Discrete least squares polynomial approximation with random evaluations – application to PDEs with random parameters

F. Nobile\textsuperscript{1}, G. Migliorati\textsuperscript{1}, R. Tempone\textsuperscript{2}

\textsuperscript{1}CSQI-MATHICSE, EPFL, Switzerland
\textsuperscript{2}AMCS and SRI-UQ Center, KAUST, Saudi Arabia

Acknowledgments: A. Cohen, A. Chkifa (UPMC - Paris VI), E. von Schwerin (KTH),

Advances in UQ Methods, Algorithms and Applications
KAUST, January 6-9, 2015
Outline

1. Introduction – PDEs with random parameters
2. Stochastic polynomial approximation
3. Discrete projection using random evaluations
   - Stability
   - Convergence results in expectation and probability
   - The case of noisy observations
4. Conclusions
Outline

1. Introduction – PDEs with random parameters

2. Stochastic polynomial approximation

3. Discrete projection using random evaluations
   - Stability
   - Convergence results in expectation and probability
   - The case of noisy observations

4. Conclusions
Consider a deterministic PDE model

\[ \text{find } u : \quad \mathcal{L}(y)(u) = F \quad \text{in } D \subset \mathbb{R}^d \]  

(1)

with suitable boundary / initial conditions.

- The operator \( \mathcal{L}(y) \) depends on a vector of \( N \) parameters:
  \( y = (y_1, \ldots, y_N) \in \mathbb{R}^N \) (\( N = \infty \) when dealing with distributed fields).

- Often in applications the parameters \( y \) are not perfectly known or are intrinsically variable. Examples are:
  - subsurface modeling: porous media flows; seismic waves; basin evolutions; ...
  - modeling of living tissues: mechanical response; growth models;
  - material science: properties of composite materials

- Probabilistic approach: \( y \) is a random vector with probability density function \( \rho : \Gamma \rightarrow \mathbb{R}_+ \).
UQ for deterministic PDE models

- Consider a deterministic PDE model
  
  \[
  \text{find } u : \quad \mathcal{L}(\mathbf{y})(u) = \mathcal{F} \quad \text{in } D \subset \mathbb{R}^d
  \]

  with suitable boundary / initial conditions.

- The operator \( \mathcal{L}(\mathbf{y}) \) depends on a vector of \( N \) parameters:
  \( \mathbf{y} = (y_1, \ldots, y_N) \in \mathbb{R}^N \) (\( N = \infty \) when dealing with distributed fields).

- Often in applications the parameters \( \mathbf{y} \) are not perfectly known or are intrinsically variable. Examples are:
  - subsurface modeling: porous media flows; seismic waves; basin evolutions; ...
  - modeling of living tissues: mechanical response; growth models;
  - material science: properties of composite materials

- Probabilistic approach: \( \mathbf{y} \) is a random vector with probability density function \( \rho : \Gamma \rightarrow \mathbb{R}_+ \).
UQ for deterministic PDE models

- Consider a deterministic PDE model

\[
\text{find } u : \quad \mathcal{L}(y)(u) = \mathcal{F} \quad \text{in } D \subset \mathbb{R}^d
\]  

(1)

with suitable boundary / initial conditions.

- The operator \( \mathcal{L}(y) \) depends on a vector of \( N \) parameters:
  \( y = (y_1, \ldots, y_N) \in \mathbb{R}^N \) (\( N = \infty \) when dealing with distributed fields).

- Often in applications the parameters \( y \) are not perfectly known or are intrinsically variable. Examples are:
  - subsurface modeling: porous media flows; seismic waves; basin evolutions; ...
  - modeling of living tissues: mechanical response; growth models;
  - material science: properties of composite materials

- Probabilistic approach: \( y \) is a random vector with probability density function \( \rho : \Gamma \rightarrow \mathbb{R}_+ \).
Introduction – PDEs with random parameters

UQ for deterministic PDE models

- Consider a deterministic PDE model

\[
\text{find } u : \quad \mathcal{L}(y)(u) = F \quad \text{in } D \subset \mathbb{R}^d
\]  

with suitable boundary / initial conditions.

