Eidgenössische Technische Hochschule Zürich

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

Stochastic Galerkin approximation of operator equations with infinite dimensional noise

C.J. Gittelson

Research Report No. 2011-10
February 2011

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

Stochastic Galerkin Approximation of Operator Equations with Infinite Dimensional Noise

Claude Jeffrey Gittelson*

February 28, 2011

Abstract

It is common practice in the study of stochastic Galerkin methods for boundary value problems depending on random fields to truncate a series representation of this field prior to the Galerkin discretization. We show that this is unnecessary; the projection onto a finite dimensional subspace automatically replaces the infinite series expansion by a suitable partial sum. We construct tensor product polynomial bases on infinite dimensional parameter domains, and use these to recast a random boundary value problem as a countably infinite system of deterministic equations. The stochastic Galerkin method can be interpreted as a standard finite element discretization of a finite section of this infinite system.

Introduction

Several numerical methods have emerged recently for solving boundary value problems in which some coefficients in the differential operator are random fields. The solution is then also a random field, and quantities of interest include statistics of this solution, or a parametric representation, for example in terms of polynomials of random variables appearing in a series representation of the input random field.

Stochastic Galerkin methods were introduced in [DBO01, XK02, BTZ04], and have been analyzed for example in [WK05, MK05, FST05, WK06, TS07, BS09, BAS10]. They approximate the random solution by a Galerkin projection onto a finite dimensional space of random fields. This requires the solution of a single coupled system of deterministic equations for the coefficients of the Galerkin projection with respect to a predefined set of basis functions on the parameter domain.

Stochastic collocation was studied *e.g.* in [XH05, BNT07, Bie09, WK09] as an alternative which requires only the solution of independent deterministic equations, and maintains

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

similar convergence properties. The coefficients that are computed directly by stochastic Galerkin can be obtained by evaluating certain integrals in a post-processing step, see [Xiu07].

For both methods, it is common practice to expand the input random field in a series, such as the Karhunen–Loève series, and then to truncate this expansion prior to any other approximation. This is often referred to as the *finite dimensional noise* assumption. Such an assumption is necessary in the stochastic collocation approach since, as in other sampling methods, individual realizations of the input random field must be approximated.

However, the finite dimensional noise assumption is superfluous in combination with the stochastic Galerkin method. For suitably chosen Galerkin subspaces, the approximate solution only depends on a partial sum of the series expansion of the input random field. It is therefore unnecessary to truncate this series prior to Galerkin approximation.

Keeping the full series introduces the problem of constructing an orthonormal basis on an infinite dimensional parameter domain. It turns out that the usual finite dimensional tensor product construction extends to countably infinite products by a straightforward limit argument, which is given in Section 2.2.

Passing to the coefficients of the solution with respect to an orthonormal basis on the parameter domain transforms the random boundary value problem into an equivalent countably infinite system of deterministic equations. The stochastic Galerkin method can be interpreted as a standard finite element approximation of a finite section of this infinite system.

This approach to stochastic Galerkin discretization combines all approximations into a single step: the choice of a finite dimensional subspace. It avoids errors in the representation of the random input, and is not affected by quadrature on the parameter domain. In particular, there is no need to equilibrate errors from various sources.

The countably infinite system of deterministic equations which represents a random boundary value problem with respect to an orthonormal basis on the parameter domain is analogous to the representation of a boundary value problem as a bi-infinite matrix equation using a wavelet basis. Adaptive methods for selecting a finite section of such bi-infinite matrix equations have been studied in [CDD01, GHS07, DSS09]. These techniques carry over to random boundary value problems, and allow the adaptive construction of Galerkin subspaces, see [Git11b, Git11c, Git11a].

Section 1 begins with a discussion of the isotropic diffusion equation with a stochastic diffusion coefficient. This example of a random boundary value problem motivates an abstract framework for such equations. We derive a weak formulation in this general setting, and define the Galerkin projection.

In Section 2, the weak formulation is recast as a countably infinite system of equations. We discuss orthonormal polynomials in one dimension, and tensorization of such polynomials to construct an orthonormal basis on an infinite dimensional parameter domain. Subsequently, we derive systems of deterministic equations that are equivalent to the original random boundary value problem and its Galerkin approximation, respectively.

Finally, in Section 3, we discuss some algorithmic aspects of the stochastic Galerkin method. We interpret the finite system of deterministic equations that determines the coefficients of the Galerkin approximation as a single operator-matrix equation, and consider the preconditioned conjugate gradient method as a solver.

1 Stochastic Operator Equations

1.1 The Isotropic 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

$$\begin{aligned} -\nabla \cdot (a(x)\nabla u(x)) &= f(x) \;, \quad x \in G \;, \\ u(x) &= 0 \;, \quad x \in \partial G \;. \end{aligned} \tag{1.1}$$

We view f as fixed, but allow a to vary, giving rise to a parametric operator

$$A_0(a)\colon H^1_0(G) \to H^{-1}(G) \;, \quad v \mapsto -\nabla \cdot (a \nabla v) \;, \tag{1.2}$$

which depends continuously on $a \in L^\infty(G)$.

We model the permeability a as a $L^\infty(G)$-valued random variable $\tilde{a}$ on a probability space $(\Omega, \mathcal{F}, P)$. The resulting stochastic diffusion equation is

$$A_0(\tilde{a}(\omega))U(\omega) = f \qquad \forall \omega \in \Omega \;. \tag{1.3}$$

The solution is a random variable U on $(\Omega, \mathcal{F})$ with values in $H^1_0(G)$. For all $\omega \in \Omega$, $U(\omega)$ is the weak solution of (1.1) with $a = \tilde{a}(\omega)$.

We assume that $\tilde{a}(\omega)$ is uniformly bounded from above and away from 0,

$$0 < \check{a} \le \tilde{a}(\omega, x) \le \hat{a} < \infty \qquad \forall x \in G \;, \quad \forall \omega \in \Omega \;. \tag{1.4}$$

Let $\bar{a} \in L^\infty(G)$ be some uniformly positive deterministic approximation of $\tilde{a}$. For example, $\bar{a}$ can be the mean field

$$\bar{a}\colon G \to \mathbb{R} \;, \quad \bar{a}(x) \coloneqq \int_\Omega \tilde{a}(\omega, x) \,\mathrm{d}P(\omega) \;, \tag{1.5}$$

or simply a constant $\bar{a} \coloneqq (\hat{a} + \check{a})/2$, $\bar{a} \coloneqq \sqrt{\hat{a}\check{a}}$, or $\bar{a} \coloneqq 1$.

For a countable set $\mathcal{M}$, let $(\varphi_m)_{m\in\mathcal{M}}$ be a frame of $L^2(G)$ with dual frame $(\varphi^*_m)_{m\in\mathcal{M}}$, which we interpret also as a sequence in $L^2(G)$. Define the random variables

$$Y_m(\omega) \coloneqq \frac{1}{\alpha_m} \int_G (\tilde{a}(\omega, x) - \bar{a}(x)) \, \varphi^*_m(x) \,\mathrm{d}x \;, \quad m \in \mathcal{M} \;. \tag{1.6}$$

Note that Y_m is bounded due to Hölder's inequality and (1.4). We assume that α_m is chosen such that $Y_m(\Omega) \subset [-1, 1]$ for all $m \in \mathcal{M}$. For example, this holds for

$$\alpha_m \coloneqq \sup_{\omega\in\Omega} \|\tilde{a}(\omega) - \bar{a}\|_{L^\infty(G)} \left\|\varphi^*_m\right\|_{L^1(G)} \;, \quad m \in \mathcal{M} \;. \tag{1.7}$$

Abbreviating $a_m \coloneqq \alpha_m \varphi_m$, we have

$$\tilde{a}(\omega, x) = \bar{a}(x) + \sum_{m \in \mathcal{M}} Y_m(\omega) a_m(x) \tag{1.8}$$

for all $\omega \in \Omega$ with convergence in $L^2(G)$. Let $\Gamma \coloneqq [-1,1]^{\mathcal{M}}$ and

$$a(y, x) \coloneqq \bar{a}(x) + \sum_{m \in \mathcal{M}} y_m a_m(x) \;, \quad y = (y_m)_{m \in \mathcal{M}} \in \Gamma \;. \tag{1.9}$$

Then $\tilde{a}(\omega, x) = a(Y(\omega), x)$ for all $\omega \in \Omega$, where $Y \coloneqq (Y_m)_{m \in \mathcal{M}}$.

Convergence of (1.9) is assured in $L^\infty(G)$ if the series $\sum_{m \in \mathcal{M}} |a_m|$ converges in $L^\infty(G)$. Furthermore, (1.9) defines a continuous map from Γ to $L^\infty(G)$, see [Git11a, Lemma 7.1.6]. This permits us to replace the parameter domain $L^\infty(G)$ by the product space $\Gamma = [-1,1]^{\mathcal{M}}$.

We define the parametric operator $A(y) \coloneqq A_0(a(y))$ for $y \in \Gamma$. Due to the linearity of A_0,

$$A(y) = D + R(y) \;, \quad R(y) \coloneqq \sum_{m \in \mathcal{M}} y_m R_m \qquad \forall y \in \Gamma \tag{1.10}$$

with convergence in $\mathcal{L}(H_0^1(G), H^{-1}(G))$, for

$$\begin{aligned} D &\coloneqq A_0(\bar{a}) \colon H_0^1(G) \to H^{-1}(G) \;, \quad v \mapsto -\nabla \cdot (\bar{a} \nabla v) \;, \\ R_m &\coloneqq A_0(a_m) \colon H_0^1(G) \to H^{-1}(G) \;, \quad v \mapsto -\nabla \cdot (a_m \nabla v) \;, \quad m \in \mathcal{M} \,. \end{aligned}$$

This leads to the parametric operator equation

$$A(y) u(y) = f \qquad \forall y \in \Gamma \tag{1.11}$$

with $A(y)$ from (1.10). The solution is related to the solution U of (1.3) by $U(\omega) = u(Y(\omega))$ for all $\omega \in \Omega$.

Lemma 1.1. *If there is a $\gamma \in [0,1)$ such that*

$$\operatorname*{ess\,sup}_{x \in G} \sum_{m \in \mathcal{M}} \frac{|a_m(x)|}{\bar{a}(x)} \le \gamma \;, \tag{1.12}$$

then $-\gamma D \le R(y) \le \gamma D$ for all $y \in \Gamma$, in the sense of symmetric operators on $H_0^1(G)$.

