_L = \sum_{\nu \in \Lambda_L} \sum_{\tilde{\nu} \in \Lambda_L} 1_{(\nu \leq \tilde{\nu})} 1_{(\|\tilde{\nu} - \nu\|_{\ell^\infty(\mathbb{N})} \leq 1)} (-1)^{\|\tilde{\nu} - \nu\|_{\ell^1(\mathbb{N})}} \bigotimes_{m \geq 1} i_{\nu_m} \lambda_{L-k(\tilde{\nu})} \quad (5.16)$$

and similarly

$$\hat{u}_L = \sum_{\nu \in \Lambda_L} \sum_{\tilde{\nu} \in \Lambda_L} 1_{(\nu \leq \tilde{\nu})} 1_{(\|\tilde{\nu} - \nu\|_{\ell^\infty(\mathbb{N})} \leq 1)} (-1)^{\|\tilde{\nu} - \nu\|_{\ell^1(\mathbb{N})}} \bigotimes_{m \geq 1} i_{\nu_m} u_{L-k(\tilde{\nu})}. \quad (5.17)$$

REMARK 5.19. In (5.16) and (5.17), the number of nontrivial terms in the inner sum does not affect the overall complexity. Indeed, for $\tilde{\nu} \in \Lambda_L$ the number of collocation points in $I_{\{\nu\}} = \bigotimes_{m \geq 1} i_{\tilde{\nu}_m}$ is given by $\prod_{m \geq 1} (\tilde{\nu}_m + 1)$ which by (5.3) equals $\#\{\nu \in \Lambda_L : \nu \leq \tilde{\nu}\}$. On the other hand, if $\tilde{\nu} \in \Lambda_L$ there are at most $\#\{\nu \in \Lambda_L : \nu \leq \tilde{\nu}\}$ multiindices in the outer sum which have a non-trivial contribution from $\tilde{\nu}$ in the inner sum.

REMARK 5.20. Standard hat function discretizations can be used to compute \hat{u}_L in (5.17). Indeed, let $\hat{u}_L^{(\ell)}$, $\ell = 0, \dots, L$ be the contribution of u_ℓ in (5.17) and let P_ℓ denote the prolongation operator $V_{\ell-1} \rightarrow V_\ell$. We then have

$$\hat{u}_L = \hat{u}_L^{(L)} + P_L(\hat{u}_L^{(L-1)} + P_{L-1}(\dots P_2(\hat{u}_L^{(1)} + P_1\hat{u}_L^{(0)}) \dots)),$$

with total cost being proportional to the dimension of the ansatz space V_L .

An efficient algorithm for computing Λ_L has been given in [8]. Observe that for $\nu \in \Lambda_L = \Lambda(\mu, \varepsilon_L)$ we have $\tilde{\nu} \in \Lambda_L$ and $\tilde{\nu} \geq \nu$ iff $\eta = \tilde{\nu} - \nu$ satisfies $\eta \in \Lambda(\mu, \varepsilon_L/\mu^\nu)$. Thus, the same algorithm (with a straightforward modification to take the constraint $\|\tilde{\nu} - \nu\|_{\ell^\infty(\mathbb{N})} \leq 1$ and the variable level indicator $k(\tilde{\nu})$ into account) can be used to compute the coefficients of the terms $\bigotimes_{m \geq 1} i_{\nu_m} \lambda_{L-k(\tilde{\nu})}$ in (5.16) and $\bigotimes_{m \geq 1} i_{\nu_m} u_{L-k(\tilde{\nu})}$ in (5.17) efficiently.

Finally, the total computational effort for the application of the sparse composite collocation operator to the eigenpair of the ground state (5.16) and (5.17) can be estimated using the following lemma.

LEMMA 5.21. For $k \geq 0$ assume $\#\Lambda_k \lesssim 2^{d_1 k/2}$, $d_1 > 0$ and that the work for determination of the numerical solution u_k on one collocation node is bounded by a constant multiple of $2^{d_2 k}$, $d_2 > 0$. Then the computational effort for the numerical realization of the sparse composite operator applied to the eigenpair of the ground state is bounded by an absolute multiple of $L 2^{\max\{d_1, d_2\}L}$ as $L \rightarrow \infty$.