- The operator \( \mathcal{L}(y) \) depends on a vector of \( N \) parameters:
\[
y = (y_1, \ldots, y_N) \in \mathbb{R}^N \quad (N = \infty \text{ when dealing with distributed fields}).
\]

- Often in applications the parameters \( y \) are not perfectly known or are intrinsically variable. Examples are:
  - subsurface modeling: porous media flows; seismic waves; basin evolutions; ...
  - modeling of living tissues: mechanical response; growth models;
  - material science: properties of composite materials

- **Probabilistic approach:** \( y \) is a random vector with probability density function \( \rho : \Gamma \to \mathbb{R}_+ \).
UQ for deterministic PDE models

Assumption: \( \forall y \in \Gamma \) the problem admits a unique solution \( u \in V \) in a Hilbert space \( V \). Moreover,

\[
\| u(y) \|_V \leq C(y) \| F \|_V
\]

- Then, the PDE (1) induces a map \( u = u(y) : \Gamma \rightarrow V \).
- if \( C(y) \in L^p_\rho(\Gamma) \), then \( u \in L^p_\rho(\Gamma, V) \).

Goals: Compute statistics of the solution

pointwise Expected value: \( \bar{u}(x) = \mathbb{E}[u(x, \cdot)], x \in D \)
pointwise Variance: \( \text{Var}[u](x) = \mathbb{E}[(u(x, \cdot) - \bar{u}(x))^2](x) \)
two points corr.: \( \text{Cov}_u(x_1, x_2) = \mathbb{E}[(u(x_1, \cdot) - \bar{u}(x_1))(u(x_2, \cdot) - \bar{u}(x_2))] \)
or of specific Quantities of Interest \( Q(u) : V \rightarrow \mathbb{R} \). Then \( \varphi(y) = Q(u(y)) \)
is a real-valued function of the random vector \( y \) and we would like to approximate its law.
UQ for deterministic PDE models

**Assumption:** \( \forall y \in \Gamma \) the problem admits a unique solution \( u \in V \) in a Hilbert space \( V \). Moreover,

\[
\|u(y)\|_V \leq C(y)\|F\|_V,
\]

- Then, the PDE (1) induces a map \( u = u(y) : \Gamma \rightarrow V \).
- if \( C(y) \in L^p_{\rho}(\Gamma) \), then \( u \in L^p_{\rho}(\Gamma, V) \).

**Goals:** Compute **statistics of the solution**

- **pointwise Expected value:** \( \bar{u}(x) = \mathbb{E}[u(x, \cdot)] \), \( x \in D \)
- **pointwise Variance:** \( \text{Var}[u](x) = \mathbb{E}[(u(x, \cdot) - \bar{u}(x))^2](x) \)
- **two points corr.:** \( \text{Cov}_u(x_1, x_2) = \mathbb{E}[(u(x_1, \cdot) - \bar{u}(x_1))(u(x_2, \cdot) - \bar{u}(x_2))] \)

or of specific **Quantities of Interest** \( Q(u) : V \rightarrow \mathbb{R} \). Then \( \varphi(y) = Q(u(y)) \) is a real-valued function of the random vector \( y \) and we would like to approximate its law.
UQ for deterministic PDE models

Assumption: \( \forall y \in \Gamma \) the problem admits a unique solution \( u \in V \) in a Hilbert space \( V \). Moreover,

\[
\|u(y)\|_V \leq C(y)\|F\|_V
\]

- Then, the PDE (1) induces a map \( u = u(y) : \Gamma \to V \).
- if \( C(y) \in L^p_\rho(\Gamma) \), then \( u \in L^p_\rho(\Gamma, V) \).

