Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Uniformly convergent adaptive methods for parametric operator equations

C.J. Gittelson*

Research Report No. 2011-19
April 2011

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

*Research supported in part by the Swiss National Science Foundation grant No. 200021-120290/1.

Uniformly Convergent Adaptive Methods for Parametric Operator Equations

Claude Jeffrey Gittelson*

April 4, 2011

Abstract

We derive and analyze adaptive solvers for parametric boundary value problems. These methods control the error uniformly in the parameters. Numerical computations indicate that they are more efficient than similar methods that control the error in a mean square sense.

Introduction

Boundary value problems with unknown coefficients can be interpreted as parametric equations, in which the unknown coefficients are permitted to depend on a sequence of scalar parameters, for example by applying a series expansion. It may be possible to interpret the unknown coefficients as random quantities, in which case the solution to the boundary value problem is a random field. In this probabilistic setting, Galerkin methods have been developed for approximating this random field in a parametric form, see [DBO01, XK02, BTZ04, WK05, MK05, FST05, WK06, TS07, BS09, BAS10].

These methods generally require strong assumptions on the probability distribution of the random coefficients. In particular, it is often assumed that the scalar parameters, *e.g.* coming from a series expansion of the unknown coefficients, are independent. This assumption is fundamental to the construction of polynomial chaos bases. To cover the more realistic setting of non-independent parameters, an auxiliary measure is introduced *e.g.* in [BNT07, NTW08], although this still requires fairly elusive assumptions on the probability distribution.

To circumvent such assumptions, we consider numerical methods that converge uniformly in the parameter. This implies mean square convergence with respect to any probability measure on the parameter domain, in particular with respect to whatever distribution is deemed physical.

Our goal is to compute a parametric representation of the solution that is reliable to a given accuracy on the entire parameter domain. This can be combined with Monte Carlo

*Research supported in part by the Swiss National Science Foundation grant No. 200021-120290/1.

sampling in order to compute probabilistic quantities. Instead of solving the boundary value problem independently at every sample point, one can evaluate the parametric representation of the solution, which is generally much faster, see *e.g.* [WK09].

If insufficient statistical data is available for modeling unknown coefficients as random variables, then methods based on the probabilistic structure of the parameter domain do not make sense. However, uniform convergence does not make use of a probability distribution, and is thus still well defined.

The main difficulty in applying stochastic Galerkin methods is the construction of suitable spaces in which to compute approximate solutions. In [Git11b, Git11c], we suggest adaptive methods based on techniques from the adaptive wavelet algorithms [CDD01, CDD02, GHS07, DSS09]. We use an orthonormal polynomial basis on the parameter domain in place of wavelets. An arbitrary discretization of the physical domain can be used to approximate the coefficients of the random solution with respect to this basis.

In order to ensure uniform convergence in the parameter, we deviate a bit further from adaptive wavelet methods, which are formulated in a Hilbert space setting. We follow the approach in [CDD02, Git11c], which is based on applying an iterative method directly to the full parametric boundary value problem. Individual substeps of this iteration, such as application of the parametric operator, are replaced by approximate counterparts, realized by suitable adaptive algorithms. These keep track of errors entering the computation, ensuring convergence of the algorithm, and providing an upper bound on the error of the approximate solution.

In Section 1, we study parametric operator equations in an abstract setting. We show that the parametric solution depends continuously on the parameter. In the setting that the operator has a dominant nonparametric component, we derive a perturbed stationary linear iteration, which forms the basis for our adaptive method. We also present an illustrative example for a parametric boundary value problem.

Our method is formulated on the level of coefficients with respect to a polynomial basis. In Section 2, we apply the Stone–Weierstrass theorem to show that continuous functions can be approximated uniformly by polynomials in an infinite dimensional setting. We construct suitable polynomial bases, and represent a class of parametric operators in these bases.

We present our adaptive method in Section 3. A vital component is an adaptive routine for applying the parametric operator, which is discussed in Section 3.1. In Section 4, we present a variant of our adaptive solver which has the potential to reduce the computational cost while maintaining the same accuracy.

In Section 5, we apply these adaptive solvers to a simple model problem. Numerical computations demonstrate the convergence of the algorithms and compare them to the adaptive methods from [Git11b, Git11c].

1 Parametric Operator Equations

1.1 Continuous Parameter Dependence

Let V and W be Banach spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Denote by W^* the space of bounded antilinear maps from W to $\mathbb{K}$, and by $\mathcal{L}(V, W^*)$ the Banach space of bounded linear maps from V to W^*.

Let Γ be a nonempty topological space. A parametric linear operator from V to W^* with parameter domain Γ is a continuous map

$$A\colon \Gamma \to \mathcal{L}(V, W^*)\,, \quad y \mapsto A(y)\,. \tag{1.1}$$

For a given $f\colon \Gamma \to W^*$, we are interested in determining $u\colon \Gamma \to V$ such that

$$A(y)u(y) = f(y) \qquad \forall y \in \Gamma\,. \tag{1.2}$$

Assumption 1.A. *$A(y)$ is bijective for all $y \in \Gamma$.*

By the open mapping theorem, Assumption 1.A implies that $A(y)$ is boundedly invertible for all $y \in \Gamma$.

Theorem 1.1. *Equation* (1.2) *has a unique solution $u\colon \Gamma \to V$. It is continuous if and only if $f\colon \Gamma \to W^*$ is continuous.*

Proof. By Assumption 1.A, (1.2) has the unique solution $u(y) = A(y)^{-1} f(y)$.

Let $D \in \mathcal{L}(V, W^*)$ be boundedly invertible. For example, D could be equal to $A(y)$ for some $y \in \Gamma$. Then $y \mapsto D^{-1}A(y)$ is a continuous map from Γ into $\mathcal{L}(V)$. By the abstract property [KR97, Prop. 3.1.6] of Banach algebras, the map $\mathrm{inv}\colon T \mapsto T^{-1}$ defined on the multiplicative group of $\mathcal{L}(V)$ is continuous in the topology of $\mathcal{L}(V)$. Therefore, $y \mapsto \mathrm{inv}(D^{-1}A(y)) = A(y)^{-1}D$ is continuous, and multiplying from the right by the constant D^{-1}, it follows that $y \mapsto A(y)^{-1}$ is a continuous map from Γ to $\mathcal{L}(W^*, V)$.

Note that the application of an operator to a vector, *i.e.* the map $\mathrm{mult}\colon \mathcal{L}(W^*, V) \times W^* \to V$ defined by $\mathrm{mult}(T, z) := Tz$, is continuous. Therefore, if $y \mapsto f(y)$ is continuous, then so is $y \mapsto u(y) = \mathrm{mult}(A(y)^{-1}, f(y))$. Similarly, since $\mathrm{mult}\colon \mathcal{L}(V, W^*) \times V \to W^*$ is continuous, if $y \mapsto u(y)$ is continuous, then so is $y \mapsto f(y) = \mathrm{mult}(A(y), u(y))$. □

Example 1.2. Assumption 1.A is assured to hold if $A(y)$ is a perturbation of a boundedly invertible $D \in \mathcal{L}(V, W^*)$, *i.e.*

$$A(y) = D + R(y)\,, \quad y \in \Gamma\,, \tag{1.3}$$

with a continuous $y \mapsto R(y) \in \mathcal{L}(V, W^*)$ satisfying

$$\left\|D^{-1}R(y)\right\|_{V\to V} \le \gamma < 1 \qquad \forall y \in \Gamma\,. \tag{1.4}$$

Then $A(y)$ can be decomposed as

$$A(y) = D(\mathrm{id}_V + D^{-1}R(y))\,, \quad y \in \Gamma\,, \tag{1.5}$$

and consequently, using a Neumann series in $\mathcal{L}(V)$ to invert the second factor,

$$A(y)^{-1} = \left(\sum_{n=0}^{\infty} \left(-D^{-1} R(y) \right)^n \right) D^{-1} , \quad y \in \Gamma . \tag{1.6}$$

In this setting, due to (1.4), (1.5) and (1.6), the parametric operators $A(y)$ and $A(y)^{-1}$ are uniformly bounded,

$$\left\| A(y) \right\|_{V \to W^*} \leq \|D\|_{V \to W^*} (1 + \gamma) \qquad \forall y \in \Gamma , \tag{1.7}$$

$$\left\| A(y)^{-1} \right\|_{W^* \to V} \leq \frac{\left\| D^{-1} \right\|_{W^* \to V}}{1 - \gamma} \qquad \forall y \in \Gamma . \tag{1.8}$$

A sufficient condition for (1.4) is

$$\left\| R(y) \right\|_{V \to W^*} \leq \frac{\gamma}{\left\| D^{-1} \right\|_{W^* \to V}} \qquad \forall y \in \Gamma \tag{1.9}$$

with $\gamma < 1$. Equation (1.9) does not depend on the precise structure of the operators $D^{-1}R(y)$. Therefore, Assumption 1.A is always satisfied if the parametric component $R(y)$ of $A(y)$ is sufficiently small. ⌟

Assumption 1.B. *Γ is a compact Hausdorff space.*

Lemma 1.3. *There exist constants $\hat{c}, \check{c} \in \mathbb{R}$ such that*

$$\left\| A(y) \right\|_{V \to W^*} \leq \hat{c} \qquad \text{and} \qquad \left\| A(y)^{-1} \right\|_{W^* \to V} \leq \check{c} \qquad \forall y \in \Gamma . \tag{1.10}$$

Proof. By assumption, the map $y \mapsto A(y)$ is continuous. As shown in the proof of Theorem 1.1, $y \mapsto A(y)^{-1}$ is also continuous. Consequently, the maps $y \mapsto \left\| A(y) \right\|_{V \to W^*}$ and $y \mapsto \left\| A(y)^{-1} \right\|_{W^* \to V}$ are continuous maps from Γ into $\mathbb{R}$. Since Γ is compact by Assumption 1.B, the ranges of these maps are compact in $\mathbb{R}$, and therefore bounded. □

For any Banach space X, let $C(\Gamma; X)$ denote the Banach space of continuous maps from Γ to X with norm

$$\|v\|_{C(\Gamma;X)} := \sup_{y \in \Gamma} \left\| v(y) \right\|_X , \quad v \in C(\Gamma; X) . \tag{1.11}$$

In what follows, we abbreviate $C(\Gamma) := C(\Gamma; \mathbb{K})$.

Corollary 1.4. *The operators*

$$\mathcal{A} \colon C(\Gamma; V) \to C(\Gamma; W^*) , \quad v \mapsto [y \mapsto A(y)v(y)] \qquad \text{and} \tag{1.12}$$

$$\mathcal{A}^{-1} \colon C(\Gamma; W^*) \to C(\Gamma; V) , \quad g \mapsto \left[y \mapsto A(y)^{-1} g(y) \right] \tag{1.13}$$

are well-defined, inverse to each other, and bounded with norms $\|\mathcal{A}\| \leq \hat{c}$ and $\left\| \mathcal{A}^{-1} \right\| \leq \check{c}$.

Proof. The assertion is a direct consequence of Theorem 1.1 and Lemma 1.3. □

1.2 A Perturbed Linear Iteration

We consider the setting of Example 1.2, *i.e.* $\mathcal{A}$ is a sum

$$\mathcal{A} = \mathcal{D} + \mathcal{R} \tag{1.14}$$

with $\mathcal{D}\colon C(\Gamma; V) \to C(\Gamma; W^*)$ of the form $(\mathcal{D}v)(y) = Dv(y)$ for a boundedly invertible $D \in \mathcal{L}(V, W^*)$, and $\mathcal{R}\colon C(\Gamma; V) \to C(\Gamma; W^*)$ satisfies

$$\left\|\mathcal{D}^{-1}\mathcal{R}\right\|_{C(\Gamma;V)\to C(\Gamma;V)} \le \gamma < 1 \,. \tag{1.15}$$

Condition (1.15) implies that $\mathcal{D}^{-1}\mathcal{A}$ can be inverted by a Neumann series in $C(\Gamma; V)$, and

$$\left\|(\mathcal{D}^{-1}\mathcal{A})^{-1}\right\|_{C(\Gamma;V)\to C(\Gamma;V)} \le \frac{1}{1-\gamma} \,. \tag{1.16}$$

Therefore, for any $f \in C(\Gamma; W^*)$, the solution u of the operator equation

$$\mathcal{A}u = f \tag{1.17}$$

is the limit of the sequence $(u_k)_{k=0}^{\infty}$ with arbitrary $u_0 \in C(\Gamma; V)$ and

$$u_k := \mathcal{D}^{-1}(f - \mathcal{R}u_{k-1}) \,, \quad k \in \mathbb{N} \,. \tag{1.18}$$

We generalize (1.18) by allowing errors in the computation of f, the evaluation of $\mathcal{R}$, and the inversion of $\mathcal{D}$.