Proof. For all $v \in H_0^1(G)$ and all $y \in \Gamma$,

$$\left| \langle R(y) v, v \rangle \right| \le \int_G \left(\sum_{m \in \mathcal{M}} |a_m(x)| \right) |\nabla v(x)|^2 \, \mathrm{d}x \le \gamma \int_G \bar{a}(x) \, |\nabla v(x)|^2 \, \mathrm{d}x = \gamma \, \langle D v, v \rangle \;. \qquad \square$$

1.2 Abstract Setting

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and let V be a separable Hilbert space over $\mathbb{K}$. Let $D \in \mathcal{L}(V, V^*)$ be a positive symmetric operator, *i.e.* a bounded linear map from V to V^* for which $\langle D\cdot, \cdot\rangle$ is an inner product on V. Furthermore, let $\gamma \in [0, 1)$ and let $R(y) \in \mathcal{L}(V, V^*)$ be symmetric with $-\gamma D \leq R(y) \leq \gamma D$ for all $y \in \Gamma = [-1, 1]^{\mathcal{M}}$. We define the parametric operator

$$A(y) \coloneqq D + R(y) \,, \quad y \in \Gamma \,. \tag{1.13}$$

By definition, $A(y)$ is symmetric for all $y \in \Gamma$. Furthermore, D is boundedly invertible since it is positive. By the following proposition, the assumptions on the perturbation $R(y)$ ensure that $A(y)$ is also boundedly invertible, uniformly in $y \in \Gamma$.

Proposition 1.2. *For all $y \in \Gamma$, $A(y)$ is boundedly invertible; $A(y)$ and $A(y)^{-1}$ satisfy*

$$(1 - \gamma)D \leq A(y) \leq (1 + \gamma)D \,, \tag{1.14}$$

$$\frac{1}{1+\gamma} D^{-1} \leq A(y)^{-1} \leq \frac{1}{1-\gamma} D^{-1} \,. \tag{1.15}$$

Proof. Due to $R(y) \leq \gamma D$, for any $v \in V$,

$$\langle A(y)v, v\rangle = \langle Dv, v\rangle + \langle R(y)v, v\rangle \leq (1 + \gamma)\langle Dv, v\rangle \,,$$

which shows the second inequality in (1.14). The first follows by a similar estimate using $-\gamma D \leq R(y)$.

Consequently, the spectrum of $D^{-1/2}A(y)D^{-1/2}$ is in $[1 - \gamma, 1 + \gamma]$. Due to the spectral mapping theorem, the spectrum of $D^{1/2}A(y)^{-1}D^{1/2}$ is in $[(1 + \gamma)^{-1}, (1 - \gamma)^{-1}]$. This is equivalent to the statement (1.15). □

As in (1.11), we consider the parametric operator equation

$$A(y)u(y) = f(y) \qquad \forall y \in \Gamma \,, \tag{1.16}$$

where, for the sake of generality, we allow the right hand side $f(y) \in V^*$ to depend on the parameter $y \in \Gamma$. Motivated by (1.10), we are particularly interested in parametric operators of the form

$$R(y) = \sum_{m \in \mathcal{M}} y_m R_m \,, \quad \forall y \in \Gamma \,, \tag{1.17}$$

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

1.3 Weak Formulation

Let π be a probability measure on the parameter domain Γ with Borel σ-algebra $\mathcal{B}(\Gamma)$. In the example from Section 1.1, π could be the image of the physical probability P under the map Y, or it may be any other probability measure. We derive a weak formulation of (1.16) by integrating over Γ with respect to π.

Let the map $\Gamma \ni y \mapsto A(y)v(y)$ be measurable for any measurable $v\colon \Gamma \to V$. Then due to (1.14),

$$\mathcal{A}\colon L^2_\pi(\Gamma;V) \to L^2_\pi(\Gamma;V^*) \ , \quad v \mapsto [y \mapsto A(y)v(y)] \ , \tag{1.18}$$

is well-defined and continuous with norm at most $(1+\gamma)\,\|D\|_{V\to V^*}$. We assume also that $f \in L^2_\pi(\Gamma;V^*)$.

The weak formulation of (1.16) is to find $u \in L^2_\pi(\Gamma;V)$ such that

$$\int_\Gamma \langle A(y)u(y), v(y)\rangle \,\mathrm{d}\pi(y) = \int_\Gamma \langle f(y), v(y)\rangle \,\mathrm{d}\pi(y) \qquad \forall v \in L^2_\pi(\Gamma;V) \ . \tag{1.19}$$

The left term in (1.19) is the duality pairing in $L^2_\pi(\Gamma;V)$ of $\mathcal{A}u$ with the test function v, and the right term is the duality pairing of f with v. We follow the convention that the duality pairing is linear in the first argument and antilinear in the second.

Before turning to existence and uniqueness of the solution u of the linear variational problem (1.19), we introduce some additional notation. Let

$$(v,w)_A \coloneqq \int_\Gamma \langle A(y)v(y), w(y)\rangle \,\mathrm{d}\pi(y) \qquad \text{and} \qquad \|v\|_A \coloneqq \sqrt{(v,v)_A} \tag{1.20}$$

for $v, w \in L^2_\pi(\Gamma;V)$, and let $(\cdot,\cdot)_D$ and $\|\cdot\|_D$ be defined analogously, with $A(y)$ replaced by D.

Lemma 1.3. *The sesquilinear form $(\cdot,\cdot)_A$ is an inner product on $L^2_\pi(\Gamma;V)$, and the norm $\|\cdot\|_A$ is equivalent to the standard norm on $L^2_\pi(\Gamma;V)$, with*

$$(1-\gamma)\,\|v\|_D^2 \le \|v\|_A^2 \le (1+\gamma)\,\|v\|_D^2 \qquad \forall v \in L^2_\pi(\Gamma;V) \ , \tag{1.21}$$

$$\left\|D^{-1}\right\|^{-1}_{V^*\to V} \|v\|^2_{L^2_\pi(\Gamma;V)} \le \|v\|_D^2 \le \|D\|_{V\to V^*}\,\|v\|^2_{L^2_\pi(\Gamma;V)} \qquad \forall v \in L^2_\pi(\Gamma;V) \ . \tag{1.22}$$

Proof. Let $v \in L^2_\pi(\Gamma;V)$. By Proposition 1.2,

$$\int_\Gamma \langle A(y)v(y), v(y)\rangle \,\mathrm{d}\pi(y) \le (1+\gamma)\int_\Gamma \langle Dv(y), v(y)\rangle \,\mathrm{d}\pi(y) \le (1+\gamma)\,\|D\|_{V\to V^*}\,\|v\|^2_{L^2_\pi(\Gamma;V)} \ ,$$

and similarly,

$$\int_\Gamma \langle A(y)v(y), v(y)\rangle \,\mathrm{d}\pi(y) \ge (1-\gamma)\int_\Gamma \langle Dv(y), v(y)\rangle \,\mathrm{d}\pi(y) \ge (1-\gamma)\left\|D^{-1}\right\|^{-1}_{V^*\to V}\|v\|^2_{L^2_\pi(\Gamma;V)} \ .$$

This shows positivity of $(\cdot,\cdot)_A$, and the estimates (1.21). Symmetry of $(\cdot,\cdot)_A$ follows from the symmetry of $A(y)$ for all $y \in \Gamma$. □

Theorem 1.4. *For any $f \in L^2_\pi(\Gamma;V^*)$, the solution u of* (1.16) *is in $L^2_\pi(\Gamma;V)$ and u is the unique element of this space satisfying* (1.19).

Proof. It is tempting to simply multiply (1.16) by $v(y)$ integrate against π. However, we do not know a priori that u is measurable. We therefore first show that (1.19) has a unique solution in $L^2_\pi(\Gamma; V)$, and then that this solution coincides with that of (1.16).

By assumption, the right hand side of (1.19) is a continuous linear functional on $L^2_\pi(\Gamma; V)$. Since by Lemma 1.3, $(\cdot,\cdot)_A$ is an inner product on $L^2_\pi(\Gamma; V)$ which induces a norm equivalent to the standard norm on this space, the Riesz isomorphism ensures existence and uniqueness of the weak solution u of (1.19).

For $w \in V$ and $E \in \mathcal{B}(\Gamma)$, let $v(y) \coloneqq w1_E(y)$. By linearity, (1.19) implies

$$\int_E \langle A(y)u(y) - f(y), w\rangle \, \mathrm{d}\pi(y) = 0 \, .$$

Since this holds for all measurable sets E, the integrand is zero π-a.e. in Γ for any $w \in V$, and therefore the solution u of (1.19) satisfies (1.16) for π-a.e. $y \in \Gamma$. This implies that the solution of (1.16) is a version of the solution of (1.19), *i.e.* the two are equal in $L^2_\pi(\Gamma; V)$. □

Since V is a separable Hilbert space, the Lebesgue–Bochner space $L^2_\pi(\Gamma; V)$ is isometrically isomorphic to the Hilbert tensor product $L^2_\pi(\Gamma) \otimes V$, and $L^2_\pi(\Gamma; V^*)$ is isometrically isomorphic to $L^2_\pi(\Gamma) \otimes V^*$. By Theorem 1.4, $\mathcal{A}$ is a boundedly invertible linear map between these spaces.

We define the multiplication operators

$$K_m \colon L^2_\pi(\Gamma) \to L^2_\pi(\Gamma) \, , \quad v(y) \mapsto y_m v(y) \, , \quad m \in \mathcal{M} \, . \tag{1.23}$$

Since y_m is real and $\left|y_m\right|$ is less than one, K_m is symmetric and has norm at most one.

In the case (1.17), $\mathcal{A}$ can be expanded as

$$\mathcal{A} = \mathrm{id}_{L^2_\pi(\Gamma)} \otimes D + \sum_{m \in \mathcal{M}} K_m \otimes R_m \, . \tag{1.24}$$

We decompose this as $\mathcal{A} = \mathcal{D} + \mathcal{R}$ with

$$\mathcal{D} \coloneqq \mathrm{id}_{L^2_\pi(\Gamma)} \otimes D \qquad \text{and} \qquad \mathcal{R} \coloneqq \sum_{m \in \mathcal{M}} K_m \otimes R_m \, . \tag{1.25}$$

We focus on this setting in the following.