Goals: Compute statistics of the solution

- pointwise Expected value: \( \bar{u}(x) = \mathbb{E}[u(x, \cdot)], x \in D \)
- pointwise Variance: \( Var[u](x) = \mathbb{E}[((u(x, \cdot) - \bar{u}(x))^2](x) \)
- two points corr.: \( Cov_u(x_1, x_2) = \mathbb{E}[(u(x_1, \cdot) - \bar{u}(x_1))(u(x_2, \cdot) - \bar{u}(x_2))] \)

or of specific Quantities of Interest \( Q(u) : V \to \mathbb{R} \). Then \( \varphi(y) = Q(u(y)) \) is a real-valued function of the random vector \( y \) and we would like to approximate its law.
Example: Elliptic PDE with random coefficients

\[
\begin{aligned}
- \text{div}(a(y, x) \nabla u(y, x)) &= f(x) \quad x \in D, \ y \in \Gamma, \\
u(y, x) &= 0 \quad x \in \partial D, \ y \in \Gamma
\end{aligned}
\]

with \( a_{\text{min}}(y) = \inf_{x \in D} a(y, x) > 0 \) for all \( y \in \Gamma \) and \( f \in L^2(D) \). Then

\[
\forall y \in \Gamma, \quad u(y) \in V \equiv H^1_0(D), \quad \text{and} \quad \| u(y) \|_V \leq \frac{C_P}{a_{\text{min}}(y)} \| f \|_{L^2(D)}.
\]

**Inclusions problem**

- \( y \) describes the conductivity in each inclusion
- \( a(y, x) = a_0 + \sum_{n=N}^{\infty} y_n \mathbb{1}_{D_n}(x) \)

**Random fields problem**

- \( a(y, x) \) is a random field, e.g. lognormal:
  \( a(y, x) = e^{\gamma(y, x)} \) with \( \gamma \)
- expanded e.g. in Karhunen-Loève series

\[
\gamma(y, x) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} y_n b_n(x), \quad y_n \sim N(0, 1) \text{ i.i.d.}
\]
Outline

1. Introduction – PDEs with random parameters

2. Stochastic polynomial approximation

3. Discrete projection using random evaluations
   - Stability
   - Convergence results in expectation and probability
   - The case of noisy observations

4. Conclusions
Stochastic multivariate polynomial approximation

- The parameter-to-solution map \( u(y) : \Gamma \to V \) is often smooth (even analytic for the elliptic diffusion model). It is therefore sound to approximate it by global multivariate polynomials.

- Let \( \Lambda \subset \mathbb{N}^N \) be an index set of cardinality \(|\Lambda| = M\), and consider the multivariate polynomial space

\[
P_\Lambda(\Gamma) = \text{span} \left\{ \prod_{n=1}^{N} y_n^{p_n}, \text{ with } p = (p_1, \ldots, p_N) \in \Lambda \right\}
\]

We seek an approximation \( P_\Lambda u \in P_\Lambda(\Gamma) \otimes V \).

- The optimal choice of \( \Lambda \) depends heavily on the problem at hand and the “structure” of the map \( u(y) \).

**Definition.** An index set \( \Lambda \) is downward closed if

\[
p \in \Lambda \text{ and } q \leq p \implies q \in \Lambda
\]
Stochastic multivariate polynomial approximation

- The parameter-to-solution map $u(y) : \Gamma \to V$ is often smooth (even analytic for the elliptic diffusion model). It is therefore sound to approximate it by global multivariate polynomials.

- Let $\Lambda \subset \mathbb{N}^N$ be an index set of cardinality $|\Lambda| = M$, and consider the multivariate polynomial space

$$\mathbb{P}_\Lambda(\Gamma) = \text{span}\left\{\prod_{n=1}^{N} y_n^{p_n}, \quad \text{with } \mathbf{p} = (p_1, \ldots, p_N) \in \Lambda\right\}$$

We seek an approximation $P_\Lambda u \in \mathbb{P}_\Lambda(\Gamma) \otimes V$.

- The optimal choice of $\Lambda$ depends heavily on the problem at hand and the “structure” of the map $u(y)$.