Let $\delta_0 \ge \|u - u_0\|_{C(\Gamma;V)}$. For example, if $u_0 = 0$, we may set

$$\delta_0 := \frac{\left\|D^{-1}\right\|_{W^*\to V}}{1-\gamma} \left\|f\right\|_{C(\Gamma;W^*)} \,. \tag{1.19}$$

For all $k \in \mathbb{N}$, let $g_k \in C(\Gamma; W^*)$ with

$$\left\|g_k - (f - \mathcal{R}u_{k-1})\right\|_{C(\Gamma;W^*)} \le \beta\delta_{k-1} \left\|D^{-1}\right\|_{W^*\to V}^{-1} \,, \tag{1.20}$$

and let $u_k \in C(\Gamma; V)$ satisfy

$$\left\|u_k - \mathcal{D}^{-1}g_k\right\|_{C(\Gamma;V)} \le \alpha\delta_{k-1} \,, \tag{1.21}$$

where δ_{k-1} is an upper bound for $\|u - u_{k-1}\|_{C(\Gamma;V)}$ and $\alpha, \beta \ge 0$ are independent of k.

Theorem 1.5. *Let u_k and g_k satisfy* (1.20) *and* (1.21) *for any $\delta_{k-1} \ge \|u - u_{k-1}\|_{C(\Gamma;V)}$. Then*

$$\|u - u_k\|_{C(\Gamma;V)} \le (\alpha + \beta + \gamma)\delta_{k-1} =: \delta_k \,. \tag{1.22}$$

In particular, if $\alpha + \beta < 1 - \gamma$, then $u_k \to u$ in $C(\Gamma; V)$, and

$$\|u - u_k\|_{C(\Gamma;V)} \le (\alpha + \beta + \gamma)^k \delta_0 \qquad \forall k \in \mathbb{N}_0 \,. \tag{1.23}$$

Proof. Since $\mathcal{D}u = f - \mathcal{R}u$,

$$u - u_k = \mathcal{D}^{-1}(f - \mathcal{R}u) - \mathcal{D}^{-1}(f - \mathcal{R}u_{k-1}) + \mathcal{D}^{-1}(f - \mathcal{R}u_{k-1} - g_k) + \mathcal{D}^{-1}g_k - u_k \ .$$

By triangle inequality,

$$\begin{aligned}\|u - u_k\|_{C(\Gamma;V)} \leq\ & \left\|\mathcal{D}^{-1}\mathcal{R}(u - u_{k-1})\right\|_{C(\Gamma;V)} + \left\|D^{-1}\right\|_{W^*\to V} \left\|g_k - (f - \mathcal{R}u_{k-1})\right\|_{C(\Gamma;W^*)} \\ & + \left\|u_k - \mathcal{D}^{-1}g_k\right\|_{C(\Gamma;V)} \\ \leq\ & \gamma \|u - u_{k-1}\|_{C(\Gamma;V)} + \beta\delta_{k-1} + \alpha\delta_{k-1} \ .\end{aligned}$$

Equation (1.22) follows by the assumption that δ_{k-1} is greater than $\|u - u_{k-1}\|_{C(\Gamma;V)}$. If $\alpha + \beta + \gamma < 1$, repeated application with δ_k defined as in (1.22) leads to (1.23). □

Remark 1.6. Theorem 1.5 uses a priori known quantities $\delta_k = (\alpha + \beta + \gamma)^k\delta_0$ as upper bounds for the error at iteration $k \in \mathbb{N}_0$. However, better estimates may be available or computable during an iteration. The residual at iteration $k \in \mathbb{N}_0$ is given by

$$r_k \coloneqq f - \mathcal{A}u_k = \mathcal{A}(u - u_k) \in C(\Gamma; W^*) \ . \tag{1.24}$$

Since $\mathcal{A}$ is invertible by a Neumann series,

$$\|u - u_k\|_{C(\Gamma;V)} \leq \left\|\mathcal{A}^{-1}\right\|_{C(\Gamma;W^*)\to C(\Gamma;V)} \|r_k\|_{C(\Gamma;W^*)} \leq \frac{1}{1-\gamma} \left\|D^{-1}\right\|_{W^*\to V} \|r_k\|_{C(\Gamma;W^*)} \ . \tag{1.25}$$

Therefore, if it is known that $\|r_k\|_{C(\Gamma;W^*)} \leq \rho_k$, we also have the upper bound

$$\bar{\delta}_k \coloneqq \frac{1}{1-\gamma} \left\|D^{-1}\right\|_{W^*\to V} \rho_k \tag{1.26}$$

of $\|u - u_k\|_{C(\Gamma;V)}$ for all $k \in \mathbb{N}$. ⌟

1.3 The Parametric Diffusion Equation

As an illustrative example, we consider the isotropic diffusion equation on a bounded Lipschitz domain $G \subset \mathbb{R}^d$ with homogeneous Dirichlet boundary conditions. For any uniformly positive $a \in L^\infty(G)$ and any $f \in L^2(G)$, we have

$$\begin{aligned} -\nabla \cdot (a(x)\nabla u(x)) &= f(x) \ , \quad x \in G \ , \\ u(x) &= 0 \ , \quad x \in \partial G \ . \end{aligned} \tag{1.27}$$

We model a as a parametric coefficient, depending affinely on a sequence of scalar parameters in $[-1, 1]$. For the compact parameter domain $\Gamma \coloneqq [-1, 1]^\infty$, we have

$$a(y, x) \coloneqq \bar{a}(x) + \sum_{m=1}^{\infty} y_m a_m(x) \ , \quad y = (y_m)_{m=1}^\infty \in \Gamma \ . \tag{1.28}$$

Thus the parameters y_m are coefficients in a series expansion of $a(y, x) - \bar{a}(x)$.

We define the parametric operator

$$A(y)\colon H_0^1(G) \to H^{-1}(G) \ , \quad v \mapsto -\nabla \cdot (a(y)\nabla v) \ , \quad y \in \Gamma \ . \tag{1.29}$$

By linearity, we can expand $A(y)$ as

$$A(y) = D + R(y) \ , \quad R(y) \coloneqq \sum_{m=1}^{\infty} y_m R_m \qquad \forall y \in \Gamma \ , \tag{1.30}$$

for

$$\begin{aligned} D\colon H_0^1(G) &\to H^{-1}(G) \ , \quad v \mapsto -\nabla \cdot (\bar{a}\nabla v) \ , \\ R_m\colon H_0^1(G) &\to H^{-1}(G) \ , \quad v \mapsto -\nabla \cdot (a_m \nabla v) \ , \quad m \in \mathbb{N} \ . \end{aligned}$$

Note that $\|R_m\|_{H_0^1(G)\to H^{-1}(G)} \leq \|a_m\|_{L^\infty(G)}$, and thus convergence in (1.30) and (1.28) is assured if $(\|a_m\|_{L^\infty(G)})_{m=1}^\infty$ is summable.

Assuming that $\bar{a}$ is bounded and uniformly positive, the operator D is invertible with

$$\left\|D^{-1}\right\|_{H^{-1}(G)\to H_0^1(G)} \leq \left(\operatorname*{ess\,inf}_{x\in G} \bar{a}(x)\right)^{-1} . \tag{1.31}$$

2 Polynomial Expansion

2.1 Uniform Approximation by Polynomials in Infinite Dimensions

We consider polynomials on the compact domain $\Gamma \coloneqq [-1,1]^\infty$. We denote a generic element of Γ by $y = (y_m)_{m=1}^\infty$.

For any finite set $F \subset \mathbb{N}$, let $\mathcal{P}_F(\Gamma)$ denote the vector space of polynomials in the variables $(y_m)_{m\in F}$. Then

$$\mathcal{P}(\Gamma) \coloneqq \bigcup_{\substack{F\subset\mathbb{N} \\ \#F<\infty}} \mathcal{P}_F(\Gamma) \tag{2.1}$$

is the vector space of polynomials on the infinite dimensional domain Γ.

Theorem 2.1. *The space $\mathcal{P}(\Gamma)$ is dense in $C(\Gamma)$.*

Proof. Since all polynomials are continuous, $\mathcal{P}(\Gamma) \subset C(\Gamma)$. Furthermore, $\mathcal{P}(\Gamma)$ is closed under addition and pointwise multiplication, so it is a subalgebra of $C(\Gamma)$. It contains the constant functions on Γ, and for any $y, z \in \Gamma$, $y \neq z$, there is an $m \in \mathbb{N}$ such that $y_m \neq z_m$. Therefore, y and z are separated by the linear polynomial y_m. Consequently, the assertion follows from the Stone–Weierstrass theorem, see *e.g.* [Sto48, Rud91]. □

A Banach space X is said to have the approximation property if, for every compact set $K \subset X$ and every $\epsilon > 0$, there is a finite rank operator $T \in \mathcal{L}(X)$ such that

$$\|x - Tx\|_X \leq \epsilon \quad \forall x \in K \ . \tag{2.2}$$

We recall that every space with a Schauder basis has the approximation property. In particular, every separable Hilbert space has this property.

Theorem 2.1 extends to functions with values in X under the assumption that X has the approximation property. For any finite set $F \subset \mathbb{N}$, let $\mathcal{P}_F(\Gamma; X)$ denote the vector space of polynomials in the variables $(y_m)_{m\in F}$ with coefficients in X. As in (2.1), we define

$$\mathcal{P}(\Gamma; X) := \bigcup_{\substack{F\subset\mathbb{N}\\ \#F<\infty}} \mathcal{P}_F(\Gamma; X) \ . \tag{2.3}$$

Theorem 2.2. *If X has the approximation property, then $\mathcal{P}(\Gamma; X)$ is dense in $C(\Gamma; X)$.*

Proof. Let $f \in C(\Gamma; X)$ and $\epsilon > 0$. Since Γ is compact, $K := f(\Gamma) \subset X$ is compact, and thus there is a finite rank operator $T \in \mathcal{L}(X)$ such that (2.2) holds. We write T as

$$Tx = \sum_{i=1}^{n} \psi_i(x) x_i$$

with $\psi_i \in X^*$ and $x_i \in X$, scaled such that $\|x_i\|_X = 1$. Since each of the maps $\psi_i \circ f$ is in $C(\Gamma)$, Theorem 2.1 implies that there is a polynomial $p_i \in \mathcal{P}(\Gamma)$ with $\left|p_i(y) - \psi(f(y))\right| \le \epsilon/n$ for all $y \in \Gamma$. Consequently, for all $y \in \Gamma$,

$$\left\| f(y) - \sum_{i=1}^{n} p_i(y) x_i \right\|_X \le \left\| f(y) - Tf(y) \right\|_X + \sum_{i=1}^{n} \left|\psi_i(f(y)) - p_i(y)\right| \|x_i\|_X \le 2\epsilon \ . \qquad \square$$

2.2 Polynomial Systems in Infinite Dimensions

Let $(P_n)_{n=0}^{\infty}$ be a sequence of polynomials on $[-1, 1]$ satisfying $P_0(\xi) = 1$, $P_1(\xi) = \xi$ and

$$\xi P_n(\xi) = \pi_n^+ P_{n+1}(\xi) + \pi_n^- P_{n-1}(\xi) \qquad \forall n \in \mathbb{N} \tag{2.4}$$

for all $\xi \in [-1, 1]$. In particular, it follows by induction that P_n is a polynomial of degree n. We define $\pi_0^+ := 1$ in order to achieve $\xi P_0(\xi) = \pi_0^+ P_1(\xi)$.