Proof. We compute from (5.17) using the fact $\nu \leq \tilde{\nu} \Rightarrow k(\nu) \leq k(\tilde{\nu})$ and (5.3)

$$\begin{aligned} \sum_{\nu \in \Lambda_L} \left(\prod_{m \geq 1} (\nu_m + 1) \right) L 2^{d_2(L-k(\nu))} &\lesssim L \sum_{k=0}^L \sum_{\nu \in \Lambda_k \setminus \Lambda_{k-1}} \left(\prod_{m \geq 1} (\nu_m + 1) \right) 2^{d_2(L-k)} \\ &\leq L \sum_{k=0}^L (\#\Lambda_k - \#\Lambda_{k-1}) \#\Lambda_k 2^{d_2(L-k)} \\ &\leq L \sum_{k=0}^L 2^{d_1 k} 2^{d_2(L-k)} \end{aligned}$$

which shows the claim. \square

We collect the foregoing in the following theorem.

THEOREM 5.22. *For the parametric eigenvalue problem (2.10) assume the particular form (4.1) with (4.2) and the Karhunen-Loève expansion (4.4). Assume (4.5) holds. Assume further that $D \subset \mathbb{R}^d$ is open, bounded and convex. Let $p \in (0, 1)$ be such that (4.6) holds. Then $B(p) \leq \sum_{m \geq 1} \|\psi_m\|_{L^\infty(D)}^p$. Assume $0 < \sigma \leq \frac{1}{p} - 1$. For $y \in U$ let $(\lambda(y), u(y)) = (\lambda_1(y), w_1(y))$ denote the eigenpair with smallest eigenvalue. Let $\varepsilon \in (0, 1)$, $c > 0$ and define $\tau_m = \min\{(1 - \varepsilon) \frac{a_{\min} \|\psi_m\|_{L^\infty(D)}^{p-1}}{2B(p)(1+\delta^{-1})}, c \|\psi_m\|_{W^{1,\infty}(D)}^{p-1}\}$, $m \geq 1$ and $E(\tau) = \{z \in \mathbb{C}^M : \text{dist}(z_m, [-1, 1]) < \tau_m\}$, where $\delta > 0$ is as in (4.9). Define $\rho_m = \tau_m + \sqrt{1 + \tau_m^2}$, $m \geq 1$ and $\mu_m = \sup_{m' \geq m} \frac{1}{\rho_{m'}}$, $m \geq 1$. Let $0 \leq t \leq 1$ and W be as in (4.12). If $p \in (0, 1)$ is small enough then there exists $0 < \varkappa < 1$ such that*

1. μ has asymptotic overhead order $\varkappa^*(\mu) \leq \varkappa$
2. defining $\varepsilon_k = 2^{-tk/(1-\varkappa)}$ and $\Lambda_k = \Lambda(\mu, \varepsilon_k)$ we have $\#\Lambda_k \lesssim \varepsilon_k^{-1/\sigma}$
3. the sparse composite approximation (5.13) satisfies for $L \geq 0$

$$\|u - \hat{u}_L\|_{L_\pi^2(U; H_0^1(D))} \leq C(u) L 2^{-tL} = C(u) L \varepsilon_L^{1-\varkappa}$$

and

$$\|u - \hat{u}_L\|_{L_\pi^2(U; H_0^1(D))} \leq C(u) (\#\Lambda_L)^{-\sigma(1-\varkappa)} \leq C(u) N_L^{-\sigma(1-\varkappa)/2}$$

where $C(u) = C\|u\|_{C^0(\overline{E(\tau)}; W)} < \infty$ with $C > 0$ independent of $L \geq 0$, and N_L denotes the number of collocation points in I_{Λ_L} .

6. Numerical examples. In the numerical examples we approximate the parametric eigenpairs by tensorized polynomials using the sparse collocation method as described in Sect. 5.

We take an elliptic stochastic operator expanded in its Karhunen-Loève series as a model example [8]. We set $D = (-1, 1) \subset \mathbb{R}$ and $U = [-1, 1]^\infty$ and let the diffusion coefficient in (4.1) be $a(x, y) = \bar{a}(x) + \sum_{m \geq 1} y_m a_m(x)$, $(x, y) \in D \times U$, where \bar{a} and $\{a_m\}_{m \geq 1} \subset L^\infty(D)$. Specifically, we set $\bar{a} \equiv 1$ and $a_m(x) = \frac{\cos(\pi m x)}{(m+1)^3}$, $x \in D$, $m \geq 1$. This implies $\|a_m\|_{L^\infty(D)} = \frac{1}{(m+1)^3}$, $\|\nabla a_m\|_{L^\infty(D)} = \frac{\pi m}{(m+1)^3}$ such that $\|a_m\|_{W^{1,\infty}(D)} = \frac{\pi}{(m+1)^2} + O(m^{-3})$ as $m \rightarrow \infty$. Hence, for all $p > \frac{1}{2}$ we have $(\|a_m\|_{W^{1,\infty}(D)})_{m \geq 1} \in \ell^p(\mathbb{N})$, which implies (4.6). In the computation we set $p := 0.6$, $B(p) := \sum_{m=1}^{M_0} \|a_m\|_{L^\infty(D)}^p$ with $M_0 = 10^5$, $a_{\min} := \inf_{x \in D} \bar{a} - B(1)$, $\delta := 2.8$ (empirical estimate from a few samples), $\varepsilon = 0$ and $c = 10$ in the definition of μ , see Theorem 5.22. This now completely defines the multiindex sets $\Lambda(\mu, \varepsilon)$ for all $\varepsilon > 0$.