**Definition.** An index set $\Lambda$ is downward closed if

$$p \in \Lambda \quad \text{and} \quad q \leq p \quad \implies \quad q \in \Lambda$$
Stochastic multivariate polynomial approximation

- The parameter-to-solution map $u(y) : \Gamma \rightarrow V$ is often smooth (even analytic for the elliptic diffusion model). It is therefore sound to approximate it by global multivariate polynomials.

- Let $\Lambda \subset \mathbb{N}^N$ be an index set of cardinality $|\Lambda| = M$, and consider the multivariate polynomial space

$$P_\Lambda(\Gamma) = \text{span} \left\{ \prod_{n=1}^N y_n^{p_n}, \quad \text{with } p = (p_1, \ldots, p_N) \in \Lambda \right\}$$

We seek an approximation $P_\Lambda u \in P_\Lambda(\Gamma) \otimes V$.

- The optimal choice of $\Lambda$ depends heavily on the problem at hand and the “structure” of the map $u(y)$.

**Definition.** An index set $\Lambda$ is downward closed if

$$p \in \Lambda \text{ and } q \leq p \implies q \in \Lambda$$
Stochastic multivariate polynomial approximation

- The parameter-to-solution map \( u(\mathbf{y}) : \Gamma \to V \) is often smooth (even analytic for the elliptic diffusion model). It is therefore sound to approximate it by global multivariate polynomials.

- Let \( \Lambda \subset \mathbb{N}^N \) be an index set of cardinality \(|\Lambda| = M\), and consider the multivariate polynomial space

\[
P_\Lambda(\Gamma) = \text{span} \left\{ \prod_{n=1}^{N} y_n^{p_n}, \text{ with } \mathbf{p} = (p_1, \ldots, p_N) \in \Lambda \right\}
\]

We seek an approximation \( P_\Lambda u \in P_\Lambda(\Gamma) \otimes V \).

- The optimal choice of \( \Lambda \) depends heavily on the problem at hand and the “structure” of the map \( u(\mathbf{y}) \).

**Definition.** An index set \( \Lambda \) is downward closed if

\[
\mathbf{p} \in \Lambda \text{ and } \mathbf{q} \leq \mathbf{p} \implies \mathbf{q} \in \Lambda
\]
Common choices of polynomial spaces

- **tensor product (TP)**
  \[ \Lambda(w) = \{ \mathbf{p} : \max_n p_n \leq w \} \]

- **total degree (TD)**
  \[ \Lambda(w) = \{ \mathbf{p} : \sum_{n=1}^{N} p_n \leq w \} \]

- **hyperbolic cross (HC)**
  \[ \Lambda(w) = \{ \mathbf{p} : \prod_{n=1}^{N} (p_n + 1) \leq w + 1 \} \]

- Anisotropic versions are also possible.
- All these index sets are all downward closed.
Common choices of polynomial spaces

- tensor product (TP)
  \[ \Lambda(w) = \{ p : \max_n p_n \leq w \} \]

- total degree (TD)
  \[ \Lambda(w) = \{ p : \sum_{n=1}^{N} p_n \leq w \} \]

- hyperbolic cross (HC)
  \[ \Lambda(w) = \{ p : \prod_{n=1}^{N} (p_n + 1) \leq w + 1 \} \]

- Anisotropic versions are also possible.
- All these index sets are all downward closed.
Outline

1 Introduction – PDEs with random parameters

2 Stochastic polynomial approximation

3 Discrete projection using random evaluations
   - Stability
   - Convergence results in expectation and probability
   - The case of noisy observations

4 Conclusions
Discrete $L^2$ projection using random evaluations


1. Generate $M$ random i.i.d. samples $y(\omega_k) \sim \rho(y) dy$, $k = 1, \ldots, M$

2. Compute the corresponding solutions $u^{(k)} = u(y(\omega_k))$

3. Find the discrete least squares approximation $\Pi^M_{\Lambda} u \in \mathbb{P}_\Lambda(\Gamma) \otimes V$:

$$\Pi^M_{\Lambda} u = \operatorname*{argmin}_{v \in \mathbb{P}_\Lambda(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} \| u^{(k)} - v(y(\omega_k)) \|^2_V$$