For example, $P_n(\xi) = \xi^n$ if $\pi_n^+ = 1$ and $\pi_n^- = 0$ for all $n \in \mathbb{N}$. If π_n^+ and π_n^- are given by

$$\pi_n^+ := \frac{1}{2} \qquad \text{and} \qquad \pi_n^- := \frac{1}{2} \ , \tag{2.5}$$

then $(P_n)_{n=0}^{\infty}$ are Chebyshev polynomials of the first kind. Alternatively, the values

$$\pi_n^+ := \frac{n+1}{2n+1} \qquad \text{and} \qquad \pi_n^- := \frac{n}{2n+1} \tag{2.6}$$

lead to Legendre polynomials. More generally, identities of the type (2.4) follow from recursion formulas for families of orthonormal polynomials with respect to symmetric measures on $[-1, 1]$, see *e.g.* [Gau04, Sze75].

In both of the above examples,

$$|P_n(\xi)| \le 1 \quad \forall \xi \in [-1,1]\,, \quad \forall n \in \mathbb{N}_0\,. \tag{2.7}$$

We assume that the polynomials $(P_n)_{n=0}^{\infty}$ are scaled in such a way that (2.7) holds.

We define the set of finitely supported sequences in $\mathbb{N}_0$ as

$$\Lambda := \left\{\mu \in \mathbb{N}_0^{\mathbb{N}}\,;\, \#\operatorname{supp}\mu < \infty\right\}\,, \tag{2.8}$$

where the support is defined by

$$\operatorname{supp}\mu := \{m \in \mathbb{N}\,;\, \mu_m \neq 0\}\,, \quad \mu \in \mathbb{N}_0^{\mathbb{N}}\,. \tag{2.9}$$

Then countably infinite tensor product polynomials are given by

$$(P_\mu)_{\mu\in\Lambda}\,, \quad P_\mu := \bigotimes_{m=1}^{\infty} P_{\mu_m}\,, \quad \mu \in \Lambda\,. \tag{2.10}$$

Note that each of these functions depends on only finitely many dimensions,

$$P_\mu(y) = \prod_{m=1}^{\infty} P_{\mu_m}(y_m) = \prod_{m\in\operatorname{supp}\mu} P_{\mu_m}(y_m)\,, \quad \mu \in \Lambda\,, \tag{2.11}$$

since $P_0 = 1$.

Proposition 2.3. *If X is a Banach space with the approximation property, then for any $f \in C(\Gamma;X)$ and any $\epsilon > 0$, there is a finite set $\Xi \subset \Lambda$ and $x_\mu \in X$, $\mu \in \Xi$, such that*

$$\max_{y\in\Gamma} \left\| f(y) - \sum_{\mu\in\Xi} x_\mu P_\mu(y) \right\|_X \le \epsilon\,. \tag{2.12}$$

Proof. The assertion follows from Theorem 2.2 since $(P_\mu)_{\mu\in\Lambda}$ is an algebraic basis of the vector space $\mathcal{P}(\Gamma;X)$. □

2.3 Representation of a Class of Parametric Operators

Let V and W be Banach spaces. Motivated by Example 1.2 and (1.30), we consider $\Gamma = [-1,1]^\infty$ and operators $\mathcal{A}\colon C(\Gamma;V) \to C(\Gamma;W^*)$ of the form $\mathcal{A} = \mathcal{D}+\mathcal{R}$ with $(\mathcal{D}v)(y) = Dv(y)$ for a boundedly invertible $D \in \mathcal{L}(V,W^*)$ and

$$(\mathcal{R}v)(y) := \sum_{m=1}^{\infty} y_m R_m v(y)\,, \quad y \in \Gamma\,, \quad v \in C(\Gamma;V)\,, \tag{2.13}$$

for $R_m \in \mathcal{L}(V,W^*)$ satisfying

$$\sum_{m=1}^{\infty} \left\|D^{-1}R_m\right\|_{V\to V} \le \gamma < 1\,. \tag{2.14}$$

We note that these assumptions imply (1.15). Truncating the series in (2.13), we approximate $\mathcal{R}$ by

$$(\mathcal{R}_{[M]}v)(y) \coloneqq \sum_{m=1}^{M} y_m R_m v(y) \;, \quad y \in \Gamma \;, \quad v \in C(\Gamma; V) \;, \tag{2.15}$$

for $M \in \mathbb{N}$, and $\mathcal{R}_{[0]} \coloneqq 0$.

Lemma 2.4. *For all $M \in \mathbb{N}_0$,*

$$\left\|\mathcal{R} - \mathcal{R}_{[M]}\right\|_{C(\Gamma;V)\to C(\Gamma;W^*)} \leq \sum_{m=M+1}^{\infty} \|R_m\|_{V\to W^*} \;. \tag{2.16}$$

In particular, $\mathcal{R}_{[M]} \to \mathcal{R}$ in $\mathcal{L}(C(\Gamma;V), C(\Gamma;W^*))$.

Proof. For any $M \in \mathbb{N}_0$, $v \in C(\Gamma;V)$ and $y \in \Gamma$, since $\left|y_m\right| \leq 1$,

$$\left\|(\mathcal{R}v)(y) - (\mathcal{R}_{[M]}v)(y)\right\|_{W^*} \leq \sum_{m=M+1}^{\infty} \left|y_m\right| \left\|R_m v(y)\right\|_{W^*} \leq \sum_{m=M+1}^{\infty} \|R_m\|_{V\to W^*} \|v\|_{C(\Gamma;V)} \;.$$

Furthermore, (2.14) implies that $(R_m)_{m\in\mathbb{N}} \in \ell^1(\mathbb{N}; \mathcal{L}(V, W^*))$. □

According to the following statement, $\mathcal{R}_{[M]}$ maps $\mathcal{P}(\Gamma;V)$ into $\mathcal{P}(\Gamma;W^*)$. We determine the coefficients of $\mathcal{R}_{[M]}v$ in terms of those of $v \in \mathcal{P}(\Gamma;V)$ with respect to a polynomial basis $(P_\mu)_{\mu\in\Lambda}$ from Section 2.2.

Lemma 2.5. *For any $M \in \mathbb{N}$ and any $v \in \mathcal{P}(\Gamma;V)$, represented as*

$$v(y) = \sum_{\mu\in\Xi} v_\mu P_\mu(y) \;, \quad y \in \Gamma \;, \tag{2.17}$$

for a finite set $\Xi \subset \Lambda$, $\mathcal{R}_{[M]}v \in \mathcal{P}(\Gamma;W^*)$ has the form

$$(\mathcal{R}_{[M]}v)(y) = \sum_{\mu\in\Xi} \sum_{m=1}^{M} R_m v_\mu \Big(\pi^+_{\mu_m} P_{\mu+\epsilon_m}(y) + \pi^-_{\mu_m} P_{\mu-\epsilon_m}(y)\Big) \;, \quad y \in \Gamma \;, \tag{2.18}$$

where $\epsilon_m \in \Lambda$ is the Kronecker sequence $(\epsilon_m)_n \coloneqq \delta_{mn}$, and we set $P_\mu \coloneqq 0$ if any $\mu_m < 0$.

Proof. By the definitions (2.15) and (2.17),

$$(\mathcal{R}_{[M]}v)(y) = \sum_{\mu\in\Xi} \sum_{m=1}^{M} R_m v_\mu y_m P_\mu(y) \;.$$

Equation (2.4) implies

$$y_m P_\mu(y) = \pi^+_{\mu_m} P_{\mu+\epsilon_m}(y) + \pi^-_{\mu_m} P_{\mu-\epsilon_m}(y) \;.$$ □

Combining Proposition 2.3, Lemma 2.4 and Lemma 2.5, one can represent $\mathcal{R}v$ as a limit of terms of the form (2.18) for any $v \in C(\Gamma;V)$, provided that V has the approximation property.

3 A Uniformly Convergent Adaptive Solver

3.1 Adaptive Application of Parametric Operators

We consider operators $\mathcal{R}$ of the form (2.13). For all $M \in \mathbb{N}$, let $\bar{e}_{\mathcal{R},M}$ be given such that

$$\left\|\mathcal{R} - \mathcal{R}_{[M]}\right\|_{C(\Gamma;V)\to C(\Gamma;W^*)} \le \bar{e}_{\mathcal{R},M} \,. \tag{3.1}$$

For example, by Lemma 2.4, these bounds can be chosen as

$$\bar{e}_{\mathcal{R},M} := \sum_{m=M+1}^{\infty} \|R_m\|_{V\to W^*} \,, \tag{3.2}$$

or as estimates for these sums. We assume that $(\bar{e}_{\mathcal{R},M})_{M=0}^{\infty}$ is nonincreasing and converges to 0, and also that the sequence of differences $(\bar{e}_{\mathcal{R},M} - \bar{e}_{\mathcal{R},M+1})_{M=0}^{\infty}$ is nonincreasing. For (3.2), the latter property holds if R_m are arranged in decreasing order of $\|R_m\|_{V\to W^*}$.

Alternative values for $\bar{e}_{\mathcal{R},M}$ are provided by the following elementary estimate, which is a direct consequence of Lemma 2.4, see [Git11b, Prop. 4.4].

Proposition 3.1. *Let $s > 0$. If either*

$$\|R_m\|_{V\to W^*} \le s\delta_{\mathcal{R},s}(m+1)^{-s-1} \qquad \forall m \in \mathbb{N} \tag{3.3}$$

or the sequence $(\|R_m\|_{V\to W^*})_{m=1}^{\infty}$ is nonincreasing and

$$\left(\sum_{m=1}^{\infty} \|R_m\|_{V\to W^*}^{\frac{1}{s+1}}\right)^{s+1} \le \delta_{\mathcal{R},s} \,, \tag{3.4}$$

then

$$\left\|\mathcal{R} - \mathcal{R}_{[M]}\right\|_{C(\Gamma;V)\to C(\Gamma;W^*)} \le \delta_{\mathcal{R},s}(M+1)^{-s} \qquad \forall M \in \mathbb{N}_0 \,. \tag{3.5}$$

We consider the approximation of a vector $\boldsymbol{w} \in \ell^1(\Lambda)$ by a sum $\boldsymbol{w}_{[1]} + \cdots + \boldsymbol{w}_{[P]}$ for sections $\boldsymbol{w}_{[p]} := \boldsymbol{w}|_{\Lambda_p}$, $p = 1, \dots, P$, with mutually disjoint sets $\Lambda_p \subset \Lambda$. The section $\boldsymbol{w}_{[1]}$ can be thought of as containing the largest elements of $\boldsymbol{w}$, $\boldsymbol{w}_{[2]}$ the next largest, and so on.