1.4 Galerkin Projection

Let $\mathcal{W}$ be a closed subspace of $L^2_\pi(\Gamma; V)$. The Galerkin solution $\bar{u} \in \mathcal{W}$ is defined through the linear variational problem

$$\int_\Gamma \langle A(y)\bar{u}(y), w(y)\rangle \, \mathrm{d}\pi(y) = \int_\Gamma \langle f(y), w(y)\rangle \, \mathrm{d}\pi(y) \qquad \forall w \in \mathcal{W} \, . \tag{1.26}$$

Proposition 1.5. *There is a unique $\bar{u} \in \mathcal{W}$ satisfying* (1.26)*. Furthermore,*

$$\|\bar{u} - u\|_{L^2_\pi(\Gamma;V)} \leq \sqrt{\frac{1+\gamma}{1-\gamma}\,\kappa(D)} \inf_{w\in\mathcal{W}} \|w - u\|_{L^2_\pi(\Gamma;V)} \ , \tag{1.27}$$

where $\kappa(D) \coloneqq \|D\|_{V\to V^*} \left\|D^{-1}\right\|_{V^*\to V}$ is the condition number of D.

Proof. Existence and uniqueness of $\bar{u}$ are ensured by Lemma 1.3 since $\bar{u}$ is just the $\|\cdot\|_A$-orthogonal projection of u onto the closed subspace $\mathcal{W}$ of $L^2_\pi(\Gamma;V)$. Furthermore, (1.21) implies

$$(1-\gamma)\left\|D^{-1}\right\|^{-1}_{V^*\to V} \|\bar{u} - u\|^2_{L^2_\pi(\Gamma;V)} \leq \|\bar{u} - u\|^2_A \leq \|w - u\|^2_A \leq (1+\gamma)\,\|D\|_{V\to V^*}\,\|w - u\|^2_{L^2_\pi(\Gamma;V)}$$

for any $w \in \mathcal{W}$, and (1.27) follows by taking the infimum over all such w. □

2 Transformation to a System of Deterministic Equations

2.1 Orthonormal Polynomials

Let μ be a Borel measure on $[-1,1]$. Let $\Delta \coloneqq \{0,1,\dots,N-1\}$ if the support of μ has cardinality $N \in \mathbb{N}$, and $\Delta \coloneqq \mathbb{N}_0$ otherwise. Let $P_{-1} \coloneqq 0$, $P_0 \coloneqq 1$ and

$$\beta_n P_n(\xi) \coloneqq (\xi - \alpha_{n-1})P_{n-1}(\xi) - \beta_{n-1}P_{n-2}(\xi) \ , \quad n \in \Delta \setminus \{0\} \ , \tag{2.1}$$

with

$$\alpha_n \coloneqq \int_{-1}^{1} \xi P_n(\xi)^2 \,\mathrm{d}\mu(\xi) \qquad \text{and} \qquad \beta_n \coloneqq \frac{c_{n-1}}{c_n} \ , \tag{2.2}$$

where c_n is the leading coefficient of P_n, $\beta_0 \coloneqq 1$, and P_n is chosen as normalized in $L^2_\mu([0,1])$, with a positive leading coefficient. The values $(\alpha_n)_{n\in\Delta}$ and $(\beta_n)_{n\in\Delta}$ are conveniently tabulated for many common distributions μ; see Table 1 for the coefficients of a few classical polynomials, or *e.g.* [Gau04], which tabulates β_n^2 in place of β_n. The following proposition is shown *e.g.* in [Gau04, Sze75].

Proposition 2.1. *The sequence $(P_n)_{n\in\Delta}$ is an orthonormal basis of $L^2_\mu([-1,1])$. Furthermore, P_n is a polynomial of degree n for all $n \in \Delta$.*

Remark 2.2. The above construction generalizes to Borel measures μ on $\mathbb{R}$. The polynomials P_n are well-defined by (2.1) if the moments

$$M_n \coloneqq \int \xi^n \,\mathrm{d}\mu(\xi) \ , \quad n \in \mathbb{N}_0 \ , \tag{2.3}$$

are finite, and they are orthonormal by construction. They form an orthonormal basis of $L^2_\mu(\mathbb{R})$ if the measure μ is uniquely characterized by its moments $(M_n)_{n\in\mathbb{N}_0}$, see *e.g.* [EMSU10, Theorem 3.2], [Fre71, Theorem 4.3], [Ber96, Theorem 2.1] and [Rie23] for details. We note that the construction in Section 2.2 below also goes through in this setting. ⌟

Table 1: Recursion coefficients of orthonormal polynomials on $[-1,1]$ w.r.t. $w(\xi)\,\mathrm{d}\xi$.

Name	$w(\xi)$	α_n	β_n
Legendre	$\frac{1}{2}$	0	$\frac{1}{\sqrt{4-n^{-2}}}$
Chebyshev #1	$\frac{1}{\pi}(1-\xi^2)^{-1/2}$	0	$\frac{1}{\sqrt{2}}$, $n=1$ $\frac{1}{2}$, $n\geq 2$
Chebyshev #2	$\frac{2}{\pi}(1-\xi^2)^{1/2}$	0	$\frac{1}{2}$
Chebyshev #3	$\frac{1}{\pi}(1-\xi)^{-1/2}(1+\xi)^{1/2}$	$\frac{1}{2}$, $n=0$ 0, $n\geq 1$	$\frac{1}{2}$
Chebyshev #4	$\frac{1}{\pi}(1-\xi)^{1/2}(1+\xi)^{-1/2}$	$-\frac{1}{2}$, $n=0$ 0, $n\geq 1$	$\frac{1}{2}$
Gegenbauer, $\lambda > -\frac{1}{2}$	$\frac{\Gamma(\lambda+1)}{\sqrt{\pi}\Gamma(\lambda+\frac{1}{2})}(1-\xi^2)^{\lambda-1/2}$	0	$\frac{1}{\sqrt{2+2\lambda}}$, $n=1$ $\frac{1}{2}\sqrt{\frac{n(n+2\lambda-1)}{(n+\lambda)(n+\lambda-1)}}$

Example 2.3 (Legendre Polynomials). If μ is the uniform probability measure on $[-1,1]$, *i.e.* $\mathrm{d}\mu(\xi) = \frac{1}{2}\,\mathrm{d}\xi$, then $(P_n)_{n\in\mathbb{N}_0}$ consists of the normalized Legendre polynomials defined by Rodrigues' formula

$$P_n(\xi) = L_n(\xi) \coloneqq \frac{\sqrt{2n+1}}{2^n n!}\,\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}(\xi^2-1)^n\,, \quad n\in\mathbb{N}_0\,. \tag{2.4}$$

Normalized Legendre polynomials satisfy the three term recursion

$$\frac{n+1}{\sqrt{2n+3}\sqrt{2n+1}}\,L_{n+1}(\xi) = \xi L_n(\xi) - \frac{n}{\sqrt{2n+1}\sqrt{2n-1}}\,L_{n-1}(\xi)\,, \quad n\in\mathbb{N}_0\,, \tag{2.5}$$

with $L_{-1} \coloneqq 0$. In particular, $\alpha_n = 0$ for all $n\in\mathbb{N}_0$ and

$$\beta_n = \frac{n}{\sqrt{2n+1}\sqrt{2n-1}} = \frac{1}{\sqrt{4-n^{-2}}} \in \left(\frac{1}{2}, \frac{1}{\sqrt{3}}\right], \quad n\in\mathbb{N}\,. \tag{2.6}$$

The first few Legendre polynomials are

$$L_0(\xi) = 1\,, \quad L_1(\xi) = \sqrt{3}\,\xi\,, \quad L_2(\xi) = \frac{\sqrt{5}}{2}\,(3\xi^2-1)\,. \tag{2.7}$$

Note that these polynomials differ from the standard definition of Legendre polynomials by a constant factor. ⌟

Example 2.4 (Jacobi Polynomials). Jacobi polynomials generalize Legendre polynomials to certain nonuniform distributions on $[-1,1]$. For parameters $a > -1$ and $b > -1$, we consider the probability measure $\mathrm{d}\mu(\xi) = w(\xi)\,\mathrm{d}\xi$ for the weight function

$$w(\xi) = 2^{-(a+b+1)}\frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b+1)}(1-\xi)^a(1+\xi)^b\,, \quad \xi\in[-1,1]\,. \tag{2.8}$$

The Jacobi polynomials can be constructed through the recursion (2.1) with the coefficients

$$\alpha_0 = \frac{b-a}{a+b+2}\,, \quad \alpha_n = \frac{b^2-a^2}{(2n+1+b)(2n+a+b+2)}\,, \quad n \geq 1\,, \tag{2.9}$$

and

$$\beta_n = \begin{cases} \sqrt{\dfrac{4(a+1)(b+1)}{(a+b+2)^2(a+b+3)}} & \text{if } n = 1, \\ \sqrt{\dfrac{4n(n+a)(n+b)(n+a+b)}{(2n+a+b)^2(2n+a+b+1)(2n+a+b-1)}} & \text{if } n \geq 2. \end{cases} \tag{2.10}$$

See Table 1 for the coefficients of a few particular cases of Jacobi polynomials. ⌟

2.2 Tensor Product Bases

We return to the parameter domain $\Gamma = [-1,1]^{\mathcal{M}}$ from Section 1. Let π be a probability measure on $(\Gamma, \mathcal{B}(\Gamma))$; further assumptions will be made on π below, as they are needed.

Let $\mathcal{F}(\mathcal{M})$ denote the set of finite subsets of $\mathcal{M}$. For any $I \in \mathcal{F}(\mathcal{M})$, we define the finite product domain $\Gamma_I := [-1,1]^I$. The coordinate maps $y_I \colon \Gamma \to \Gamma_I$ are continuous, and thus in particular Borel measurable. They generate σ-algebras on Γ, which we denote by $\mathcal{B}_I := \sigma(y_I)$. We define $L^2_{\pi|I}(\Gamma)$ to be the space of $\mathcal{B}_I$-measurable elements of $L^2_\pi(\Gamma)$. Furthermore, we denote by $\pi_I := y_I(\pi)$ the image of π under the map y_I, which is a probability measure on $(\Gamma_I, \mathcal{B}(\Gamma_I))$.

Lemma 2.5. *For all $I \in \mathcal{F}(\mathcal{M})$, the map*

$$L^2_{\pi_I}(\Gamma_I) \to L^2_\pi(\Gamma)\,, \quad v \mapsto v \circ y_I\,, \tag{2.11}$$

is an isometry with range $L^2_{\pi|I}(\Gamma)$.