For a quantity of interest $\varphi(y) = Q(u(y))$ this reads simply

$$\Pi^M_{\Lambda} \varphi = \operatorname*{argmin}_{v \in \mathbb{P}_\Lambda(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} | \varphi^{(k)} - v(y(\omega_k)) |^2$$
Discrete $L^2$ projection using random evaluations

(see e.g. [Hosder-Walters et al. 2010, Blatman-Sudret 2008, Burkardt-Eldred
Migliorati et al 2011-2014])

1. Generate $M$ random i.i.d. samples $y(\omega_k) \sim \rho(y)dy$, $k = 1, \ldots, M$
2. Compute the corresponding solutions $u^{(k)} = u(y(\omega_k))$
3. Find the discrete least squares approximation $\Pi^M u \in \mathcal{P}_\Lambda(\Gamma) \otimes V$:

$$
\Pi^M u = \arg\min_{v \in \mathcal{P}_\Lambda(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} \| u^{(k)} - v(y(\omega_k)) \|^2_V
$$

For a quantity of interest $\varphi(y) = Q(u(y))$ this reads simply

$$
\Pi^M \varphi = \arg\min_{v \in \mathcal{P}_\Lambda(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} | \varphi^{(k)} - v(y(\omega_k)) |^2
$$
Notation

- **continuous norm**: \[ \| v \|_{L^2_{\rho}(\Gamma, \mathcal{V})}^2 = \int_{\Gamma} \| v(y) \|_{\mathcal{V}_\rho(y)}^2 d\mathbf{y} \]

- **discrete norm**: \[ \| v \|_{M, \mathcal{V}}^2 = \frac{1}{M} \sum_{i=1}^{M} \| v(y(\omega_i)) \|_{\mathcal{V}}^2 \]

- \( \{ \psi_p \}_{p \in \Lambda} \): orthonormal basis of \( P_\Lambda(\Gamma) \)
Algebraic formulation (for Q.o.I.)

Design matrix: \( D \in \mathbb{R}^{|\Lambda| \times M}, \quad D_{ip} = \psi_p(y(\omega_i)), \quad p \in \Lambda, 1 \leq i \leq M. \)

Then, expanding \( \Pi_M^\Lambda \varphi \) onto the basis: \( \Pi_M^\Lambda \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y), \) and setting \( (\varphi)_i = \varphi(y(\omega_i)) \), the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

Equivalent reformulation:

\[
Gc = J\varphi, \quad \text{with} \quad G = \frac{1}{M} D^T D, \quad J = \frac{1}{M} D^T
\]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \|G^{-1}\|. \)
- It holds:

\[
\|G\| = \sup_{v \in \mathcal{P}^\Lambda(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}, \quad \|G^{-1}\| = \sup_{v \in \mathcal{P}^\Lambda(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}.
\]
Discrete projection using random evaluations

Algebraic formulation (for Q.o.I.)

**Design matrix:** \( D \in \mathbb{R}^{\Lambda \times M} \), \( D_{ip} = \psi_p(y(\omega_i)) \), \( p \in \Lambda, 1 \leq i \leq M \).

Then, expanding \( \Pi_M \varphi \) onto the basis: \( \Pi_M \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y) \), and setting \( (\varphi)_i = \varphi(y(\omega_i)) \), the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

Equivalent reformulation:

\[
Gc = J \varphi, \quad \text{with} \quad G = \frac{1}{M} D^T D, \quad J = \frac{1}{M} D^T
\]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \|G^{-1}\| \).
- It holds

\[
\|G\| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\|v\|_{L^2, \mathbb{R}}^2}{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}, \quad \|G^{-1}\| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{L^2, \mathbb{R}}^2}.
\]
Algebraic formulation (for Q.o.I.)