Such a partitioning can be constructed by the approximate sorting algorithm

$$\texttt{BucketSort}[\boldsymbol{w}, \epsilon] \mapsto \left[(\boldsymbol{w}_{[p]})_{p=1}^{P}, (\Lambda_p)_{p=1}^{P}\right] \,, \tag{3.6}$$

which, given a finitely supported $\boldsymbol{w} \in \ell^1(\Lambda)$ and a threshold $\epsilon > 0$, returns index sets

$$\Lambda_p := \left\{\mu \in \Lambda \,;\, |w_\mu| \in (2^{-p}\,\|\boldsymbol{w}\|_{\ell^\infty}, 2^{-(p-1)}\,\|\boldsymbol{w}\|_{\ell^\infty}]\right\} \tag{3.7}$$

and $\boldsymbol{w}_{[p]} := \boldsymbol{w}|_{\Lambda_p}$, see [Met02, Bar05, GHS07, DSS09]. The integer P is minimal with

$$2^{-P}\,\|\boldsymbol{w}\|_{\ell^\infty(\Lambda)}\,\#\operatorname{supp}\boldsymbol{w} \le \epsilon \,. \tag{3.8}$$

$\texttt{Apply}_{\mathcal{R}}[v, \epsilon] \mapsto z$

$[\cdot, (\Lambda_p)_{p=1}^P] \longleftarrow \texttt{BucketSort}\left[(\left\|v_\mu\right\|_V)_{\mu\in\Lambda}, \dfrac{\epsilon}{2\bar{e}_{\mathcal{R},0}}\right]$

for $p = 1, \dots, P$ **do** $v_{[p]} \longleftarrow (v_\mu)_{\mu\in\Lambda_p}$

Compute the minimal $\ell \in \{0, 1, \dots, P\}$ s.t. $\delta := \bar{e}_{\mathcal{R},0} \left\| v - \sum_{p=1}^{\ell} v_{[p]} \right\|_{\ell^1(\Lambda;V)} \le \dfrac{\epsilon}{2}$

for $p = 1, \dots, \ell$ **do** $M_p \longleftarrow 0$

while $\sum_{p=1}^{\ell} \bar{e}_{\mathcal{R},M_p} \left\|v_{[p]}\right\|_{\ell^1(\Lambda;V)} > \epsilon - \delta$ **do**

  $q \longleftarrow \operatorname{argmax}_{p=1,\dots,\ell}(\bar{e}_{\mathcal{R},M_p} - \bar{e}_{\mathcal{R},M_p+1}) \left\|v_{[p]}\right\|_{\ell^1(\Lambda;V)} / \#\Lambda_p$

  $M_q \longleftarrow M_q + 1$

$z = (z_\nu)_{\nu\in\Lambda} \longleftarrow \mathbf{0}$

for $p = 1, \dots, \ell$ **do**

  forall $\mu \in \Lambda_p$ **do**

    for $m = 1, \dots, M_p$ **do**

      $w \longleftarrow R_m v_\mu$

      $z_{\mu+\epsilon_m} \longleftarrow z_{\mu+\epsilon_m} + \pi^+_{\mu_m} w$

      if $\mu_m \ge 1$ **then** $z_{\mu-\epsilon_m} \longleftarrow z_{\mu-\epsilon_m} + \pi^-_{\mu_m} w$

By [GHS07, Rem. 2.3] or [DSS09, Prop. 4.4], the number of operations and storage locations required by a call of $\texttt{BucketSort}[w, \epsilon]$ is bounded by

$$\#\operatorname{supp} w + \max(1, \lceil \log(\|w\|_{\ell^\infty(\Lambda)} (\#\operatorname{supp} w)/\epsilon) \rceil) \,, \tag{3.9}$$

which is faster than exact comparison-based sorting algorithms.

Let $v = (v_\mu)_{\mu\in\Lambda}$ be a finitely supported sequence in V, indexed by Λ. Such a vector represents a polynomial $v \in \mathcal{P}(\Gamma; V)$ by

$$v(y) = \sum_{\mu\in\operatorname{supp} v} v_\mu P_\mu(y) \,, \quad y \in \Gamma \,, \tag{3.10}$$

where $\operatorname{supp} v = \{\mu \in \Lambda \,;\, v_\mu \neq 0\}$ is a finite subset of Λ by assumption. Due to the normalization (2.7),

$$\|v\|_{C(\Gamma;V)} = \max_{y\in\Gamma} \left\|v(y)\right\|_V \le \sum_{\mu\in\Lambda} \left\|v_\mu\right\|_V = \|v\|_{\ell^1(\Lambda;V)} \,. \tag{3.11}$$

The routine $\texttt{Apply}_{\mathcal{R}}[v, \epsilon]$ adaptively approximates $\mathcal{R}v$ for $v \in \mathcal{P}(\Gamma; V)$ in three distinct steps. First, the elements of the coefficient vector v of v are grouped according to their norm. Elements smaller than a certain tolerance are discarded. This truncation of the vector v induces an error of at most $\delta \le \epsilon/2$.

Next, a greedy algorithm is used to assign to each segment $v_{[p]} = v|_{\Lambda_p}$ of v an approximation $\mathcal{R}_{[M_p]}$ of $\mathcal{R}$. Starting with $\mathcal{R}_{[M_p]} = 0$ for all $p = 1, \dots, \ell$, these approximations are refined iteratively until an estimate of the error is smaller than $\epsilon - \delta$.

Finally, the operations determined by the previous two steps are performed. Each multiplication $R_m v_\mu$ is performed just once, and copied to the appropriate entries of $\boldsymbol{z}$. Then the polynomial

$$z(y) \coloneqq \sum_{\mu \in \operatorname{supp} \boldsymbol{z}} z_\mu P_\mu(y) \,, \quad y \in \Gamma \,, \tag{3.12}$$

is an approximation of $\mathcal{R}v$ with error at most ϵ.

Proposition 3.2. *For any $\epsilon > 0$ and any $v \in \mathcal{P}(\Gamma; V)$ with coefficient vector $\boldsymbol{v}$ as in* (3.10)*, $\texttt{Apply}_{\mathcal{R}}[v, \epsilon]$ produces a finitely supported $\boldsymbol{z} \in \ell^1(\Lambda; W^*)$ such that*

$$\#\operatorname{supp} \boldsymbol{z} \le 2 \sum_{p=1}^{\ell} M_p \# \Lambda_p \tag{3.13}$$

and the polynomial $z \in \mathcal{P}(\Gamma; W^*)$ from (3.12) *satisfies*

$$\|\mathcal{R}v - z\|_{C(\Gamma; W^*)} \le \delta + \sum_{p=1}^{\ell} \bar{e}_{\mathcal{R}, M_p} \left\| \boldsymbol{v}_{[p]} \right\|_{\ell^1(\Lambda; V)} \,, \tag{3.14}$$

where M_p refers to the final value of this variable in the call of $\texttt{Apply}_{\mathcal{R}}$. The total number of products $R_m v_\mu$ computed in $\texttt{Apply}_{\mathcal{R}}[v, \epsilon]$ is $\sum_{p=1}^{\ell} M_p \# \Lambda_p$.

Proof. For each $\mu \in \Lambda_p$, $R_m v_\mu$ is computed for $m = 1, \dots, M_p$, and passed on to at most two coefficients of $\boldsymbol{z}$. This shows (3.13) and the bound on the number of multiplications. Since $\|\mathcal{R}\|_{C(\Gamma; V) \to C(\Gamma; W^*)} \le \bar{e}_{\mathcal{R}, 0}$, using (3.11),

$$\|\mathcal{R}v - \mathcal{R}w\|_{C(\Gamma; W^*)} \le \bar{e}_{\mathcal{R}, 0} \|v - w\|_{C(\Gamma; W)} \le \bar{e}_{\mathcal{R}, 0} \|\boldsymbol{v} - \boldsymbol{w}\|_{\ell^1(\Lambda; V)} = \delta \le \frac{\epsilon}{2} \,,$$

where $\boldsymbol{w} \coloneqq \sum_{p=1}^{\ell} \boldsymbol{v}_{[p]}$ and w is the polynomial (3.10) with coefficients $\boldsymbol{w}$. For all $p = 1, \dots, \ell$, let $v_{[p]} \in \mathcal{P}(\Gamma; V)$ denote the polynomial with coefficients $\boldsymbol{v}_{[p]}$. Due to (3.1) and the termination criterion of the greedy subroutine in $\texttt{Apply}_{\mathcal{R}}$,

$$\sum_{p=1}^{\ell} \left\| \mathcal{R} v_{[p]} - \mathcal{R}_{[M_p]} v_{[p]} \right\|_{C(\Gamma; W^*)} \le \sum_{p=1}^{\ell} \bar{e}_{\mathcal{R}, M_p} \left\| v_{[p]} \right\|_{C(\Gamma; V)} \le \sum_{p=1}^{\ell} \bar{e}_{\mathcal{R}, M_p} \left\| \boldsymbol{v}_{[p]} \right\|_{\ell^1(\Lambda; V)} \le \epsilon - \delta \,.$$

The assertion follows since $z = \sum_{p=1}^{\ell} \mathcal{R}_{[M_p]} v_{[p]}$. □

Remark 3.3. By Proposition 3.2, the cost of $\texttt{Apply}_{\mathcal{R}}$ is described by $\sum_{p=1}^{\ell} M_p \# \Lambda_p$, and up to the term δ from the truncation of $\boldsymbol{v}$, the error is bounded by

$$\sum_{p=1}^{\ell} \bar{e}_{\mathcal{R}, M_p} \left\| \boldsymbol{v}_{[p]} \right\|_{\ell^1(\Lambda; V)} \,. \tag{3.15}$$

Due to the assumption that $(\bar{e}_{\mathcal{R}, M} - \bar{e}_{\mathcal{R}, M+1})_{M=0}^{\infty}$ is nonincreasing, the greedy algorithm used in $\texttt{Apply}_{\mathcal{R}}$ to determine M_p is guaranteed to minimize $\sum_{p=1}^{\ell} M_p \# \Lambda_p$ under the condition that (3.15) is at most $\epsilon - \delta$. ⌟

3.2 Formulation of the Method

The adaptive application routine from Section 3.1 efficiently realizes the approximate application routine of the operator $\mathcal{R}$, which is a crucial component of the perturbed linear iteration from Section 1.2. We assume that polynomial approximations of the right hand side $f \in C(\Gamma; W^*)$ in (1.17) are available with arbitrary precision. By Theorem 2.2, such approximations are guaranteed to exist if W^* has the approximation property. We assume that a routine

$$\mathtt{RHS}_f[\epsilon] \mapsto \tilde{\boldsymbol{f}} \tag{3.16}$$

is available which, for any $\epsilon > 0$, returns a finitely supported $\tilde{\boldsymbol{f}} = (\tilde{f}_\nu)_{\nu\in\Lambda} \in \ell^1(\Lambda; W^*)$ with

$$\left\|f - \tilde{f}\right\|_{C(\Gamma;W^*)} \le \epsilon \qquad \text{for} \qquad \tilde{f}(y) \coloneqq \sum_{\nu\in\Lambda} \tilde{f}_\nu P_\nu(y)\,, \quad y \in \Gamma\,. \tag{3.17}$$

Of course, $\mathtt{RHS}_f$ is trivial if f does not depend on $y \in \Gamma$.

Furthermore, let $\mathtt{Solve}_D$ be a solver for D such that for any $g \in W^*$ and any $\epsilon > 0$,

$$\mathtt{Solve}_D[g, \epsilon] \mapsto v\,, \quad \left\|v - D^{-1}g\right\|_V \le \epsilon\,. \tag{3.18}$$

For example, $\mathtt{Solve}_D$ could be an adaptive wavelet method, see *e.g.* [CDD01, CDD02, GHS07], an adaptive frame method, see *e.g.* [Ste03, DFR07, DRW$^+$07], or a finite element method with a posteriori error estimation, see *e.g.* [Dör96, MNS00, BDD04].

A realization of the iteration from Section 1.2 using the above approximations is given in $\mathtt{SolveDirect}_{\mathcal{A},f}$. We write $\tilde{u}^{(k)}$, u_ϵ and $g^{(k)}$ for the polynomials with coefficients $\tilde{\boldsymbol{u}}^{(k)}$, $\boldsymbol{u}_\epsilon$ and $\boldsymbol{g}^{(k)}$, respectively. The initial values can be set to

$$\tilde{\boldsymbol{u}}^{(0)} \coloneqq \boldsymbol{0} \quad \text{and} \quad \delta_0 \coloneqq (1-\gamma)^{-1} \left\|D^{-1}\right\|_{W^*\to V} \left\|f\right\|_{C(\Gamma;W^*)}\,. \tag{3.19}$$

Note that δ_0 is an upper bound for the initial error $\left\|u - \tilde{u}^{(0)}\right\|_{C(\Gamma;V)}$. Values for the other arguments are given below.