Proof. Let $I \in \mathcal{F}(\mathcal{M})$ and $v \in L^2_{\pi_I}(\Gamma_I)$. Then $v \circ y_I \in L^2_\pi(\Gamma)$ and $v \circ y_I$ is $\mathcal{B}_I = \sigma(y_I)$-measurable, so $v \circ y_I \in L^2_{\pi|I}(\Gamma)$. Conversely, let $w \in L^2_{\pi|I}(\Gamma)$. By the Doob–Dynkin lemma, there is a measurable function v on $(\Gamma_I, \mathcal{B}(\Gamma_I))$ such that $w = v \circ y_I$. Furthermore, since $\pi_I = y_I(\pi)$,

$$\int_{\Gamma_I} |v|^2 \,\mathrm{d}\pi_I = \int_\Gamma \left|v \circ y_I\right|^2 \mathrm{d}\pi = \int_\Gamma |w|^2 \,\mathrm{d}\pi\,.$$

This shows $v \in L^2_{\pi_I}(\Gamma_I)$ and that the map is an isometry. □

We recall the monotone class theorem, see for example [Pro05, Theorem I.8]. A set $\mathfrak{M}$ of real-valued functions on Γ is *multiplicative* if $v, w \in \mathfrak{M}$ implies $vw \in \mathfrak{M}$. A *monotone vector space* over Γ is a real vector space $\mathfrak{H}$ of bounded, real-valued functions on Γ such that all constants are in $\mathfrak{H}$ and if $(v_n)_{n\in\mathbb{N}}$ is a sequence in $\mathfrak{H}$ with $0 \leq v_n \leq v_{n+1}$ for all $n \in \mathbb{N}$ and $v := \sup_n v_n$ is a bounded function on Γ, then $v \in \mathfrak{H}$.

Theorem 2.6 (Monotone Class Theorem). *Let $\mathfrak{M}$ be a multiplicative class of bounded, real-valued functions on Γ, and let $\mathfrak{H}$ be a monotone vector space containing $\mathfrak{M}$. Then $\mathfrak{H}$ contains all bounded $\sigma(\mathfrak{M})$-measurable functions.*

Proposition 2.7. *$\bigcup_{I\in\mathcal{F}(\mathcal{M})} L^2_{\pi|I}(\Gamma)$ is a dense subspace of $L^2_\pi(\Gamma)$.*

Proof. Let $\mathfrak{V} \coloneqq \overline{\bigcup_{I\in\mathcal{F}(\mathcal{M})} L^2_{\pi|I}(\Gamma)} \subset L^2_\pi(\Gamma)$ and define $\mathfrak{H} \coloneqq \mathfrak{V} \cap L^\infty_\pi(\Gamma)$ as the vector space of bounded functions in $\mathfrak{V}$. Let $\mathfrak{M} \coloneqq \{1_S\,;\, S \in \bigcup_{I\in\mathcal{F}(\mathcal{M})} \mathcal{B}_I\}$ be the set of indicator functions that are in $L^2_{\pi|I}(\Gamma)$ for some $I \in \mathcal{F}(\mathcal{M})$. Then $\mathfrak{M} \subset \mathfrak{H}$, $1 \in \mathfrak{H}$, and $\mathfrak{M}$ is closed under multiplication. Let $0 \le v_1 \le v_2 \le \cdots$ be a pointwise monotonic sequence in $\mathfrak{H}$ and $v \coloneqq \sup_n v_n$ its supremum. If $v \in L^\infty_\pi(\Gamma) \subset L^2_\pi(\Gamma)$, then $(v_n)_n$ converges to v in $L^2_\pi(\Gamma)$ by dominated convergence. Since $\mathfrak{V}$ is closed in $L^2_\pi(\Gamma)$, $v \in \mathfrak{V}$ and therefore $v \in \mathfrak{H}$. Thus $\mathfrak{H}$ is a monotone vector space and, using $\mathcal{B}(\Gamma) = \sigma(\mathfrak{M})$, the monotone class theorem implies $\mathfrak{H} = L^\infty_\pi(\Gamma)$.

If $v \in L^2_\pi(\Gamma)$, then for any $N \in \mathbb{N}$, $v1_{\{|v|\le N\}} \in L^\infty_\pi(\Gamma) = \mathfrak{H} \subset \mathfrak{V}$ and $v \in \mathfrak{V}$ by dominated convergence. □

In order to construct an orthonormal polynomial basis of $L^2_\pi(\Gamma)$, we assume that π is a product measure. Let

$$\pi = \bigotimes_{m\in\mathcal{M}} \pi_m \tag{2.12}$$

for probability measures π_m on $([-1,1], \mathcal{B}([-1,1]))$; see *e.g.* [Bau02, Section 9] for a general construction of arbitrary products of probability measures. Then $\pi_{\{m\}} = \pi_m$ and $\pi_I = \bigotimes_{m\in I} \pi_m$ for any $I \in \mathcal{F}(\mathcal{M})$.

For all $m \in \mathcal{M}$, let $(P^m_n)_{n\in\Lambda_m}$ denote the orthonormal polynomial basis of $L^2_{\pi_m}([-1,1])$ from Proposition 2.1, with recursion coefficients $(\alpha^m_n)_{n\in\Lambda_m}$ and $(\beta^m_n)_{n\in\Lambda_m}$. We define the set of finitely supported sequences in $\mathbb{N}_0$, indexed by $\mathcal{M}$, as

$$\Lambda \coloneqq \left\{\nu \in \mathbb{N}_0^{\mathcal{M}}\,;\, \nu_m \in \Lambda_m\ \forall m \in \mathcal{M},\ \#\operatorname{supp}\nu < \infty\right\}, \tag{2.13}$$

where the support is defined by

$$\operatorname{supp}\nu \coloneqq \{m \in \mathcal{M}\,;\, \nu_m \neq 0\}, \quad \nu \in \mathbb{N}_0^{\mathcal{M}}. \tag{2.14}$$

Dropping all zeros, Λ can also be interpreted as the set of sequences ν in $\mathbb{N}$ indexed by any $I \in \mathcal{F}(\mathcal{M})$, with $\nu_m \in \Lambda_m \setminus \{0\}$ for all $m \in I$.

We define the countable tensor product polynomials

$$\boldsymbol{P} \coloneqq (P_\nu)_{\nu\in\Lambda}, \quad P_\nu \coloneqq \bigotimes_{m\in\mathcal{M}} P^m_{\nu_m}, \quad \nu \in \Lambda. \tag{2.15}$$

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

$$P_\nu(y) = \prod_{m\in\mathcal{M}} P^m_{\nu_m}(y_m) = \prod_{m\in\operatorname{supp}\nu} P^m_{\nu_m}(y_m), \quad \nu \in \Lambda, \tag{2.16}$$

since $P^m_0 = 1$ for all $m \in \mathcal{M}$.

Theorem 2.8. *$\boldsymbol{P}$ is an orthonormal basis of $L^2_\pi(\Gamma)$.*

Proof. By Proposition 2.1, $(P^m_n)_{n\in\Lambda_m}$ is an orthonormal basis of $L^2_{\pi_m}([-1,1])$ for all $m \in \mathcal{M}$. Consequently, for any $I \in \mathcal{F}(\mathcal{M})$, the products

$$\bigotimes_{m\in I} P^m_{\nu_m} , \quad \nu = (\nu_m)_{m\in I} , \quad \nu_m \in \Lambda_m \quad \forall m \in I ,$$

form an orthonormal basis of $L^2_{\pi_I}(\Gamma_I)$. By Lemma 2.5, interpreting these functions on Γ rather than Γ_I, they form an orthonormal basis of $L^2_{\pi|I}(\Gamma)$. Proposition 2.7 implies that the union $\boldsymbol{P}$ of all of these bases spans $L^2_\pi(\Gamma)$. □

Theorem 2.8 is similar to [EMSU10, Theorem 3.6]. The latter is formulated in a more general setting, but does not provide an explicit construction for an orthonormal basis. It uses [Bob05, Corollary 3.6.8], which is very close to Proposition 2.7.

By Parseval's identity, Theorem 2.8 is equivalent to the statement that the map

$$T\colon \ell^2(\Lambda) \to L^2_\pi(\Gamma) , \quad (c_\nu)_{\nu\in\Lambda} \mapsto \sum_{\nu\in\Lambda} c_\nu P_\nu , \tag{2.17}$$

is a unitary isomorphism. The inverse of T is

$$T^{-1} = T^*\colon L^2_\pi(\Gamma) \to \ell^2(\Lambda) , \quad g \mapsto \left(\int_\Gamma g(y)\overline{P_\nu(y)}\,\mathrm{d}\pi(y)\right)_{\nu\in\Lambda} . \tag{2.18}$$

2.3 Discrete Operator Equation

We use the isomorphism T from (2.17) to recast the weak stochastic operator equation (1.19) and the equation (1.26) for the Galerkin approximation as an equivalent discrete operator equation.

For all $\nu \in \Lambda$, let W_ν be a closed subspace of V, and let

$$\mathcal{S} \coloneqq \left(\prod_{\nu\in\Lambda} W_\nu\right) \cap \ell^2(\Lambda; V) . \tag{2.19}$$

Since coordinate projections are continuous on $\ell^2(\Lambda; V)$ and $W_\nu \subset V$ is closed for all $\nu \in \Lambda$, $\mathcal{S}$ is an intersection of closed subspaces of $\ell^2(\Lambda; V)$, and as such is again a closed subspace of $\ell^2(\Lambda; V)$. The dimension of $\mathcal{S}$ is

$$\dim \mathcal{S} = \sum_{\nu\in\Lambda} \dim W_\nu . \tag{2.20}$$

Consequently, $\mathcal{S}$ is finite dimensional if and only if W_ν is finite dimensional for all $\nu \in \Lambda$, and $W_\nu = \{0\}$ for all but finitely many $\nu \in \Lambda$. If $W_\nu = V$ for all $\nu \in \Lambda$, then $\mathcal{S} = \ell^2(\Lambda; V)$.