Approximate mean $\mathbb{E}[u] = \int_U u(y) d\pi(y)$ and variance $\mathbb{E}[u^2] - \mathbb{E}[u]^2$ of the first eigenfunction u are shown in Fig. 6.1.

In Fig. 6.2 convergence of the approximate mean $\mathbb{E}[I_{\Lambda(\mu, \varepsilon)} \lambda]$ and the number of collocation points in $I_{\Lambda(\mu, \varepsilon)}$ as function of ε are shown. In space, and for the reference value of $\mathbb{E}[\lambda]$ an overkill discretization is used. We observe an algebraic rate *four* for the decay of the error $\mathbb{E}[I_{\Lambda(\mu, \varepsilon)} \lambda] - \mathbb{E}[\lambda]$ as $\varepsilon \rightarrow 0$. The number of collocation points in $I_{\Lambda(\mu, \varepsilon)}$ behaves like $\varepsilon^{-3/2}$ as $\varepsilon \rightarrow 0$.

In order to verify convergence of the parametric eigenvalue in $L_\pi^2(U)$, we employ the parameterization via the Legendre polynomials, see Lemma 5.7, which allows an exact computation of the $L_\pi^2(U)$ norm. We consider finite element spaces based either on first or second order splines on an equidistant mesh, and compare the collocation operator $P_{\Lambda_\ell} \lambda_\ell$ and the sparse composite collocation operator $\hat{\lambda}_\ell$ for $\ell = 0, 1, 2$ against an overkill reference solution. In the computation we set $\varepsilon_\ell = 2^{-\ell}$ for simplicity. The first order spaces have 15, 31, 63, the second order space have 16, 32, 64 degrees of

freedom on levels $\ell = 0, 1, 2$ respectively. The results are shown in Fig. 6.3, showing the error of the approximate parametric eigenvalue versus the total number of degrees of freedom in space, that is the sum of degrees of freedom of all EVPs solved.

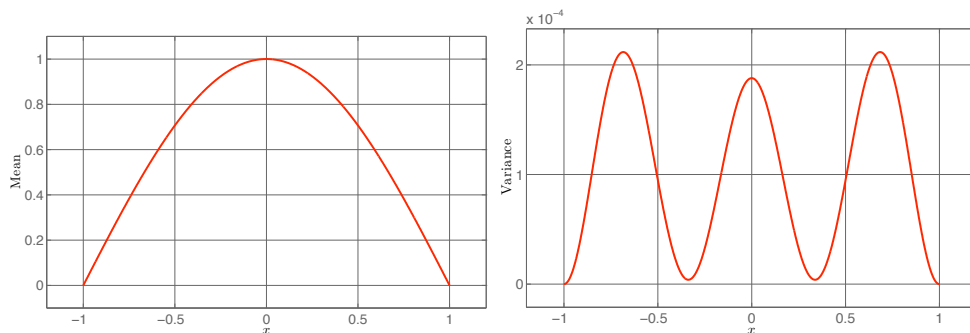


FIG. 6.1. Mean and variance of the ground state of the parametric diffusion equation (4.3) as described in Sect. 6.

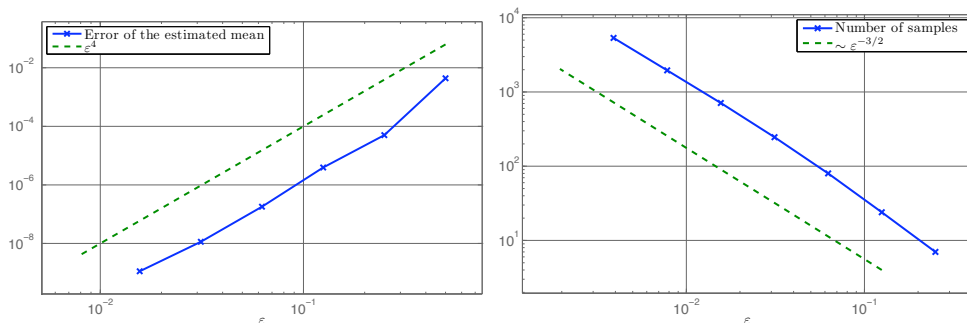


FIG. 6.2. Convergence of the mean of the computed parametric eigenvalue $\mathbb{E}[I_{\Lambda(\mu, \varepsilon)} \lambda]$ and the corresponding number of collocation points in $I_{\Lambda(\mu, \varepsilon)}$. See Sect. 6 for details.