Theorem 3.4. *For any $\epsilon > 0$ and any $\tilde{u}^{(0)} \in \mathcal{P}(\Gamma; V)$, if $\left\|u - \tilde{u}^{(0)}\right\|_{C(\Gamma;V)} \le \delta_0$, $\alpha > 0$, $\beta_0, \beta_1 > 0$ and $\alpha + \beta_0 + \beta_1 + \gamma < 1$, then $\mathtt{SolveDirect}_{\mathcal{A},f}[\tilde{\boldsymbol{u}}^{(0)}, \delta_0, \epsilon, \alpha, \beta_0, \beta_1, \gamma]$ terminates with*

$$\|u - u_\epsilon\|_{C(\Gamma;V)} \le \bar{\epsilon} \le \epsilon\,. \tag{3.20}$$

Furthermore, for all $k \in \mathbb{N}$ reached in the iteration,

$$\left\|u - \tilde{u}^{(k)}\right\|_{C(\Gamma;V)} \le \min(\delta_k, \bar{\delta}_k) \le (\alpha + \beta_0 + \beta_1 + \gamma)^k \delta_0\,. \tag{3.21}$$

Proof. We show that for all $k \in \mathbb{N}$,

$$\left\|u - \tilde{u}^{(k)}\right\|_{C(\Gamma;V)} \le \min(\delta_k, \bar{\delta}_k)\,.$$

Let $\left\|u - \tilde{u}^{(k-1)}\right\|_{C(\Gamma;V)} \le \min(\delta_{k-1}, \bar{\delta}_{k-1})$. Then as in the proof of Theorem 1.5, since $\mathcal{D}u = f - \mathcal{R}u$,

$$u - \tilde{u}^{(k)} = \mathcal{D}^{-1}(f - \mathcal{R}u) - \mathcal{D}^{-1}(f - \mathcal{R}\tilde{u}^{(k-1)}) + \mathcal{D}^{-1}(f - \mathcal{R}\tilde{u}^{(k-1)} - g^{(k)}) + \mathcal{D}^{-1}g^{(k)} - \tilde{u}^{(k)}\,.$$

$\texttt{SolveDirect}_{\mathcal{A},f}[\tilde{\boldsymbol{u}}^{(0)}, \delta_0, \epsilon, \alpha, \beta_0, \beta_1, \gamma] \mapsto [\boldsymbol{u}_\epsilon, \bar{\epsilon}]$

for $k = 1, 2, \dots$ **do**

$\eta_k \longleftarrow \delta_{k-1} \left\| D^{-1} \right\|^{-1}_{W^* \to V}$

$\boldsymbol{g}^{(k)} = (g_\mu^{(k)})_{\mu \in \Lambda} \longleftarrow \texttt{RHS}_f[\beta_0 \eta_k] - \texttt{Apply}_{\mathcal{R}}[\tilde{\boldsymbol{u}}^{(k-1)}, \beta_1 \eta_k]$

$\zeta_k \longleftarrow \alpha \delta_{k-1} (\#\operatorname{supp} \boldsymbol{g}^{(k)})^{-1}$

forall $\mu \in \operatorname{supp} \boldsymbol{g}^{(k)}$ **do** $\tilde{u}_\mu^{(k)} \longleftarrow \texttt{Solve}_D[g_\mu^{(k)}, \zeta_k]$

$\tilde{\boldsymbol{u}}^{(k)} \longleftarrow (\tilde{u}_\mu^{(k)})_{\mu \in \Lambda}$

$\bar{\delta}_{k-1} \longleftarrow (1-\gamma)^{-1} \left(\left\| \tilde{\boldsymbol{u}}^{(k)} - \tilde{\boldsymbol{u}}^{(k-1)} \right\|_{\ell^1(\Lambda; V)} + (\alpha + \beta_0 + \beta_1) \delta_{k-1} \right)$

$\delta_k \longleftarrow (\alpha + \beta_0 + \beta_1) \delta_{k-1} + \gamma \min(\delta_{k-1}, \bar{\delta}_{k-1})$

if $\delta_k \leq \epsilon$ **then break**

$\boldsymbol{u}_\epsilon \longleftarrow \tilde{\boldsymbol{u}}^{(k)}$

$\bar{\epsilon} \longleftarrow \delta_k$

Due to (1.15),

$$\left\| \mathcal{D}^{-1}(f - \mathcal{R}u) - \mathcal{D}^{-1}(f - \mathcal{R}\tilde{u}^{(k-1)}) \right\|_{C(\Gamma; V)} = \left\| \mathcal{D}^{-1} \mathcal{R}(u - \tilde{u}^{(k-1)}) \right\|_{C(\Gamma; V)} \leq \gamma \min(\delta_{k-1}, \bar{\delta}_{k-1}) \, .$$

Furthermore, using $\left\| \mathcal{D}^{-1} \right\|_{C(\Gamma; W^*) \to C(\Gamma; V)} = \left\| D^{-1} \right\|_{W^* \to V}$, Proposition 3.2 and (3.17),

$$\left\| \mathcal{D}^{-1}(f - \mathcal{R}\tilde{u}^{(k-1)} - g^{(k)}) \right\|_{C(\Gamma; V)} \leq \left\| D^{-1} \right\|_{W^* \to V} (\beta_0 \eta_k + \beta_1 \eta_k) = (\beta_0 + \beta_1) \delta_{k-1} \, .$$

Finally, due to (3.11), (3.18) and $\zeta_k = \alpha \delta_{k-1} (\#\operatorname{supp} \boldsymbol{g}^{(k)})^{-1}$,

$$\left\| \mathcal{D}^{-1} g^{(k)} - \tilde{u}^{(k)} \right\|_{C(\Gamma; V)} \leq \sum_{\mu \in \operatorname{supp} \boldsymbol{g}^{(k)}} \alpha \delta_{k-1} (\#\operatorname{supp} \boldsymbol{g}^{(k)})^{-1} = \alpha \delta_{k-1} \, .$$

By triangle inequality, these estimates imply

$$\left\| u - \tilde{u}^{(k)} \right\|_{C(\Gamma; V)} \leq \gamma \min(\delta_{k-1}, \bar{\delta}_{k-1}) + (\beta_0 + \beta_1) \delta_{k-1} + \alpha \delta_{k-1} = \delta_k \, .$$

The residual at iteration k is $r^{(k)} := f - \mathcal{A}\tilde{u}^{(k)} = A(u - \tilde{u}^{(k)})$. We observe that by (1.16),

$$\left\| u - \tilde{u}^{(k)} \right\|_{C(\Gamma; V)} \leq \left\| (\mathcal{D}^{-1}\mathcal{A})^{-1} \right\| \left\| \mathcal{D}^{-1} r^{(k)} \right\|_{C(\Gamma; V)} \leq \frac{1}{1-\gamma} \left\| \mathcal{D}^{-1} r^{(k)} \right\|_{C(\Gamma; V)} \, .$$

Furthermore, $\mathcal{D}^{-1} r^{(k)}$ can be approximated by known quantities since, similarly to above,

$$\left\| \tilde{u}^{(k+1)} - \tilde{u}^{(k)} - \mathcal{D}^{-1} r^{(k)} \right\|_{C(\Gamma; V)} = \left\| \tilde{u}^{(k+1)} - \mathcal{D}^{-1}(f - \mathcal{R}\tilde{u}^{(k)}) \right\|_{C(\Gamma; V)} \leq (\alpha + \beta_0 + \beta_1) \delta_k \, .$$

Consequently, using (3.11),

$$\left\| u - \tilde{u}^{(k)} \right\|_{C(\Gamma; V)} \leq \frac{1}{1-\gamma} \left(\left\| \tilde{u}^{(k+1)} - \tilde{u}^{(k)} \right\|_{C(\Gamma; V)} + (\alpha + \beta_0 + \beta_1) \delta_k \right) \leq \bar{\delta}_k \, .$$

Equation (3.21) follows since $\delta_k \leq (\alpha + \beta_0 + \beta_1 + \gamma) \delta_{k-1}$. □

Remark 3.5. The error bounds in $\texttt{SolveDirect}_{\mathcal{A},f}$ can be improved if each of the subroutines $\texttt{RHS}_f$, $\texttt{Apply}_{\mathcal{R}}$ and $\texttt{Solve}_D$ returns an estimate of the error it attains. These values can replace $\alpha\delta_{k-1}$, $\beta_0\delta_{k-1}$ and $\beta_1\delta_{k-1}$ in the definitions of $\bar{\delta}_{k-1}$ and δ_k. For better legibility, we refrain from making this explicit. ⌟

4 Alternating Subspace Correction

4.1 Motivation

By (2.13) and (2.15), if $v \in C(\Gamma; V)$ is an even function, *i.e.* $v(-y) = v(y)$, then $\mathcal{R}v$ and $\mathcal{R}_{[M]}v$ are odd functions, *i.e.* $v(-y) = -v(y)$. Similarly, if v is odd, then $\mathcal{R}v$ and $\mathcal{R}_{[M]}v$ are even. Since $\mathcal{D}$ does not depend on y, $\mathcal{D}v$ is even if v is even, and odd if v is odd.

Let $[n] \coloneqq n + 2\mathbb{Z}$ denote the equivalence class modulo two of $n \in \mathbb{Z}$, *i.e.* $[n] = [m]$ if $n - m$ is even.

The right hand side $f \in C(\Gamma; W^*)$ of (1.17) can be divided into even and odd parts as $f = f^{[0]} + f^{[1]}$ for

$$f^{[0]}(y) \coloneqq \frac{1}{2}\Big(f(y) + f(-y)\Big) \qquad \text{and} \qquad f^{[1]}(y) \coloneqq \frac{1}{2}\Big(f(y) - f(-y)\Big)\,. \tag{4.1}$$

Then the iteration (1.18) is equivalent to $u_k = u_k^{[k]} + u_k^{[k-1]}$ for

$$u_k^{[k]} \coloneqq \mathcal{D}^{-1}\Big(f^{[k]} - \mathcal{R}u_{k-1}^{[k-1]}\Big) \qquad \text{and} \qquad u_k^{[k-1]} \coloneqq \mathcal{D}^{-1}\Big(f^{[k-1]} - \mathcal{R}u_{k-1}^{[k]}\Big)\,. \tag{4.2}$$

We note that $u_k^{[k]}$ only depends on $u_{k-1}^{[k-1]}$ and $u_k^{[k-1]}$ only depends on $u_{k-1}^{[k]}$. We can therefore perform just one of these iterations, say

$$u_k^{[k]} \coloneqq \mathcal{D}^{-1}\Big(f^{[k]} - \mathcal{R}u_{k-1}^{[k-1]}\Big)\,, \tag{4.3}$$

and reconstruct an approximation of u as $u_k^{[k]} + u_{k-1}^{[k-1]}$.

For polynomials $v \in \mathcal{P}(\Gamma; V)$, the separation into even and odd parts carries over to the coefficients of v in V. We define the index sets

$$\Lambda^{[n]} \coloneqq \Big\{\mu \in \Lambda\,;\, [\big|\mu\big|] = [n]\Big\}\,, \quad n \in \mathbb{Z}\,, \tag{4.4}$$

where $\big|\mu\big| = \big\|\mu\big\|_{\ell^1(\mathbb{N})}$. Then

$$\Lambda = \Lambda^{[0]} \sqcup \Lambda^{[1]}\,. \tag{4.5}$$

We call $\mu \in \Lambda$ even if $\mu \in \Lambda^{[0]}$ and odd if $\mu \in \Lambda^{[1]}$.

Remark 4.1. A finitely supported sequence $(v_\mu)_{\mu\in\Lambda}$ defines a polynomial function

$$v(y) = \sum_{\mu\in\Lambda} v_\mu P_\mu(y)\,, \quad y \in \Gamma\,. \tag{4.6}$$

The function v is even if and only if $v_\mu = 0$ for all $\mu \in \Lambda^{[1]}$ and odd if and only if $v_\mu = 0$ for all $\mu \in \Lambda^{[0]}$ since P_μ is even for $\mu \in \Lambda^{[0]}$ and odd for $\mu \in \Lambda^{[1]}$, and the representation (4.6) is unique. ⌟

4.2 Formulation of the Method

We assume that routines $\texttt{RHS}_f^{[0]}$ and $\texttt{RHS}_f^{[1]}$ are available similar to $\texttt{RHS}_f$ from (3.16) to construct approximations of $f^{[0]}$ and $f^{[1]}$ from (4.1), such that the approximations of $f^{[0]}$ are even and those of $f^{[1]}$ are odd.