Since T from (2.17) is a unitary map from $\ell^2(\Lambda)$ to $L^2_\pi(\Gamma)$, the tensor product operator $T \otimes \mathrm{id}_V$ is an isometric isomorphism from $\ell^2(\Lambda; V)$ to $L^2_\pi(\Gamma; V)$. We define T_V as the

restriction of $T \otimes \mathrm{id}_V$ to $\mathcal{S}$. Since $\mathcal{S}$ is a closed subspace of $\ell^2(\Lambda; V)$, its range $\mathcal{W} \coloneqq \mathrm{range}(T_V)$ is a closed subspace of $L^2_\pi(\Gamma)$. The inverse of T_V is the restriction of $T^{-1} \otimes \mathrm{id}_V$ to $\mathcal{W}$. By definition, $w \in \mathcal{W}$ and $\boldsymbol{w} = (w_\nu)_{\nu\in\Lambda} \in \mathcal{S}$ are related by $w = T_V \boldsymbol{w}$ if

$$w(y) = \sum_{\nu\in\Lambda} w_\nu P_\nu(y) \qquad \text{or} \qquad w_\nu = \int_\Gamma w(y)\overline{P_\nu(y)}\,\mathrm{d}\pi(y) \quad \forall \nu \in \Lambda\,, \tag{2.21}$$

and either of these properties implies the other. The series in (2.21) converges unconditionally in $L^2_\pi(\Gamma; V)$, and the integral can be interpreted as a Bochner integral in V.

Proposition 2.9. *The Galerin solution $\bar{u} \in \mathcal{W}$ from Proposition 1.5 satisfies $\bar{u} = T_V \bar{\boldsymbol{u}}$ for $\bar{\boldsymbol{u}} \in \mathcal{S}$ with*

$$\boldsymbol{A}\bar{\boldsymbol{u}} = \boldsymbol{f} \qquad \text{for} \quad \boldsymbol{A} \coloneqq T_V^* \mathcal{A} T_V \quad \text{and} \quad \boldsymbol{f} \coloneqq T_V^* f\,. \tag{2.22}$$

In particular, (2.22) *has a unique solution $\bar{\boldsymbol{u}} \in \mathcal{S}$.*

Proof. By Parseval's identity, $T \otimes \mathrm{id}_V$ is an isometric isomorphism from $\ell^2(\Lambda; V)$ to $L^2_\pi(\Gamma; V)$. Therefore, its restriction T_V to $\mathcal{S}$ is an isometric isomorphism onto its range $\mathcal{W}$, and T_V^* is an isomorphism from $\mathcal{W}^*$ to $\mathcal{S}^*$. This shows the equivalence of (2.22) to (1.26), and existence and uniqueness of $\bar{\boldsymbol{u}} \in \mathcal{S}$ follows from Proposition 1.5. □

Proposition 2.9 implies that $\boldsymbol{A}$ is an isomorphism from $\mathcal{S}$ to $\mathcal{S}^*$. It can be written as $\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{R}$ for $\boldsymbol{D} \coloneqq T_V^* \mathcal{D} T_V$ and $\boldsymbol{R} \coloneqq T_V^* \mathcal{R} T_V$. We note that $\boldsymbol{D}$ is a boundedly invertible linear map from $\mathcal{S}$ to $\mathcal{S}^*$ due to Proposition 2.9 for $\boldsymbol{D}$ in place of $\boldsymbol{A}$, *i.e.* if $\mathcal{R} = 0$.

Proposition 2.10. *The operators $\boldsymbol{A}$ and $\boldsymbol{D}$ satisfy*

$$(1-\gamma)\boldsymbol{D} \le \boldsymbol{A} \le (1+\gamma)\boldsymbol{D}\,, \tag{2.23}$$

$$\frac{1}{1+\gamma}\boldsymbol{D}^{-1} \le \boldsymbol{A}^{-1} \le \frac{1}{1-\gamma}\boldsymbol{D}^{-1}\,. \tag{2.24}$$

Proof. We first show that $-\gamma \boldsymbol{D} \le \boldsymbol{R} \le \gamma \boldsymbol{D}$. For all $\boldsymbol{w} \in \mathcal{S}$, if $w \coloneqq T_V \boldsymbol{w} \in \mathcal{W}$,

$$\langle \boldsymbol{R}\boldsymbol{w}, \boldsymbol{w}\rangle = \int_\Gamma \langle R(y)w(y), w(y)\rangle\,\mathrm{d}\pi(y) \le \gamma \int_\Gamma \langle Dw(y), w(y)\rangle\,\mathrm{d}\pi(y) = \gamma\,\langle \boldsymbol{D}\boldsymbol{w}, \boldsymbol{w}\rangle\,,$$

and a similar estimate implies $\langle \boldsymbol{R}\boldsymbol{w}, \boldsymbol{w}\rangle \ge -\gamma\,\langle \boldsymbol{D}\boldsymbol{w}, \boldsymbol{w}\rangle$. Since $\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{R}$, this shows (2.23), and (2.24) follows from the spectral mapping theorem as in Proposition 1.2. □

In particular, using $\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{A}$, (2.24) leads to

$$\frac{1}{1+\gamma}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A} \le \boldsymbol{A} \le \frac{1}{1-\gamma}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A}\,. \tag{2.25}$$

Proposition 2.10 implies that $\boldsymbol{A}$ and $\boldsymbol{D}$ induce equivalent norms on $\mathcal{S}$, which we denote by $\|\boldsymbol{w}\|_{\boldsymbol{A}} \coloneqq \sqrt{\langle \boldsymbol{A}\boldsymbol{w}, \boldsymbol{w}\rangle}$ and $\|\boldsymbol{w}\|_{\boldsymbol{D}} \coloneqq \sqrt{\langle \boldsymbol{D}\boldsymbol{w}, \boldsymbol{w}\rangle}$. By definition, $\|\boldsymbol{w}\|_{\boldsymbol{A}} = \|T_V \boldsymbol{w}\|_A$ and $\|\boldsymbol{w}\|_{\boldsymbol{D}} = \|T_V \boldsymbol{w}\|_D$ for all $\boldsymbol{w} \in \mathcal{S}$. Therefore, Lemma 1.3 implies that these norms are equivalent to the standard $\ell^2(\Lambda; V)$-norm on $\mathcal{S}$.

2.4 System of Deterministic Equations

We interpret the discrete operator equation (2.22) as a system of deterministic equations for $\bar{\boldsymbol{u}} = (\bar{u}_\mu)_{\mu\in\Lambda}$.

Lemma 2.11. *For all $m \in \mathcal{M}$, the operator $\boldsymbol{K}_m = T^* K_m T \in \mathcal{L}(\ell^2(\Lambda))$ has the form*

$$(\boldsymbol{K}_m \boldsymbol{c})_\mu = \beta^m_{\mu_m+1} c_{\mu+\epsilon_m} + \alpha^m_{\mu_m} c_\mu + \beta^m_{\mu_m} c_{\mu-\epsilon_m} , \quad \mu \in \Lambda , \tag{2.26}$$

for $\boldsymbol{c} = (c_\mu)_{\mu\in\Lambda} \in \ell^2(\Lambda)$, where ϵ_m is the Kronecker sequence $(\epsilon_m)_n = \delta_{nm}$, and $c_\mu := 0$ if $\mu_m < 0$ for any $m \in \mathcal{M}$. It satisfies $\boldsymbol{K}_m^* = \boldsymbol{K}_m$ and $\|\boldsymbol{K}_m\|_{\ell^2(\Lambda)\to\ell^2(\Lambda)} \le 1$.

Proof. As noted after (1.23), K_m is symmetric and has norm at most one. Since T is unitary, these properties carry over to $\boldsymbol{K}_m$. Let $\boldsymbol{c} = (c_\mu)_{\mu\in\Lambda} \in \ell^2(\Lambda)$ and $m \in \mathcal{M}$. The recursion formula (2.1) can be rearranged to read

$$\xi P^m_n(\xi) = \beta^m_{n+1} P^m_{n+1}(\xi) + \alpha^m_n P^m_n(\xi) + \beta^m_n P^m_{n-1}(\xi) , \quad n \in \Lambda_m , \quad \xi \in [-1,1] ,$$

where $P^m_n := 0$ for $n \in \mathbb{Z} \setminus \Lambda_m$. Therefore,

$$\begin{aligned} K_m T \boldsymbol{c} &= \sum_{\mu\in\Lambda} c_\mu K_m P_\mu = \sum_{\mu\in\Lambda} c_\mu (\beta^m_{\mu_m+1} P_{\mu+\epsilon_m} + \alpha^m_{\mu_m} P_\mu + \beta^m_{\mu_m} P_{\mu-\epsilon_m}) \\ &= \sum_{\mu\in\Lambda} (\beta^m_{\mu_m+1} c_{\mu+\epsilon_m} + \alpha^m_{\mu_m} c_\mu + \beta^m_{\mu_m} c_{\mu-\epsilon_m}) P_\mu . \end{aligned}$$

Equation (2.26) follows since $T^* = T^{-1}$. □

Remark 2.12. If π_m is a symmetric measure on $[-1,1]$, then $\alpha^m_n = 0$ for all $n \in \Lambda_m$ by symmetry of the integral (2.2). This simplifies (2.26). ⌟

Lemma 2.13. *The operator $\boldsymbol{A}\colon \mathcal{S} \to \mathcal{S}^*$ has the form*

$$\boldsymbol{A} = \boldsymbol{I} \otimes D + \sum_{m\in\mathcal{M}} \boldsymbol{K}_m \otimes R_m , \tag{2.27}$$

where $\boldsymbol{I} := \mathrm{id}_{\ell^2(\Lambda)}$.