Design matrix: \( D \in \mathbb{R}^{|\Lambda| \times M} \), \( D_{ip} = \psi_p(y(\omega_i)), \ p \in \Lambda, 1 \leq i \leq M \).

Then, expanding \( \Pi_M^\Lambda \varphi \) onto the basis: \( \Pi_M^\Lambda \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y) \), and setting \( (\varphi)_i = \varphi(y(\omega_i)) \), the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

Equivalent reformulation:

\[
Gc = J\varphi, \quad \text{with} \quad G = \frac{1}{M}D^T D, \quad J = \frac{1}{M}D^T
\]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \|G^{-1}\| \).
- It holds

\[
\|G\| = \sup_{v \in p_\Lambda(\Gamma)} \frac{\|v\|_{M,\mathbb{R}}^2}{\|v\|_{L_p^2(\Gamma,\mathbb{R})}^2}, \quad \|G^{-1}\| = \sup_{v \in p_\Lambda(\Gamma)} \frac{\|v\|_{L_p^2(\Gamma,\mathbb{R})}^2}{\|v\|_{M,\mathbb{R}}^2}.
\]
Algebraic formulation (for Q.o.I.)

Design matrix: \( D \in \mathbb{R}^{\Lambda | \times M}, \quad D_{ip} = \psi_p(y(\omega_i)), \quad p \in \Lambda, 1 \leq i \leq M. \)

Then, expanding \( \Pi_M \varphi \) onto the basis: \( \Pi_M \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y), \) and setting \( (\varphi)_i = \varphi(y(\omega_i)), \) the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

Equivalent reformulation:

\[
Gc = J \varphi, \quad \text{with} \quad G = \frac{1}{M} D^T D, \quad J = \frac{1}{M} D^T
\]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \|G^{-1}\|. \)
- It holds

\[
\|G\| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{M, \mathbb{R}}^2}, \quad \|G^{-1}\| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{M, \mathbb{R}}^2}
\]
Algebraic formulation (for Q.o.I.)

Design matrix: \( D \in \mathbb{R}^{\Lambda \times M}, \quad D_{ip} = \psi_p(y(\omega_i)), \quad p \in \Lambda, 1 \leq i \leq M. \)

Then, expanding \( \prod_\Lambda^M \varphi \) onto the basis: \( \prod_\Lambda^M \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y), \) and setting \( (\varphi)_i = \varphi(y(\omega_i)), \) the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

Equivalent reformulation:

\[ Gc = J\varphi, \quad \text{with} \quad G = \frac{1}{M} D^T D, \quad J = \frac{1}{M} D^T \]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \| G^{-1} \| \).
- It holds

\[
\| G \| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\| v \|^2_{M,\mathbb{R}}}{\| v \|^2_{L^2(\Gamma,\mathbb{R})}}, \quad \| G^{-1} \| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma)} \frac{\| v \|^2_{L^2(\Gamma,\mathbb{R})}}{\| v \|^2_{M,\mathbb{R}}},
\]

\( \| \cdot \| \) denotes a norm.
Algebraic formulation (for Q.o.I.)

**Design matrix:** \( D \in \mathbb{R}^{\Lambda \times M}, \ D_{ip} = \psi_p(y(\omega_i)), \ p \in \Lambda, 1 \leq i \leq M. \)

Then, expanding \( \prod_{\Lambda}^M \varphi \) onto the basis: \( \prod_{\Lambda}^M \varphi(y) = \sum_{p \in \Lambda} c_p \psi_p(y) \), and setting \( (\varphi)_i = \varphi(y(\omega_i)) \), the vector \( c = \{c_p\}_p \) of Fourier coefficients satisfies the normal equations

\[
(D^T D)c = D^T \varphi.
\]

**Equivalent reformulation:**

\[
Gc = J\varphi, \quad \text{with} \quad G = \frac{1}{M} D^T D, \quad J = \frac{1}{M} D^T
\]