7. Summary. We have quantified the analytic dependence of an isolated eigenpair of a linear operator depending affinely on a vector of parameters in an abstract setting. We have then specialized the discussion on stochastic differential operators expanded in its Karhunen-Loève series. Analyticity has been used to prove convergence of the sparse composite operator applied to the eigenpair. Our numerical example of an infinite dimensional parametric eigenvalue problem confirms exponential convergence of the sparse composite collocation method in the number of levels, and shows that sparse composite tensorization can be an effective tool to reduce the complexity of the problem.

In the sequel we will address the case of non-selfadjoint operators, as well as eigenpair computation with eigenvalue “crossings”.

REFERENCES

- [1] I. BABUŠKA AND B. Q. GUO, *Regularity of the solution of elliptic problems with piecewise analytic data. I. Boundary value problems for linear elliptic equation of second order*, SIAM J. Math. Anal., 19 (1988), pp. 172–203.

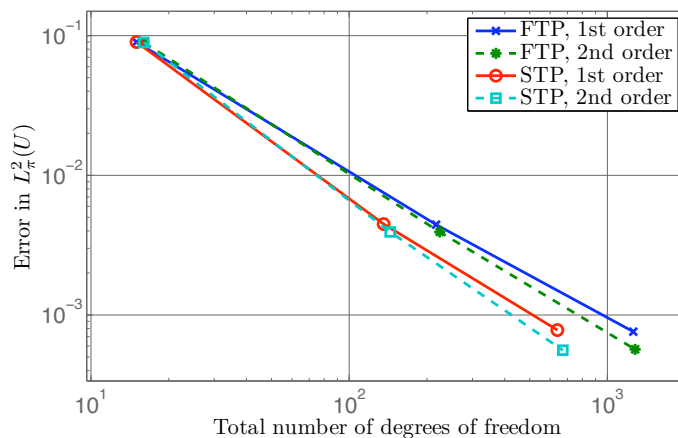


FIG. 6.3. Convergence of the eigenvalue in mean square sense for first and second order finite element, using the collocation operator $P_{\Lambda_\ell} \lambda_\ell$, $\ell = 0, 1, 2$ (FTP) and the sparse composite operator $\hat{\lambda}_\ell$, $\ell = 0, 1, 2$ (STP), see Sect. 6.

- [2] I. BABUŠKA AND J. OSBORN, *Eigenvalue problems*, in Handb. Numer. Anal., Vol. II, Handb. Numer. Anal., II, North-Holland, Amsterdam, 1991, pp. 641–787.
- [3] I. BABUŠKA, R. TEMPONE, AND G. E. ZOURARIS, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Num. Anal., 42 (2002), pp. 800–825.
- [4] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Num. Anal., 45 (2007), pp. 1005–1034.
- [5] J. BÄCK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Stochastic Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison*, Tech. Report 09-33, ICES, 2009.
- [6] M. BIERI, *A sparse composite collocation finite element method for elliptic sPDEs*, Tech. Report 2009-08, Seminar for Applied Mathematics, ETH Zürich, 2009. in review. Available via <http://www.sam.math.ethz.ch/reports/2009/08>.
- [7] ———, *Sparse tensor discretizations of elliptic PDEs with random input data*, PhD thesis, ETH Zürich, 2009. Diss ETH No. 18598 Available via <http://e-collection.ethbib.ethz.ch/>.
- [8] M. BIERI, R. ANDREEV, AND CH. SCHWAB, *Sparse tensor discretization of elliptic spdes*, SIAM J. Sci. Comput., 31 (2009), pp. 4281–4304.
- [9] M. BIERI AND CH. SCHWAB, *Sparse high order FEM for elliptic sPDEs*, Comp. Meth. Appl. Mech. Engrg., 198 (2009), pp. 1149–1170.
- [10] D. BRAESS, *Finite Elemente*, Springer, Berlin, 3rd ed., 2002.
- [11] A. COHEN, R. DEVORE, AND CH. SCHWAB, *Convergence rates of best N -term stochastic Galerkin FE-approximations for a class of elliptic sPDEs*, Tech. Report 2009-02, Seminar for Applied Mathematics, ETH Zürich, 2009. to appear in J. Found. of Comp. Math.
- [12] ———, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, Tech. Report 2010-03, Seminar for Applied Mathematics, ETH Zürich, 2010.
- [13] W. DAHMEN, T. ROHWEDDER, R. SCHNEIDER, AND A. ZEISER, *Adaptive eigenvalue computation: complexity estimates*, Numer. Math., 110 (2008), pp. 277–312.
- [14] P. J. DAVIS, *Interpolation and approximation*, Introductions to higher mathematics, Blaisdell Publishing Company, 1963.
- [15] W. GAUTSCHI, *Orthogonal polynomials. Computation and Approximation*, Numer. Math. Sci. Comput., Oxford University Press Inc., 2004.
- [16] R. GEUS, *The Jacobi-Davidson algorithm for solving large sparse symmetric eigenvalue problems with application to the design of accelerator cavities*, PhD thesis, ETH Zürich, 2002. Diss. Nr. 14734.
- [17] G. GOLUB AND C. VAN LOAN, *Matrix computations*, The Johns Hopkins University Press, London, 1996.
- [18] W. HACKBUSCH, *On the computation of approximate eigenvalues and eigenfunctions of elliptic*