Proof. Since (1.24) holds with convergence in $\mathcal{L}(L^2_\pi(\Gamma;V), L^2_\pi(\Gamma;V^*))$,

$$\begin{aligned} (T^* \otimes \mathrm{id}_{V^*}) \mathcal{A} (T \otimes \mathrm{id}_V) = {} & (T^* \otimes \mathrm{id}_{V^*})(\mathrm{id}_{L^2_\pi(\Gamma)} \otimes D)(T \otimes \mathrm{id}_V) \\ & + \sum_{m\in\mathcal{M}} (T^* \otimes \mathrm{id}_{V^*})(K_m \otimes R_m)(T \otimes \mathrm{id}_V) \\ = {} & \boldsymbol{I} \otimes D + \sum_{m\in\mathcal{M}} \boldsymbol{K}_m \otimes R_m . \end{aligned}$$

Equation (2.27) follows by restricting to $\mathcal{S}$. □

Remark 2.14. Combining Lemmas 2.11 and 2.13, we can interpret $\boldsymbol{A}$ as a bi-infinite operator matrix. For any $\nu \in \Lambda$, let I_ν be the embedding of W_ν into V. Its adjoint I_ν^* is the restriction of functionals on V onto the subspace W_ν. Then

$$\boldsymbol{A} = [A_{\nu\mu}]_{\nu,\mu\in\Lambda} , \quad A_{\nu\mu} \colon W_\mu \to W_\nu^* , \tag{2.28}$$

with entries

$$\begin{aligned} A_{\nu\nu} &= I_\nu^* \left(D + \sum_{m\in\mathcal{M}} \alpha_{\nu_m}^m R_m \right) I_\nu , \quad \nu \in \Lambda , \\ A_{\nu\mu} &= \beta_{\max(\nu_m,\mu_m)}^m I_\nu^* R_m I_\mu , \quad \nu,\mu \in \Lambda , \quad \nu - \mu = \pm\epsilon_m , \end{aligned} \tag{2.29}$$

and $A_{\nu\mu} = 0$ otherwise. ⌟

Similarly, for $f \in L^2_\pi(\Gamma; V^*)$, we have $(T^* \otimes \mathrm{id}_{V^*})f = (f_\nu)_{\nu\in\Lambda} \in \ell^2(\Lambda; V)$ for

$$f(y) = \sum_{\nu\in\Lambda} f_\nu P_\nu(y) , \quad f_\nu = \int_\Gamma f(y)\overline{P_\nu(y)}\,\mathrm{d}\pi(y) \in V^* . \tag{2.30}$$

Restricting f_ν to W_ν, *i.e.* replacing f_ν by $I_\nu^* f_\nu$, defines a vector $\boldsymbol{f} = (f_\nu)_{\nu\in\Lambda} \in \mathcal{S}^*$.

Theorem 2.15. *The Galerkin solution $\bar{\boldsymbol{u}} = (\bar u_\mu)_{\mu\in\Lambda}$ is the unique solution in $\mathcal{S}$ of*

$$D\bar u_\nu + \sum_{m\in\mathcal{M}} R_m(\beta_{\nu_m+1}^m \bar u_{\nu+\epsilon_m} + \alpha_{\nu_m}^m \bar u_\nu + \beta_{\nu_m}^m \bar u_{\nu-\epsilon_m}) = f_\nu \qquad \text{in } W_\nu \quad \forall \nu \in \Lambda . \tag{2.31}$$

Proof. The assertion is a direct consequence of Proposition 2.9, using the structure of $\boldsymbol{A}$ from Lemma 2.13 as described in Remark 2.14, and the representation (2.30) of f. □

For any $\nu \in \Lambda$, equation (2.31) only holds in W_ν, *i.e.*

$$\left\langle D\bar u_\nu + \sum_{m\in\mathcal{M}} R_m(\beta_{\nu_m+1}^m \bar u_{\nu+\epsilon_m} + \alpha_{\nu_m}^m \bar u_\nu + \beta_{\nu_m}^m \bar u_{\nu-\epsilon_m}), w \right\rangle = \langle f_\nu, w\rangle \qquad \forall w \in W_\nu . \tag{2.32}$$

Corollary 2.16. *For any $f \in L^2_\pi(\Gamma; V^*)$, $u \in L^2_\pi(\Gamma; V)$ solves* (1.19) *if and only if $u = (T\otimes\mathrm{id}_V)\boldsymbol{u}$, $\boldsymbol{u} = (u_\mu)_{\mu\in\Lambda} \in \ell^2(\Lambda; V)$, with*

$$Du_\nu + \sum_{m\in\mathcal{M}} R_m(\beta_{\nu_m+1}^m u_{\nu+\epsilon_m} + \alpha_{\nu_m}^m u_\nu + \beta_{\nu_m}^m u_{\nu-\epsilon_m}) = f_\nu \qquad \forall \nu \in \Lambda . \tag{2.33}$$

Proof. The assertion follows from Theorem 2.15 with $W_\nu = V$ for all $\nu \in \Lambda$. □

3 Algorithmic Aspects

3.1 Sparsity of the Operator Matrix

We estimate the number of nonzero entries in $\boldsymbol{A}$, interpreted as a bi-infinite operator matrix as in Remark 2.14. For any $\Xi \subset \Lambda$, let

$$\mathcal{N}(\Xi) := \left\{ \{\mu, \nu\} \ ; \ \mu, \nu \in \Xi \, , \ \left|\mu - \nu\right| = 1 \right\} \, . \tag{3.1}$$

Thus $\{\mu, \nu\} \in \mathcal{N}(\Xi)$ if and only if $\nu = \mu \pm \epsilon_m$. We call such indices *neighbors*. Furthermore, we call a set $\Xi \subset \Lambda$ *monotonic* if for any $\mu \in \Xi$ and any $\nu \in \Lambda$, $\nu_m \leq \mu_m$ for all $m \in \mathbb{N}$ implies $\nu \in \Xi$. The average length of indices in Ξ,

$$\bar{\lambda}(\Xi) := \frac{1}{\#\Xi} \sum_{\mu \in \Xi} \#\operatorname{supp} \mu \, , \tag{3.2}$$

provides a bound for the size of $\mathcal{N}(\Xi)$ compared to the size of Ξ.

Lemma 3.1. *For any finite $\Xi \subset \Lambda$,*

$$\#\mathcal{N}(\Xi) \leq \bar{\lambda}(\Xi) \, \#\Xi \, . \tag{3.3}$$

Equality holds if and only if Ξ is monotonic.

Proof. We assume initially that Ξ is monotonic. Then $\mu \in \Xi$ has the neighbors $\nu = \mu - \epsilon_m$ for all $m \in \operatorname{supp} \mu$. All neighbor pairs in Ξ are of this form since if $\nu = \mu + \epsilon_m \in \Xi$ for some $m \in \mathbb{N}$, then $\mu = \nu - \epsilon_m$ and $m \in \operatorname{supp} \nu$. Consequently,

$$\#\mathcal{N}(\Xi) = \sum_{\mu \in \Xi} \#\operatorname{supp} \mu = \bar{\lambda}(\Xi) \, \#\Xi \, .$$

If Ξ is not monotonic, then there is a $\mu \in \Xi$ and an $m \in \operatorname{supp} \mu$ such that $\nu = \mu - \epsilon_m$ is not in Ξ. Therefore,

$$\#\mathcal{N}(\Xi) < \sum_{\mu \in \Xi} \#\operatorname{supp} \mu = \bar{\lambda}(\Xi) \, \#\Xi \, . \qquad \square$$

Proposition 3.2. *If $W_\nu = \{0\}$ for all $\nu \in \Lambda \setminus \Xi$, then $A_{\nu\mu} \neq 0$ for no more than $(1 + 2\bar{\lambda}(\Xi))\#\Xi$ pairs $(\nu, \mu) \in \Lambda \times \Lambda$.*

Proof. By (2.29), $A_{\nu\mu} \neq 0$ only if $W_\nu \neq \{0\}$, $W_\mu \neq \{0\}$ and $\left|\mu - \nu\right| \leq 1$. Then the assertion follows from Lemma 3.1. $\square$

Remark 3.3. The average index length $\bar{\lambda}(\Xi)$ is generally small compared to $\#\Xi$. For certain monotonic sets Ξ, [BAS10, Corollary 4.9] estimates the maximal index length by $\log \#\Xi$. See also [Git11a] for a numerical study, which suggests logarithmic growth also for adaptively constructed sets Ξ. $\lrcorner$

3.2 Approximation of the Galerkin Solution

Let $\mathcal{S} \subset \ell^2(\Lambda; V)$ be as in (2.19), with $\Xi := \{\nu \in \Lambda\,;\, W_\nu \neq \{0\}\}$ finite. Even if it is possible to perform operation in V exactly, it is generally infeasible to compute the solution $\bar{\boldsymbol{u}}$ of (2.22).

The source term f may not be accessible. We assume the availability of a routine

$$\texttt{InRHS}_f[\Xi, \epsilon] \mapsto \tilde{f} \tag{3.4}$$

which, for any $\epsilon > 0$, computes an approximation $\tilde{f} \in \mathcal{S}^*$ of f satisfying

$$\left\|f - \tilde{f}\right\|_{\mathcal{S}^*} = \sup_{\boldsymbol{w} \in \mathcal{S}\setminus\{0\}} \frac{\left|\left\langle f - \tilde{f}, \boldsymbol{w}\right\rangle\right|}{\|\boldsymbol{w}\|_{\ell^2(\Lambda;V)}} \leq \epsilon\,. \tag{3.5}$$

Iterative solvers for (2.22) require a routine for evaluating $\boldsymbol{A}\boldsymbol{v}$ for any $\boldsymbol{v} \in \mathcal{S}$. Such a method is provided by $\texttt{InApply}_{\boldsymbol{A}}$. Due to the sparsity of $\boldsymbol{A}$, we are able to compute these products efficiently.

$\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v] \mapsto z$

---

forall $\nu \in \Xi$ **do** $z_\nu \longleftarrow A_{\nu\nu} v_\nu$
forall $\mu \in \Xi$ **do**
  forall $\nu \in \Xi$, $\nu = \mu - \epsilon_m$, $m \in \operatorname{supp} \mu$ **do**
    $z_\nu \longleftarrow z_\nu + \beta^m_{\mu_m} I^*_\nu R_m I_\mu v_\mu$
  forall $\nu \in \Xi$, $\nu = \mu + \epsilon_m$, $m \in \mathbb{N}$ **do**
    $z_\nu \longleftarrow z_\nu + \beta^m_{\nu_m} I^*_\nu R_m I_\mu v_\mu$

Remark 3.4. In the first line of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$, the diagonal components $A_{\nu\nu}$ of $\boldsymbol{A}$ are applied to the coefficients of v. These are given by an infinite series in (2.29). If all of the distributions π_m, $m \in \mathbb{N}$, are symmetric, then $A_{\nu\nu} = D$ for all $\nu \in \Xi$ by Remark 2.12. More generally, we assume that the operators $A_{\nu\nu}$ are available, and can be accessed without computing the infinite sum in (2.29). For example, in the setting of Section 1.1,

$$A_{\nu\nu} = I^*_\nu A_0 \left(\bar{a} + \sum_{m \in \mathcal{M}} \alpha^m_{\nu_m} a_m\right) I_\nu\,, \tag{3.6}$$

and thus this reduces to evaluating a series to construct the coefficient in the parametric operator. ⌟

Proposition 3.5. *The routine* $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ *computes* $\boldsymbol{A}\boldsymbol{v}$ *using one application of* $A_{\nu\nu}$ *for each* $\nu \in \Xi$ *and a total of no more than* $2\bar{\lambda}(\Xi)\#\Xi$ *applications of* R_m, $m \in \mathbb{N}$, *for any* $\boldsymbol{v} \in \mathcal{S}$.