- \( G \) is symmetric and (semi)-positive definite.
- The stability of the discrete least squares is related to \( \|G^{-1}\| \).
- It holds

\[
\|G\| = \sup_{v \in \mathcal{P}_{\Lambda}(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{M, \mathbb{R}}^2}, \quad \|G^{-1}\| = \sup_{v \in \mathcal{P}_{\Lambda}(\Gamma)} \frac{\|v\|_{L^2(\Gamma, \mathbb{R})}^2}{\|v\|_{M, \mathbb{R}}^2}
\]

F. Nobile (EPFL) Discrete least squares for random PDEs UQAW 2015, KAUST
More on the matrix $G$

Recall $G = \frac{1}{M} D^T D$. Hence

$$G = \frac{1}{M} \sum_{i=1}^{M} G^{(i)}$$

with $G^{(i)}_{pq} = \psi_q(y(\omega_i))\psi_p(y(\omega_i))$

**Remarks:**

- The matrices $G^{(i)}$, $i = 1, \ldots, M$ are i.i.d.
- $\mathbb{E}[G^{(i)}] = I$. Indeed $\mathbb{E}[G^{(i)}_{pq}] = \mathbb{E}[\psi_p \psi_q] = \delta_{pq}$.
- Define

$$K(\Lambda) = \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) = \sup_{v \in \mathcal{P}_\Lambda} \frac{\|v\|_{L^\infty(\Gamma)}^2}{\|v\|_{L^2_{\rho}(\Gamma)}^2}$$

Note that $K(\Lambda)$ does not depend on the orthonormal basis chosen.

Then $G^{(i)}$ satisfies a uniform bound

$$\|G^{(i)}\| = \sup_{v \in \mathcal{P}_\Lambda(\Gamma) \otimes V} \frac{\|v(y(\omega_i))\|_V^2}{\|v\|_{L^p(\Gamma, V)}^2} \leq K(\Lambda)$$
More on the matrix $G$

Recall $G = \frac{1}{M} D^T D$. Hence

$$G = \frac{1}{M} \sum_{i=1}^{M} G^{(i)}, \quad \text{with} \quad G_{pq}^{(i)} = \psi_q(y(\omega_i)) \psi_p(y(\omega_i))$$

Remarks:

- The matrices $G^{(i)}$, $i = 1, \ldots, M$ are i.i.d.
- $E[G^{(i)}] = I$. Indeed $E[G_{pq}^{(i)}] = E[\psi_p \psi_q] = \delta_{pq}$.
- Define

$$K(\Lambda) = \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) = \sup_{v \in P_\Lambda} \frac{\|v\|^2_{L^\infty(\Gamma)}}{\|v\|^2_{L^2(\rho(\Gamma))}}$$

Note that $K(\Lambda)$ does not depend on the orthonormal basis chosen.

Then $G^{(i)}$ satisfies a uniform bound

$$\|G^{(i)}\| = \sup_{v \in P_\Lambda(\Gamma) \otimes V} \frac{\|v(y(\omega_i))\|^2_V}{\|v\|^2_{L^2(\rho(\Gamma), V)}} \leq K(\Lambda)$$
More on the matrix $G$

Recall $G = \frac{1}{M} D^T D$. Hence

$$G = \frac{1}{M} \sum_{i=1}^{M} G^{(i)} , \quad \text{with} \quad G^{(i)}_{pq} = \psi_q(y(\omega_i))\psi_p(y(\omega_i))$$

Remarks:
- The matrices $G^{(i)}$, $i = 1, \ldots, M$ are i.i.d.
- $\mathbb{E}[G^{(i)}] = I$. Indeed $\mathbb{E}[G^{(i)}_{pq}] = \mathbb{E}[\psi_p \psi_q] = \delta_{pq}$.
- Define

$$K(\Lambda) = \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) = \sup_{\nu \in \mathcal{P}_\Lambda} \frac{||\nu||^2_{L^\infty(\Gamma)}}{||\nu||^2_{L^2(\Gamma)}}$$

Note that $K(\Lambda)$ does not depend on the orthonormal basis chosen.