The method $\texttt{Apply}_{\boldsymbol{R}}$ from Section 3.1 already respects even and odd functions in the sense that if v in $z := \texttt{Apply}_{\mathcal{R}}[v, \epsilon]$ is supported in $\Lambda^{[n]}$, then z is supported in $\Lambda^{[n+1]}$.

$\texttt{SolveAlternate}_{\mathcal{A},f}[\tilde{\boldsymbol{u}}^{(-1)}, \delta_{-1}, \epsilon, \alpha, \beta_0, \beta_1, \gamma] \mapsto [\boldsymbol{u}_\epsilon, \bar{\epsilon}]$

---

for $k = 0, 1, 2, \dots$ **do**
- $\eta_k \longleftarrow \delta_{k-1} \left\| D^{-1} \right\|_{W^* \to V}^{-1}$
- $\boldsymbol{g}^{(k)} = (g_\mu^{(k)})_{\mu \in \Lambda^{[k]}} \longleftarrow \texttt{RHS}_f^{[k]}[\beta_0 \eta_k] - \texttt{Apply}_{\mathcal{R}}[\tilde{\boldsymbol{u}}^{(k-1)}, \beta_1 \eta_k]$
- $\zeta_k \longleftarrow \alpha \delta_{k-1} (\#\operatorname{supp} \boldsymbol{g}^{(k)})^{-1}$
- **forall** $\mu \in \operatorname{supp} \boldsymbol{g}^{(k)}$ **do** $\tilde{u}_\mu^{(k)} \longleftarrow \texttt{Solve}_D[g_\mu^{(k)}, \zeta_k]$
- $\tilde{\boldsymbol{u}}^{(k)} \longleftarrow (\tilde{u}_\mu^{(k)})_{\mu \in \Lambda^{[k]}}$
- $\delta_k \longleftarrow (\alpha + \beta_0 + \beta_1 + \gamma) \delta_{k-1}$
- **if** $\delta_{k-1} + \delta_k \le \epsilon$ **then break**

$\boldsymbol{u}_\epsilon \longleftarrow \tilde{\boldsymbol{u}}^{(k-1)} + \tilde{\boldsymbol{u}}^{(k)}$

$\bar{\epsilon} \longleftarrow \delta_{k-1} + \delta_k$

As in (4.1), let

$$u^{[k]}(y) := \frac{1}{2}\Big(u(y) + (-1)^k u(-y)\Big)\,, \tag{4.7}$$

such that $u = u^{[k-1]} + u^{[k]}$ for any $k \in \mathbb{Z}$.

Theorem 4.2. *For any $\epsilon > 0$ and any finitely supported $\tilde{\boldsymbol{u}}^{(-1)} \in \ell^1(\Lambda^{[-1]}; V)$, if $\alpha + \beta_0 + \beta_1 + \gamma < 1$ and $\left\| u^{[-1]} - \tilde{u}^{(-1)} \right\|_{C(\Gamma;V)} \le \delta_{-1}$, then $\texttt{SolveAlternate}_{\mathcal{A},f}[\tilde{\boldsymbol{u}}^{(-1)}, \delta_{-1}, \epsilon, \alpha, \beta_0, \beta_1, \gamma]$ terminates with*

$$\|u - u_\epsilon\|_{C(\Gamma;V)} \le \bar{\epsilon} \le \epsilon\,. \tag{4.8}$$

Furthermore, for all $k \in \mathbb{N}$ reached in the iteration,

$$\left\| u^{[k]} - \tilde{u}^{(k)} \right\|_{C(\Gamma;V)} \le \delta_k \le (\alpha + \beta_0 + \beta_1 + \gamma)^{k+1} \delta_{-1}\,. \tag{4.9}$$

Proof. Since $\mathcal{D} u^{[k]} = f^{[k]} - \mathcal{R} u^{[k-1]}$,

$$\begin{aligned} u^{[k]} - \tilde{u}^{(k)} = {} & \mathcal{D}^{-1}(f^{[k]} - \mathcal{R} u^{[k-1]}) - \mathcal{D}^{-1}(f^{[k]} - \mathcal{R}\tilde{u}^{(k-1)}) \\ & + \mathcal{D}^{-1}(f^{[k]} - \mathcal{R}\tilde{u}^{(k-1)} - g^{(k)}) + \mathcal{D}^{-1} g^{(k)} - \tilde{u}^{(k)}\,. \end{aligned}$$

Due to (1.15),

$$\left\| \mathcal{D}^{-1}(f^{[k]} - \mathcal{R} u^{[k-1]}) - \mathcal{D}^{-1}(f^{[k]} - \mathcal{R}\tilde{u}^{(k-1)}) \right\|_{C(\Gamma;V)} \le \gamma \left\| u^{[k-1]} - \tilde{u}^{(k-1)} \right\|_{C(\Gamma;V)}\,.$$

By definition of $g^{(k)}$, using $\|\mathcal{D}^{-1}\|_{C(\Gamma;W^*)\to C(\Gamma;V)} = \|D^{-1}\|_{W^*\to V'}$

$$\|\mathcal{D}^{-1}(f^{[k]} - \mathcal{R}\tilde{u}^{(k-1)} - g^{(k)})\|_{C(\Gamma;V)} \le \|D^{-1}\|_{W^*\to V} (\beta_0\eta_k + \beta_1\eta_k) = (\beta_0+\beta_1)\delta_{k-1} .$$

Also, by (3.18),

$$\|\mathcal{D}^{-1}g^{(k)} - \tilde{u}^{(k)}\|_{C(\Gamma;V)} \le \|(D^{-1}g_\mu^{(k)}) - \tilde{\boldsymbol{u}}^{(k)}\|_{\ell^1(\Lambda;V)} \le \sum_{\mu\in\operatorname{supp}\boldsymbol{g}^{(k)}} \alpha\delta_{k-1}(\#\operatorname{supp}\boldsymbol{g}^{(k)})^{-1} = \alpha\delta_{k-1} .$$

Combining these estimates leads to

$$\|u^{[k]} - \tilde{u}^{(k)}\|_{C(\Gamma;V)} \le (\alpha+\beta_0+\beta_1)\delta_{k-1} + \gamma\,\|u^{[k-1]} - \tilde{u}^{(k-1)}\|_{C(\Gamma;V)} .$$

Consequently, if $\|u^{[k-1]} - \tilde{u}^{(k-1)}\|_{C(\Gamma;V)} \le \delta_{k-1}$, then $\|u^{[k]} - \tilde{u}^{(k)}\|_{C(\Gamma;V)} \le \delta_k$, and (4.9) follows by induction. □

Remark 4.3. As in Remark 1.6 the error bounds δ_k can be refined using an approximation of the residual analogously to $\texttt{SolveDirect}_{\mathcal{A},f}$. As in the proof of Theorem 3.4, it follows that

$$\|u - (\tilde{u}^{(k-1)} + \tilde{u}^{(k)})\|_{C(\Gamma;V)} \le \frac{1}{1-\gamma}\Big(\|\tilde{\boldsymbol{u}}^{(k+1)} - \tilde{\boldsymbol{u}}^{(k-1)}\|_{\ell^1(\Lambda;V)} + (\alpha+\beta_0+\beta_1)(\delta_{k-1}+\delta_k)\Big) . \quad (4.10)$$

This term can be used as an alternative upper bound for the errors $\|u^{[k]} - \tilde{u}^{(k)}\|_{C(\Gamma;V)}$ and $\|u^{[k-1]} - \tilde{u}^{(k-1)}\|_{C(\Gamma;V)}$. However, since it applies to the total error instead of directly to the even or odd part, we expect it to be less useful than the bound $\bar{\delta}_k$ in $\texttt{SolveDirect}_{\mathcal{A},f}$. ⌟

Remark 4.4. Comparing the convergence estimates (3.21) and (4.9), it appears that $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$ converge at the same rate. Therefore, since the latter method updates only half of the solution vector in each iteration, it should be roughly twice as efficient. However, Remark 4.3 suggests that $\texttt{SolveDirect}_{\mathcal{A},f}$ may provide a sharper bound for the error. It is not clear a priori which of these effects is more significant; numerical computations presented in Section 5.2 indicate that the two solvers are equally efficient. ⌟

5 Numerical Computations

5.1 A Model Problem

We consider as a model problem the diffusion equation (1.27) on the one dimensional domain $G = (0,1)$. For two parameters k and γ, the diffusion coefficient has the form

$$a(y,x) = 1 + \frac{1}{c}\sum_{m=1}^{\infty} y_m \frac{1}{m^k}\sin(m\pi x) , \quad x\in(0,1) , \quad y\in\Gamma = [-1,1]^\infty , \quad (5.1)$$

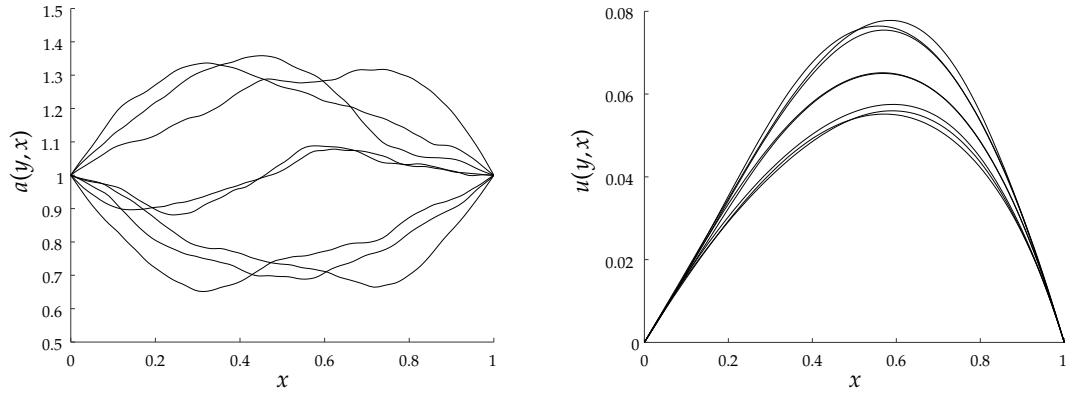

Figure 1: Realizations of $a(y,x)$ (left) and $u(y,x)$ (right).

where c is chosen as

$$c = \gamma \sum_{m=1}^{\infty} \frac{1}{m^k}\,, \tag{5.2}$$

such that $\left|a(y,x) - 1\right|$ is always less than γ. We set the parameters to $k = 2$ and $\gamma = 1/2$. A few realizations of $a(y)$ and the resulting solutions $u(y)$ of (1.27) are plotted in Figure 1.

On the parameter domain, we consider Chebyshev polynomials of the first kind and Legendre polynomials. We use a multilevel finite element discretization with piecewise linear basis functions on uniform meshes. The residual-based a posteriori error estimator from [Git11b] is used to estimate the error in $\texttt{Solve}_D$. In order to isolate the discretization of the parameter domain, we also consider a fixed spatial discretization, using linear finite elements on a uniform mesh of $(0,1)$ with 1024 elements to approximate all coefficients. We refer to these simpler versions of the numerical methods as single level discretizations. All computations were performed in Matlab on a workstation with an AMD Athlon™ 64 X2 5200+ processor and 4GB of memory.

5.2 Convergence of Solvers with Uniform Error Control

The convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$ is plotted in Figure 2. We use Chebyshev polynomials on the parameter domain Γ, and the parameters of both methods are set to $\alpha = 1/20$, $\beta_0 = 0$ and $\beta_1 = 1/10$.