Proof. It follows from (2.29) that $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ does indeed compute $\boldsymbol{A}\boldsymbol{v}$. The multiplications $A_{\nu\nu} v_\nu$ appear in the first line of the algorithm. The total number of subsequent products $R_m v_\mu$ is bounded by $2\#\mathcal{N}(\Xi)$. Thus the assertion follows using Lemma 3.1. □

We assume that an iterative method

$$\mathtt{PCG}_{\boldsymbol{A}}[\mathcal{S}, \tilde{f}, \tilde{u}_0, \epsilon] \mapsto \tilde{u} \tag{3.7}$$

is available which, starting from the initial approximation $\tilde{u}_0 \in \mathcal{S}$, computes $\tilde{u} \in \mathcal{S}$ satisfying

$$\|\tilde{u} - \tilde{u}^*\|_{\boldsymbol{A}} \leq \frac{\epsilon}{\sqrt{1-\gamma}}\,, \tag{3.8}$$

where $\tilde{u}^* := A^{-1}\tilde{f}$. Such a method would call the function $\mathtt{InApply}_{\boldsymbol{A}}$ to evaluate the application of the operator A to a $v \in \mathcal{S}$. A realization of $\mathtt{PCG}_{\boldsymbol{A}}$ by a preconditioned conjugate gradient iteration is provided in Section 3.3, see Proposition 3.9.

The method $\mathtt{Galerkin}_{\boldsymbol{A},f}$ combines $\mathtt{PCG}_{\boldsymbol{A}}$ with $\mathtt{InRHS}_f$ to approximate $\bar{u}$ with an ensured error bound in the norm $\|\cdot\|_{\boldsymbol{A}}$.

$\mathtt{Galerkin}_{\boldsymbol{A},f}[\mathcal{S}, \tilde{u}_0, \epsilon, \vartheta, \gamma] \mapsto \bar{u}_\epsilon$

$\epsilon_f \longleftarrow \vartheta\sqrt{1-\gamma}\left\|D^{-1}\right\|_{V^*\to V}^{-1/2}\epsilon$
$\epsilon_\downarrow \longleftarrow (1-\vartheta)\sqrt{1-\gamma}\,\epsilon$
$\tilde{f} \longleftarrow \mathtt{InRHS}_f[\mathcal{S}, \epsilon_f]$
$\bar{u}_\epsilon \longleftarrow \mathtt{PCG}_{\boldsymbol{A}}[\mathcal{S}, \tilde{f}, \tilde{u}_0, \epsilon_\downarrow]$

Proposition 3.6. *For any $\tilde{u}_0 \in \mathcal{S}$, $\epsilon > 0$ and $\vartheta \in (0,1)$, a call of $\mathtt{Galerkin}_{\boldsymbol{A},f}[\mathcal{S}, \tilde{u}_0, \epsilon, \vartheta, \gamma]$ computes $\bar{u}_\epsilon \in \mathcal{S}$ with*

$$\|\bar{u}_\epsilon - \bar{u}\|_{\boldsymbol{A}} \leq \epsilon\,. \tag{3.9}$$

If f is available, $\vartheta = 0$ is admissible.

Proof. Due to the assumption (3.5), $\left\|f - \tilde{f}\right\|_{\mathcal{S}^*} \leq \epsilon_f$. Since $\bar{u} = A^{-1}f$ and $\tilde{u}^* = A^{-1}\tilde{f}$, Proposition 2.10 implies

$$\begin{aligned}\|\bar{u} - \tilde{u}^*\|_{\boldsymbol{A}}^2 &= \left\|A^{-1}(f - \tilde{f})\right\|_{\boldsymbol{A}}^2 = \left\langle f - \tilde{f}, A^{-1}(f - \tilde{f})\right\rangle \\ &\leq \frac{1}{1-\gamma}\left\langle f - \tilde{f}, \boldsymbol{D}^{-1}(f - \tilde{f})\right\rangle \leq \frac{1}{1-\gamma}\left\|D^{-1}\right\|\epsilon_f^2 = (\vartheta\epsilon)^2\,.\end{aligned}$$

Furthermore, (3.8) implies

$$\|\bar{u}_\epsilon - \tilde{u}^*\|_{\boldsymbol{A}} \leq \frac{\epsilon_\downarrow}{\sqrt{1-\gamma}} = (1-\vartheta)\epsilon\,.$$

The assertion follows by triangle inequality. □

3.3 Conjugate Gradient Iteration

We use the preconditioned conjugate gradient method with preconditioner $\boldsymbol{D}^{-1}$ to approximate the Galerkin projection $\bar{\boldsymbol{u}}$ onto $\mathcal{S}$.

Theorem 3.7. *The conjugate gradient iteration for $\boldsymbol{A}\tilde{\boldsymbol{u}}^* = \tilde{\boldsymbol{f}}$, $\tilde{\boldsymbol{f}} \in \mathcal{W}^*$ with initial approximation $\tilde{\boldsymbol{u}}_0 \in \mathcal{S}$ and preconditioner $\boldsymbol{D}^{-1}$ constructs $\tilde{\boldsymbol{u}}_i \in \mathcal{S}$ satisfying*

$$\|\tilde{\boldsymbol{u}}_i - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} \leq 2\frac{q^i}{1+q^{2i}} \|\tilde{\boldsymbol{u}}_0 - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} \ , \quad q = \frac{\gamma}{1+\sqrt{1-\gamma^2}} \ , \tag{3.10}$$

for all $i \in \mathbb{N}_0$.

Proof. The assertion follows from [Hac91, Satz 9.4.14], which also holds in separable Hilbert spaces, with

$$q = \frac{\sqrt{1+\gamma} - \sqrt{1-\gamma}}{\sqrt{1+\gamma} + \sqrt{1-\gamma}} = \frac{\gamma}{1+\sqrt{1-\gamma^2}} \ ,$$

and using (2.23) from Proposition 2.10. □

$\mathtt{PCG}_{\boldsymbol{A}}[\mathcal{S}, \tilde{\boldsymbol{f}}, \tilde{\boldsymbol{u}}_0, \epsilon] \mapsto \tilde{\boldsymbol{u}}$

$\boldsymbol{r}^0 = (r_\nu^0)_{\nu\in\Xi} \longleftarrow \tilde{\boldsymbol{f}} - \mathtt{InApply}_{\boldsymbol{A}}[\mathcal{S}, \tilde{\boldsymbol{u}}_0]$
$\boldsymbol{s}^0 = (s_\nu^0)_{\nu\in\Xi} \longleftarrow (D_\nu^{-1} r_\nu^0)_{\nu\in\Xi}$
$\boldsymbol{v}^0 \longleftarrow \boldsymbol{s}^0$
$\eta_0 \longleftarrow \left\langle \boldsymbol{r}^0, \boldsymbol{s}^0 \right\rangle_{\ell^2(\Xi;V)}$
for $i \in \mathbb{N}$ **do**
  if $\eta_{i-1} \leq \epsilon^2$ **then**
    return $\tilde{\boldsymbol{u}} = \tilde{\boldsymbol{u}}_{i-1}$
  $\boldsymbol{z} \longleftarrow \mathtt{InApply}_{\boldsymbol{A}}[\mathcal{S}, \boldsymbol{v}^{i-1}]$
  $\alpha \longleftarrow \left\langle \boldsymbol{z}, \boldsymbol{v}^{i-1} \right\rangle_{\ell^2(\Xi;V)}$
  $\tilde{\boldsymbol{u}}_i \longleftarrow \tilde{\boldsymbol{u}}_{i-1} + \frac{\eta_{i-1}}{\alpha} \boldsymbol{v}^{i-1}$
  $\boldsymbol{r}^i \longleftarrow \boldsymbol{r}^{i-1} - \frac{\eta_{i-1}}{\alpha} \boldsymbol{z}$
  $\boldsymbol{s}^i = (s_\nu^i)_{\nu\in\Xi} \longleftarrow (D_\nu^{-1} r_\nu^i)_{\nu\in\Xi}$
  $\eta_i \longleftarrow \left\langle \boldsymbol{r}^i, \boldsymbol{s}^i \right\rangle_{\ell^2(\Xi;V)}$
  $\boldsymbol{v}^i \longleftarrow \boldsymbol{s}^i + \frac{\eta_i}{\eta_{i-1}} \boldsymbol{v}^{i-1}$

A version of the preconditioned conjugate gradient method is given in $\mathtt{PCG}_{\boldsymbol{A}}$. It uses a termination criterion based on the following norm equivalence. We abbreviate $D_\nu := I_\nu^* D I_\nu \colon W_\nu \to W_\nu^*$.

Lemma 3.8. *For all $i \in \mathbb{N}_0$,*

$$\frac{1}{1+\gamma}\eta_i \leq \|\tilde{\boldsymbol{u}}_i - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}}^2 \leq \frac{1}{1-\gamma}\eta_i \ , \tag{3.11}$$

where $\tilde{\boldsymbol{u}}^* \in \mathcal{S}$ is the solution of $\boldsymbol{A}\tilde{\boldsymbol{u}}^* = \tilde{\boldsymbol{f}}$.