- operators by means of a multi-grid method*, SIAM J. Numer. Anal., 16 (1979), pp. 201–215.
- [19] A. HENROT, *Extremum Problems for Eigenvalues of Elliptic Operators*, vol. 8 of Frontiers in Mathematics, Birkhäuser Basel, 2006.
- [20] M. HERVÉ, *Analyticity in Infinite Dimensional Spaces*, vol. 10 of De Gruyter studies in mathematics, Walter de Gruyter, 1989.
- [21] L. HÖRMANDER, *An Introduction to Complex Analysis in Several Variables*, The University Series in Higher Mathematics, D. van Nostrand Company, 1st ed., 1966.
- [22] T. KATO, *Perturbation theory for linear operators*, vol. 132 of Grundlehren Math. Wiss., Springer Berlin, Heidelberg, New-York, 2 ed., 1976.
- [23] F. NOBILE, R. TEMPONE, AND C.G. WEBSTER, *An anisotropic sparse grid stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442.
- [24] ———, *A sparse grid stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345.
- [25] M. REED AND B. SIMON, *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York, 1972.
- [26] ———, *Methods of modern mathematical physics. IV. Analysis of operators*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.
- [27] F. RELICH, *Perturbation theory of eigenvalue problems*, Notes on mathematics and its applications, Gordon and Breach, New York, London, Paris, 1969.
- [28] D. C. SORENSEN, *Numerical methods for large eigenvalue problems*, Acta Numer., 11 (2002), pp. 519–584.
- [29] G. W. WASILKOWSKI AND H. WOŹNIAKOWSKI, *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity, 11 (1995), pp. 1–56.

Research Reports

No.	Authors/Title
10-40	<i>R. Andreev and Ch. Schwab</i> Sparse tensor approximation of parametric eigenvalue problems
10-39	<i>R. Hiptmair, A. Moiola and I. Perugia</i> Stability results for the time-harmonic Maxwell equations with impedance boundary conditions
10-38	<i>I. Hnětynková, M. Plešinger, D.M. Sima, Z. Strakoš and S. Van Huffel</i> The total least squares problem in $AX \approx B$. A new classification with the relationship to the classical works
10-37	<i>S. Mishra</i> Robust finite volume schemes for simulating waves in the solar atmosphere
10-36	<i>C. Effenberger, D. Kressner and C. Engström</i> Linearization techniques for band structure calculations in absorbing photonic crystals
10-35	<i>R. Hiptmair and C. Jerez-Hanckes</i> Multiple traces boundary integral formulation for Helmholtz transmission problems
10-34	<i>H. Harbrecht and Ch. Schwab</i> Sparse tensor finite elements for elliptic multiple scale problems
10-33	<i>K. Grella and C. Schwab</i> Sparse tensor spherical harmonics approximation in radiative transfer
10-32	<i>P. Kauf, M. Torrilhon and M. Junk</i> Scale-induced closure for approximations of kinetic equations
10-31	<i>M. Hansen</i> On tensor products of quasi-Banach spaces
10-30	<i>P. Corti</i> Stable numerical scheme for the magnetic induction equation with Hall effect
10-29	<i>H. Kumar</i> Finite volume methods for the two-fluid MHD equations
10-28	<i>S. Kurz and H. Heumann</i> Transmission conditions in pre-metric electrodynamics