The solid lines in Figure 2 refer to the error estimate δ_k, which is an upper bound for the error in $C(\Gamma; V)$. We use Remark 4.3 to take advantage of information on the residual when determining this upper bound for $\texttt{SolveAlternate}_{\mathcal{A},f}$. The original formulation of $\texttt{SolveDirect}_{\mathcal{A},f}$ already makes use of this.

The dashed lines represent the maximal error on a sample set Γ_S of 64 points y in Γ, which is a lower bound for the actual error. The sample set Γ_S is chosen as a subset of the boundary of Γ. Each $y \in \Gamma_S$ consists of randomly chosen $y_m \in \{-1, 0, 1\}$ for $m \leq 250$ and $y_m = 0$ for $m > 250$. Realizations of $a(y)$ and $u(y)$ for the first eight points $y \in \Gamma_S$ are

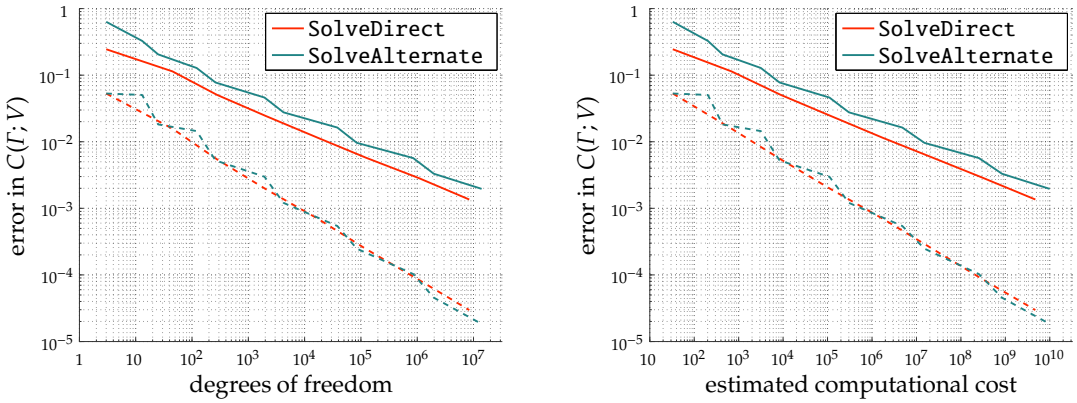

Figure 2: Convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$. Solid lines refer to the error estimate δ_k, dashed lines are maximal errors on Γ_S.

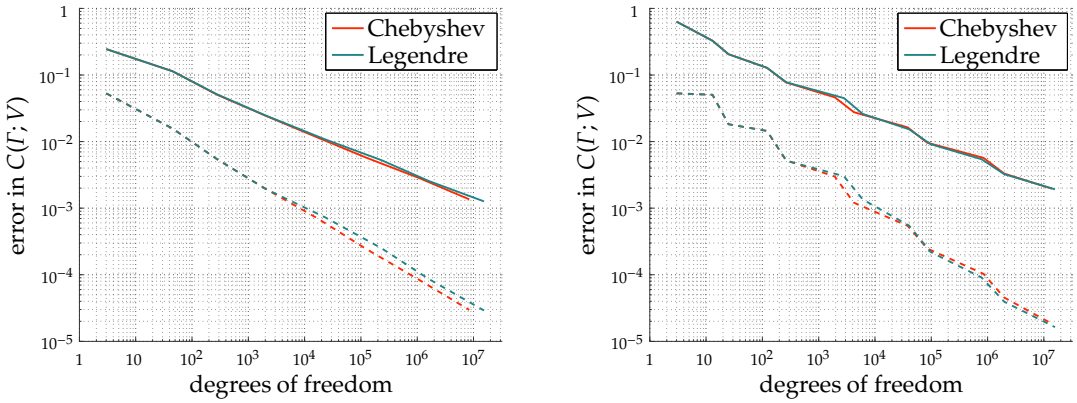

Figure 3: Convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ (left) and $\texttt{SolveAlternate}_{\mathcal{A},f}$ (right) with Chebyshev and Legendre polynomial bases. Solid lines refer to the error estimate δ_k, dashed lines are maximal errors on Γ_S.

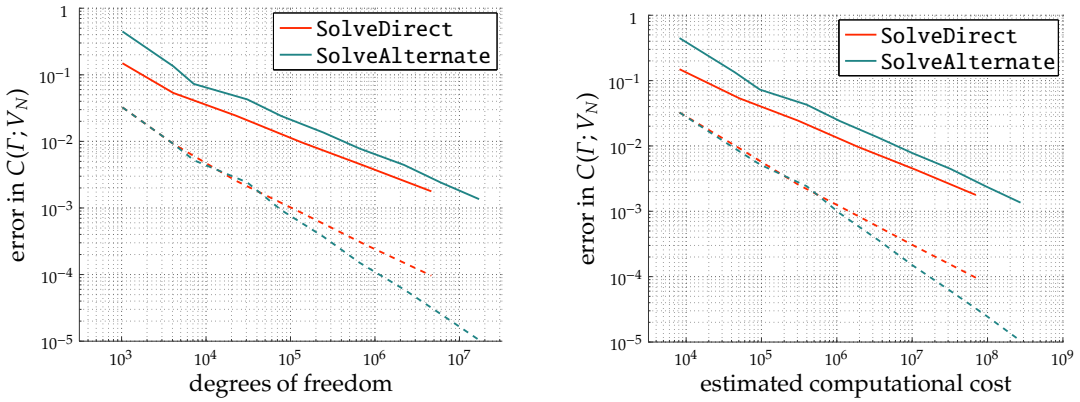

Figure 4: Convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$ with a fixed finite element discretization. Solid lines refer to the error estimate δ_k, dashed lines are maximal errors on Γ_S, compared with a finite element solution on the same level.

given in Figure 1. We assume that the maximal error on Γ_S is a good approximation of the maximal error on all of Γ.

On the left, the errors are plotted against the total number of basis functions used in the discretization. On the right, we plot the errors against an estimate of the computational cost. This estimate takes scalar products, matrix-vector multiplications and linear solves into account. The total number of each of these operations on each discretization level is tabulated during the computation, weighted by the number of degrees of freedom on the discretization level, and summed over all levels. The estimate is equal to seven times the resulting sum for linear solves, plus three times the value for matrix-vector multiplications, plus the sum for scalar products. These weights were determined empirically by timing the operations for tridiagonal sparse matrices in Matlab.

We can see in Figure 2 that δ_k is a coarse upper bound for the actual error. Furthermore, the convergence rate of δ_k in this example is only 1/3, compared to a rate of 1/2 for the estimate of the error. This leads to an over-estimation of the error by almost two orders of magnitude towards the end of the computation. This effect is not visible in [Git11a], where a larger k is used in the model problem, *i.e.* the series in (5.1) converges faster.

As anticipated, the error bound for $\texttt{SolveAlternate}_{\mathcal{A},f}$ is slightly coarser than that of $\texttt{SolveDirect}_{\mathcal{A},f}$. However, the convergence of the two methods is very similar.

Figure 3 compares the Chebyshev basis used in Figure 2 to the Legendre basis. There does not seem to be much of a difference between these two choices.

Figure 4 shows the convergence of the single level variants of $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$ that use a fixed finite element discretization. The spatial discretization error of approximately $2 \cdot 10^{-4}$ is suppressed in the convergence plot. The dashed lines refer to the maximal difference between the adaptively computed parametric solutions on the sample set, and the Galerkin projections computed individually for each point $y \in \Gamma_S$ in this set. The solid lines represent the error bounds δ_k, which are only an

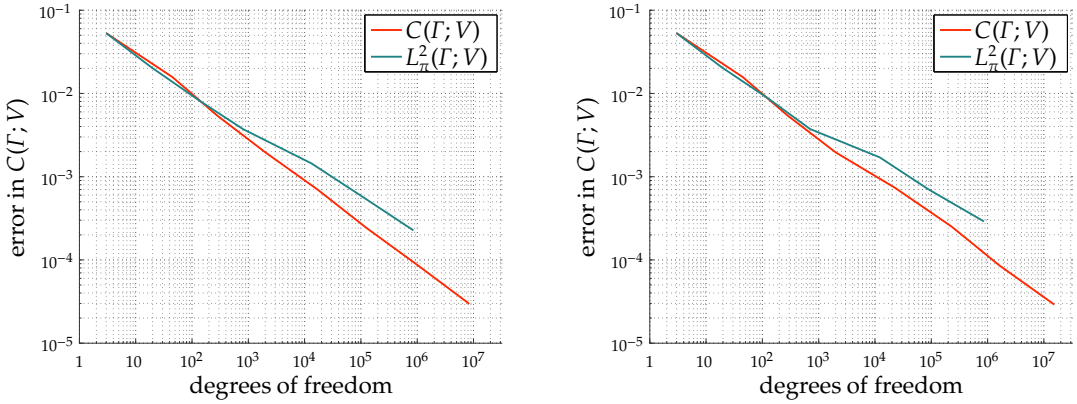

Figure 5: Convergence in $C(\Gamma;V)$ of $\texttt{SolveDirect}_{\mathcal{A},f}$ with Chebyshev polynomials (left) and Legendre polynomials (right). The two versions of $\texttt{SolveDirect}_{\mathcal{A},f}$ control the error in $C(\Gamma;V)$ and $L^2_\pi(\Gamma;V)$, respectively.

upper bound for the parametric error for single level methods, and do not capture the spatial discretization error.

The single level solvers in Figure 4 simulate $\texttt{SolveDirect}_{\mathcal{A},f}$ and $\texttt{SolveAlternate}_{\mathcal{A},f}$ with no spatial discretization. In this setting, a theoretical asymptotic approximation rate of 1 is shown in [CDS10b, CDS10a]. We observe a rate of approximately 1/2 for our adaptive solvers, although the convergence rate of the error of $\texttt{SolveAlternate}_{\mathcal{A},f}$ on the sample set approaches one.

For the fully discrete system, *i.e.* Figures 2 and 3, the assumptions of the approximation results in [CDS10b, CDS10a] are not satisfied, and thus no convergence is shown there. However, we still observe the rates 1/3 for the error bounds δ_k and 1/2 for the actual error, as mentioned above.

5.3 Comparison to Other Adaptive Methods

We compare the convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ to that of similar methods from [Git11b, Git11c], which control the error in $L^2_\pi(\Gamma;V)$ for a probability measure π on the parameter domain Γ. In the following, this probability measure is always chosen in such a way that the polynomial basis $(P_\mu)_{\mu\in\Lambda}$ is orthogonal. For example, for Legendre polynomials, π is a countable product of uniform distributions on $[-1,1]$.

Since the solvers from [Git11b, Git11c] do not provide bounds for the error in $C(\Gamma;V)$, we do not consider computationally accessible error estimates such as δ_k. We approximate the error in $C(\Gamma;V)$ by the maximal error on the finite sample set Γ_S, as in Section 5.2, and errors in $L^2_\pi(\Gamma;V)$ refer to the difference to a reference solution, evaluated using Parseval's identity.