Proof. By definition, $\eta_i = \left\langle \boldsymbol{r}^i, \boldsymbol{s}^i \right\rangle_{\ell^2(\Xi;V)}$, $\boldsymbol{r}^i = \tilde{\boldsymbol{f}} - \boldsymbol{A}\tilde{\boldsymbol{u}}_i$ and $\boldsymbol{s}^i = \boldsymbol{D}^{-1}\boldsymbol{r}^i$. We abbreviate $\boldsymbol{e}^i := \tilde{\boldsymbol{u}}_i - \tilde{\boldsymbol{u}}^*$. The assertion follows from Proposition 2.10 since

$$\left\|\boldsymbol{e}^i\right\|_{\boldsymbol{A}}^2 = \left\langle \boldsymbol{A}\boldsymbol{e}^i, \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{e}^i \right\rangle_{\ell^2(\Xi;V)}$$

and

$$\eta_i = \left\langle \boldsymbol{A}\boldsymbol{e}^i, \boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{e}^i \right\rangle_{\ell^2(\Xi;V)} . \qquad \square$$

Proposition 3.9. *The method* $\texttt{PCG}_{\boldsymbol{A}}[\mathcal{S}, \tilde{\boldsymbol{f}}, \tilde{\boldsymbol{u}}_0, \epsilon]$ *terminates and returns* $\tilde{\boldsymbol{u}}$ *satisfying*

$$\|\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} \leq \frac{\epsilon}{\sqrt{1-\gamma}} . \tag{3.12}$$

At most

$$1 + \left\lfloor \frac{\log\left(2\epsilon^{-1}\sqrt{1+\gamma}\,\|\tilde{\boldsymbol{u}}_0 - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}}\right)}{\log q} \right\rfloor \tag{3.13}$$

iterations are performed, with q from (3.10)*. Each iteration contains $\#\Xi$ evaluations of D^{-1}, one application of $A_{\nu\nu}$ for each $\nu \in \Xi$, and a total of no more than $2\bar{\lambda}(\Xi)\#\Xi$ applications of R_m, $m \in \mathbb{N}$.*

Proof. Equation (3.12) follows from Lemma 3.8. Let the final iterate be $\tilde{\boldsymbol{u}} = \tilde{\boldsymbol{u}}_N$. Then provided $N \geq 1$, using the other inequality in Lemma 3.8,

$$\|\tilde{\boldsymbol{u}}_{N-1} - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} \geq \frac{1}{\sqrt{1+\gamma}}\,\eta_{N-1} \geq \frac{\epsilon}{\sqrt{1+\gamma}} .$$

By Theorem 3.7,

$$\epsilon \leq \sqrt{1+\gamma}\,\|\tilde{\boldsymbol{u}}_{N-1} - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} \leq 2\sqrt{1+\gamma}q^{N-1}\,\|\tilde{\boldsymbol{u}}_0 - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}} .$$

Solving for N leads to

$$N - 1 \leq \frac{\log\left(2\epsilon^{-1}\sqrt{1+\gamma}\,\|\tilde{\boldsymbol{u}}_0 - \tilde{\boldsymbol{u}}^*\|_{\boldsymbol{A}}\right)}{\log q} .$$

The final part of the assertion is a consequence of Proposition 3.5. $\square$

3.4 Finite Element Approximation

The above estimates only consider the size and structure of the set of active indices $\Xi \subset \Lambda$, and not the finite element spaces W_ν, $\nu \in \Xi$. We complete the analysis by studying the cost of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$, taking into account the varying cost of operations in different W_ν.

Let $n_\nu := \#W_\nu$. We assume that the parametric operator A_0 is local, as in the example from Section 1.1. We assume further that the spaces W_ν are equipped with local bases, which is usually the case for finite elements. Then the computational cost of applying $A_{\nu\mu}$ is $O(\max(n_\nu, n_\mu))$. Consequently, the total cost of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ is on the order of

$$\sum_{\nu\in\Xi} n_\nu + 2 \sum_{\{\nu,\mu\}\in\mathcal{N}(\Xi)} \max(n_\nu, n_\mu) \ . \tag{3.14}$$

We make the natural assumption that if $\mu \in \Xi$, then for all $m \in \operatorname{supp}\mu$, $\nu := \mu - \epsilon_m \in \Xi$, and $n_\nu \geq n_\mu$. Then the second sum in (3.14) is bounded by

$$\sum_{\mu\in\Xi} \sum_{m\in\operatorname{supp}\mu} n_{\mu-\epsilon_m} = \sum_{\nu\in\Xi} q_\nu n_\nu \ , \quad q_\nu := \#\{\mu \in \Xi \ ; \ \nu = \mu - \epsilon_m \ , \ m \in \mathbb{N}\} \ . \tag{3.15}$$

Consequently, the computational cost of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ is at most

$$\sum_{\nu\in\Xi} (1 + 2q_\nu) n_\nu \ . \tag{3.16}$$

This sum can be estimated further in various ways. For example, as in Proposition 3.2,

$$\sum_{\nu\in\Xi} (1 + 2q_\nu) n_\nu \leq (1 + 2\bar{\lambda}(\Xi))\#\Xi \max_{\nu\in\Xi} n_\nu \ . \tag{3.17}$$

Alternatively, we have

$$\sum_{\nu\in\Xi} (1 + 2q_\nu) n_\nu \leq (1 + 2\max_{\nu\in\Xi} q_\nu) \sum_{\nu\in\Xi} n_\nu \ . \tag{3.18}$$

Example 3.10. Let $\Xi = \{0, \epsilon_1, \ldots, \epsilon_M\}$, with $n_{\epsilon_m} \leq n_0$ for all m. Then $q_0 = M$, and $q_{\epsilon_m} = 0$ for all m. The compuational cost of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ is on the order of

$$\sum_{\nu\in\Xi} (1 + 2q_\nu) n_\nu = (1 + 2M) n_0 + \sum_{m=1}^{M} n_{\epsilon_m} \ ,$$

and (3.17) gives the fairly sharp estimate

$$(1 + 2\bar{\lambda}(\Xi))\#\Xi \max_{\nu\in\Xi} n_\nu = (1 + 3M) n_0 \ ,$$

whereas (3.18) provides

$$(1 + 2\max_{\nu\in\Xi} q_\nu) \sum_{\nu\in\Xi} n_\nu = (1 + 2M) \left(n_0 + \sum_{m=1}^{M} n_{\epsilon_m} \right) \ ,$$

which is much coarser if n_{ϵ_m} are large. ⌟

Example 3.11. Let $\mathcal{M} = \{1\}$, and $\Xi = \{0, \ldots, k\}$, *i.e.* and $\nu \in \Xi$ are just integers less than k. Then $q_\nu = 1$ for $\nu \in \{0, \ldots, k-1\}$, and $q_k = 0$. Therefore, assuming that $(n_\nu)_{\nu=0}^k$ is decreasing, the compuational cost of $\texttt{InApply}_{\boldsymbol{A}}[\mathcal{S}, v]$ is on the order of

$$\sum_{\nu\in\Xi}(1 + 2q_\nu)n_\nu = n_k + 3\sum_{\nu=0}^{k-1} n_\nu \ .$$

The estimate (3.18) provides the bound

$$(1 + 2\max_{\nu\in\Xi} q_\nu)\sum_{\nu\in\Xi} n_\nu = 3\sum_{\nu=0}^{k} n_\nu \ ,$$

which is much sharper than

$$(1 + 2\bar{\lambda}(\Xi))\#\Xi \max_{\nu\in\Xi} n_\nu = (1 + 3k)n_0$$

from (3.17). ⌟

Remark 3.12. We have not taken into account the possibility that the functions a_m from Section 1.1 have local supports. Suppose that $(a_{\ell,i})_{\ell,i}$ forms a doubly indexed multilevel basis, with $\ell \in \mathbb{N}_0$ and i ranging from 1 to M_ℓ. We assume that the size of the support of $a_{\ell,i}$ is on the order of $1/M_\ell$. For any $\nu \in \Lambda$, let $[\nu]$ consist of all $\mu \in \Lambda$ that differ from ν only by a permutation of the i indices. For simplicity, we suppose that $n_\nu = n_{[\nu]}$ is constant on each of these equivalence classes. Then $A_{[\nu],[\mu]}$ is a finite section of the operator $\boldsymbol{A}$, and the above estimates apply verbatim with ν replaced by $[\nu]$, provided that the support of every $a_{\ell,i}$ is resolved on each finite element mesh. We will not go into details here. ⌟

References

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

[Bau02] Heinz Bauer. *Wahrscheinlichkeitstheorie.* de Gruyter Lehrbuch. [de Gruyter Textbook]. Walter de Gruyter & Co., Berlin, fifth edition, 2002.

[Ber96] Christian Berg. Moment problems and polynomial approximation. *Ann. Fac. Sci. Toulouse Math. (6)*, (Special issue):9–32, 1996. 100 ans après Th.-J. Stieltjes.

[Bie09] M. Bieri. A sparse composite collocation finite element method for elliptic sPDEs. Technical Report 2009-8, Seminar for Applied Mathematics, ETH Zürich, 2009. Submitted.

[BNT07] Ivo 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.

[Bob05] A. Bobrowski. *Functional analysis for probability and stochastic processes*. Cambridge University Press, Cambridge, 2005. An introduction.

[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 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 DeVore. Adaptive wavelet methods for elliptic operator equations: convergence rates. *Math. Comp.*, 70(233):27–75 (electronic), 2001.

[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.

[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.

[EMSU10] Oliver G. Ernst, Antje Mugler, Hans-Jörg Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. Technical Report 60, DFG Schwerpunktprogramm 1324, 2010.

[Fre71] Géza Freud. *Orthogonal Polynomials*. Pergamon Press, Oxford, 1971.

[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.

[Hac91] Wolfgang Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*, volume 69 of *Leitfäden der Angewandten Mathematik und Mechanik [Guides to Applied Mathematics and Mechanics]*. B. G. Teubner, Stuttgart, 1991. Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks].

[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.

[Pro05] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[Rie23] Friedrich Riesz. Über eine Verallgemeinerung der Parsevalschen Formel. *Math. Z.*, 18(1):117–124, 1923.

[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.

[XH05] Dongbin Xiu and Jan S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139 (electronic), 2005.

[Xiu07] Dongbin Xiu. Efficient collocational approach for parametric uncertainty analysis. *Commun. Comput. Phys.*, 2(2):293–309, 2007.

[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
11-10	*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
11-06	*G.M. Coclite, K.H. Karlsen. S. Mishra and N.H. Risebro* A hyperbolic-elliptic model of two-phase flow in porous media - Existence of entropy solutions
11-05	*U.S. Fjordholm, S. Mishra and E. Tadmor* Entropy stable ENO scheme
11-04	*M. Ganesh, S.C. Hawkins and R. Hiptmair* Convergence analysis with parameter estimates for a reduced basis acoustic scattering T-matrix method
11-03	*O. Reichmann* Optimal space-time adaptive wavelet methods for degenerate parabolic PDEs
11-02	*S. Mishra, Ch. Schwab and J. Šukys* Multi-level Monte Carlo finite volume methods for nonlinear systems of conservation laws in multi-dimensions
11-01	*V. Wheatley, R. Jeltsch and H. Kumar* Spectral performance of RKDG methods
10-49	*R. Jeltsch and H. Kumar* Three dimensional plasma arc simulation using resistive MHD
10-48	*M. Swärd and S. Mishra* Entropy stable schemes for initial-boundary-value conservation laws
10-47	*F.G. Fuchs, A.D. McMurry, S. Mishra and K. Waagan* Simulating waves in the upper solar atmosphere with Surya: A well-balanced high-order finite volume code