In Figure 5, the convergence of $\texttt{SolveDirect}_{\mathcal{A},f}$ in $C(\Gamma;V)$ is compared to an analogous method from [Git11c], which is set in $L^2_\pi(\Gamma;V)$ instead of $C(\Gamma;V)$. We observe that the method which controls the error in $C(\Gamma;V)$ converges slightly faster, although the other

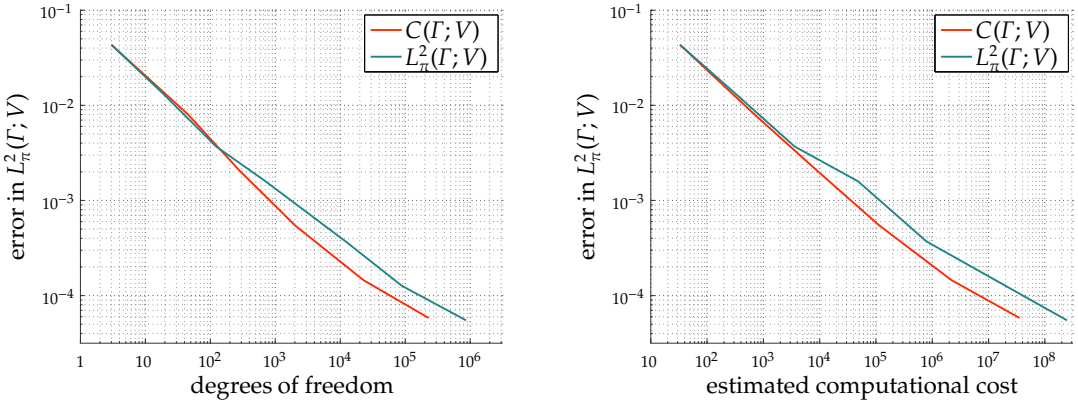

Figure 6: Convergence in $L^2_\pi(\Gamma;V)$ of $\texttt{SolveDirect}_{\mathcal{A},f}$ with Legendre polynomials. The two versions of $\texttt{SolveDirect}_{\mathcal{A},f}$ control the error in $C(\Gamma;V)$ and $L^2_\pi(\Gamma;V)$, respectively.

method also converges uniformly in the parameter $y \in \Gamma$.

Surprisingly, $\texttt{SolveDirect}_{\mathcal{A},f}$ with error control in $C(\Gamma;V)$ also converges faster in $L^2_\pi(\Gamma;V)$, as shown in Figure 6. Here, the reference solution has an error of approximately $5 \cdot 10^{-5}$, which may explain the slight flattening of the convergence curves.

In Figure 7 the convergence in $C(\Gamma;V)$ of $\texttt{SolveDirect}_{\mathcal{A},f}$ with error control in $C(\Gamma;V)$ is compared to that of the adaptive method $\texttt{SolveGalerkin}_{\mathcal{A},f}$ from [Git11b], which controls the error in the energy norm on $L^2_\pi(\Gamma;V)$. The parameters of $\texttt{SolveGalerkin}_{\mathcal{A},f}$ are chosen as in [Git11b]. The latter method includes a coarsening step, which ensures that the approximate solutions are sparse, *i.e.* for a given error tolerance in $L^2_\pi(\Gamma;V)$, the approximate solution constructed by $\texttt{SolveGalerkin}_{\mathcal{A},f}$ should contain a minimal number of degrees of freedom, up to a constant factor. Without such a coarsening procedure, $\texttt{SolveDirect}_{\mathcal{A},f}$ produces approximate solutions with almost identical sparsity if the error is measured in $C(\Gamma;V)$. However, the computational cost of $\texttt{SolveDirect}_{\mathcal{A},f}$ is two orders of magnitude lower than that of $\texttt{SolveGalerkin}_{\mathcal{A},f}$.

Conclusion and Outlook

Our adaptive methods are proven to converge uniformly in the parameter, which is assumed to be in an infinite dimensional cube. The convergence rates we observe in numerical computations presented in Section 5.2 differ from those suggested by the approximation results in [CDS10b, CDS10a]. In a semidiscrete setting, the observed convergence rates are lower, but the fully discrete algorithms converge faster than predicted by these approximation results.

The comparisons in Section 5.3 indicate that, for constructing a reliable parametric representation of the solution to a parametric boundary value problem, the adaptive methods presented here are more efficient than similar methods that control the error

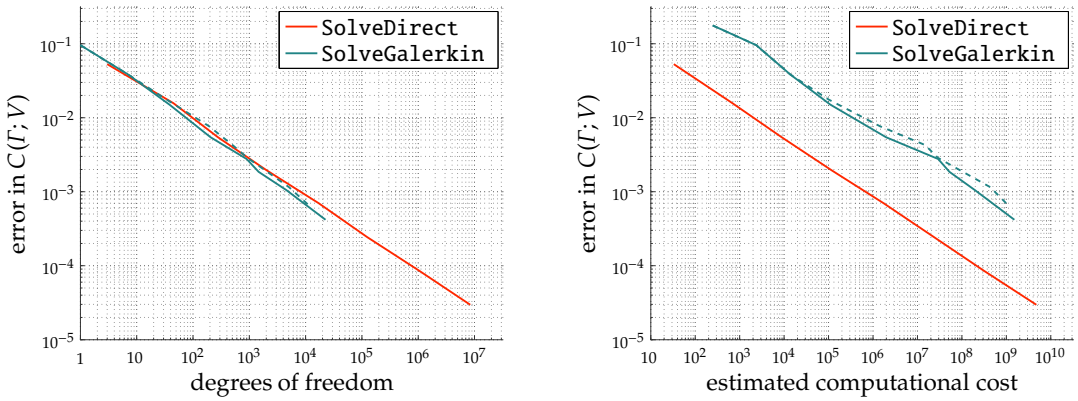

Figure 7: Convergence in $C(\Gamma;V)$ of $\texttt{SolveDirect}_{\mathcal{A},f}$ with Chebyshev polynomials and error control in $C(\Gamma;V)$ compared to that of $\texttt{SolveGalerkin}_{\mathcal{A},f}$ with Chebyshev polynomials (solid) and Legendre polynomials (dashed).

in $L^2_\pi(\Gamma;V)$ rather than $C(\Gamma;V)$.

Our methods provide a reliable upper bound for the error in $C(\Gamma;V)$. We observed that this bound may overestimate the actual error. It would be desirable to have a less conservative upper bound. Also, the addition of a coarsening step may lead to an even more efficient algorithm. These points are the subject of ongoing research.

References

[Bar05] A. Barinka. *Fast Evaluation Tools for Adaptive Wavelet Schemes*. PhD thesis, RWTH Aachen, March 2005.

[BAS10] Marcel Bieri, Roman Andreev, and Christoph Schwab. Sparse tensor discretization of elliptic SPDEs. *SIAM J. Sci. Comput.*, 31(6):4281–4304, 2009/10.

[BDD04] Peter Binev, Wolfgang Dahmen, and Ronald A. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[BNT07] Ivo M. Babuška, Fabio Nobile, and Raúl Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034 (electronic), 2007.

[BS09] Marcel Bieri and Christoph Schwab. Sparse high order FEM for elliptic sPDEs. *Comput. Methods Appl. Mech. Engrg.*, 198(37-40):1149–1170, 2009.

[BTZ04] Ivo M. Babuška, Raúl Tempone, and Georgios E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825 (electronic), 2004.

[CDD01] Albert Cohen, Wolfgang Dahmen, and Ronald A. DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Math. Comp.*, 70(233):27–75 (electronic), 2001.

[CDD02] Albert Cohen, Wolfgang Dahmen, and Ronald A. DeVore. Adaptive wavelet methods. II. Beyond the elliptic case. *Found. Comput. Math.*, 2(3):203–245, 2002.

[CDS10a] Albert Cohen, Ronald A. DeVore, and Christoph Schwab. Analytic regularity and polynomial approximation of parametric stochastic elliptic PDEs. Technical Report 2010-3, Seminar for Applied Mathematics, ETH Zürich, 2010. In review.

[CDS10b] Albert Cohen, Ronald A. DeVore, and Christoph Schwab. Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.

[DBO01] Manas K. Deb, Ivo M. Babuška, and J. Tinsley Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Engrg.*, 190(48):6359–6372, 2001.

[DFR07] Stephan Dahlke, Massimo Fornasier, and Thorsten Raasch. Adaptive frame methods for elliptic operator equations. *Adv. Comput. Math.*, 27(1):27–63, 2007.

[Dör96] Willy Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[DRW+07] Stephan Dahlke, Thorsten Raasch, Manuel Werner, Massimo Fornasier, and Rob Stevenson. Adaptive frame methods for elliptic operator equations: the steepest descent approach. *IMA J. Numer. Anal.*, 27(4):717–740, 2007.

[DSS09] Tammo Jan Dijkema, Christoph Schwab, and Rob Stevenson. An adaptive wavelet method for solving high-dimensional elliptic PDEs. *Constr. Approx.*, 30(3):423–455, 2009.

[FST05] Philipp Frauenfelder, Christoph Schwab, and Radu Alexandru Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):205–228, 2005.

[Gau04] Walter Gautschi. *Orthogonal polynomials: computation and approximation.* Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2004. Oxford Science Publications.

[GHS07] Tsogtgerel Gantumur, Helmut Harbrecht, and Rob Stevenson. An optimal adaptive wavelet method without coarsening of the iterands. *Math. Comp.*, 76(258):615–629 (electronic), 2007.

[Git11a] Claude Jeffrey Gittelson. *Adaptive Galerkin Methods for Parametric and Stochastic Operator Equations*. PhD thesis, ETH Zürich, 2011. ETH Dissertation No. 19533.

[Git11b] Claude Jeffrey Gittelson. An adaptive stochastic Galerkin method. Technical Report 2011-11, Seminar for Applied Mathematics, ETH Zürich, 2011.

[Git11c] Claude Jeffrey Gittelson. Adaptive stochastic Galerkin methods: Beyond the elliptic case. Technical Report 2011-12, Seminar for Applied Mathematics, ETH Zürich, 2011.

[KR97] Richard V. Kadison and John R. Ringrose. *Fundamentals of the theory of operator algebras. Vol. I*, volume 15 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997. Elementary theory, Reprint of the 1983 original.

[Met02] A. Metselaar. *Handling Wavelet Expansions in Numerical Methods*. PhD thesis, University of Twente, 2002.

[MK05] Hermann G. Matthies and Andreas Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16):1295–1331, 2005.

[MNS00] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38(2):466–488 (electronic), 2000.

[NTW08] Fabio Nobile, Raúl Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2008.

[Rud91] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[Ste03] Rob Stevenson. Adaptive solution of operator equations using wavelet frames. *SIAM J. Numer. Anal.*, 41(3):1074–1100 (electronic), 2003.

[Sto48] M. H. Stone. The generalized Weierstrass approximation theorem. *Math. Mag.*, 21:167–184, 237–254, 1948.

[Sze75] Gábor Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[TS07] Radu Alexandru Todor and Christoph Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232–261, 2007.

[WK05] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.*, 209(2):617–642, 2005.

[WK06] Xiaoliang Wan and George Em Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28(3):901–928 (electronic), 2006.

[WK09] Xiaoliang Wan and George Em Karniadakis. Solving elliptic problems with non-Gaussian spatially-dependent random coefficients. *Comput. Methods Appl. Mech. Engrg.*, 198(21-26):1985–1995, 2009.

[XK02] Dongbin Xiu and George Em Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644 (electronic), 2002.

Research Reports

No.	Authors/Title
11-19	*C.J. Gittelson* Uniformly convergent adaptive methods for parametric operator equations
11-18	*E. Kokiopoulou, D. Kressner, M. Zervos and N. Paragios* Optimal similarity registration of volumetric images
11-17	*D. Marazzina, O. Reichmann and Ch. Schwab* *hp*-DGFEM for Kolmogorov-Fokker-Planck equations of multivariate Lévy processes
11-16	*Ch. Schwab and A.M. Stuart* Sparse deterministic approximation of Bayesian inverse problems
11-15	*A. Barth and A. Lang* Almost sure convergence of a Galerkin–Milstein approximation for stochastic partial differential equations
11-14	*X. Claeys* A single trace integral formulation of the second kind for acoustic scattering
11-13	*W.-J. Beyn, C. Effenberger and D. Kressner* Continuation of eigenvalues and invariant pairs for parameterized nonlinear eigenvalue problems
11-12	*C.J. Gittelson* Adaptive stochastic Galerkin methods: beyond the elliptic case
11-11	*C.J. Gittelson* An adaptive stochastic Galerkin method
11-10	*C.J. Gittelson* Stochastic Galerkin approximation of operator equations with infinite dimensional noise
11-09	*R. Hiptmair, A. Moiola and I. Perugia* Error analysis of Trefftz-discontinuous Galerkin methods for the time-harmonic Maxwell equations
11-08	*W. Dahmen, C. Huang, Ch. Schwab and G. Welper* Adaptive Petrov-Galerkin methods for first order transport equations
11-07	*V.H. Hoang and Ch. Schwab* Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale PDEs