Then $G^{(i)}$ satisfies a uniform bound

$$\|G^{(i)}\| = \sup_{\nu \in \mathcal{P}_\Lambda(\Gamma) \otimes V} \frac{||\nu(\omega_i)||^2_{V}}{||\nu||^2_{L^2(\Gamma,V)}} \leq K(\Lambda)$$
More on the matrix $G$

Recall $G = \frac{1}{M} D^T D$. Hence

$$G = \frac{1}{M} \sum_{i=1}^{M} G^{(i)}$$

with $G^{(i)}_{pq} = \psi_q(y(\omega_i)) \psi_p(y(\omega_i))$

**Remarks:**

- The matrices $G^{(i)}$, $i = 1, \ldots, M$ are i.i.d.
- $\mathbb{E}[G^{(i)}] = I$. Indeed $\mathbb{E}[G^{(i)}_{pq}] = \mathbb{E}[\psi_p \psi_q] = \delta_{pq}$.
- Define

$$K(\Lambda) = \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) = \sup_{v \in P_\Lambda} \frac{\|v\|_{L^\infty(\Gamma)}^2}{\|v\|_{L^2(\rho; \Gamma)}}$$

Note that $K(\Lambda)$ does not depend on the orthonormal basis chosen.

Then $G^{(i)}$ satisfies a **uniform bound**

$$\|G^{(i)}\| = \sup_{v \in P_\Lambda(\Gamma) \otimes V} \frac{\|v(y(\omega_i))\|_V^2}{\|v\|_{L^\rho(\Gamma, V)}^2} \leq K(\Lambda)$$
More on the matrix $G$

$G$ is the sample average of i.i.d. positive definite and uniformly bounded random matrices and $\mathbb{E}[G] = I$.

Results on $\|G - I\| = \|G - \mathbb{E}[G]\|$ can be obtained from concentration of measure inequalities for sums of independent matrices.

**Goal**: obtain conditions under which $\|G - I\| \leq \delta$ for some $\delta \in (0, 1)$.
Observe that this implies a norm equivalence on $\mathbb{P}_\Lambda(\Gamma) \otimes V$

$$(1 - \delta)\|v\|_{L_2(\Gamma; V)}^2 \leq \|v\|_{M, V}^2 \leq (1 + \delta)\|v\|_{L_2(\Gamma; V)}^2, \quad \forall v \in \mathbb{P}_\Lambda \otimes V$$

(analogous to the Restricted Isometry Property (RIP) in compressed sensing, see [Candès-Tao ’06, Rahout-Ward ’12, ...])
More on the matrix $G$

$G$ is the sample average of i.i.d. positive definite and uniformly bounded random matrices and $\mathbb{E}[G] = I$.

Results on $\|G - I\| = \|G - \mathbb{E}[G]\|$ can be obtained from concentration of measure inequalities for sums of independent matrices.

**Goal**: obtain conditions under which $\|G - I\| \leq \delta$ for some $\delta \in (0, 1)$.

Observe that this implies a norm equivalence on $\mathbb{P}_\Lambda(\Gamma) \otimes V$

$$(1 - \delta)\|v\|_{L^2_\rho(\Gamma; V)}^2 \leq \|v\|_{M,V}^2 \leq (1 + \delta)\|v\|_{L^2_\rho(\Gamma; V)}^2, \quad \forall v \in \mathbb{P}_\Lambda \otimes V$$

(analogous to the Restricted Isometry Property (RIP) in compressed sensing, see [Candès-Tao ’06, Rahout-Ward ’12, ...])
More on the matrix $G$

$G$ is the sample average of i.i.d. positive definite and uniformly bounded random matrices and $\mathbb{E}[G] = I$.

Results on $\|G - I\| = \|G - \mathbb{E}[G]\|$ can be obtained from concentration of measure inequalities for sums of independent matrices.

**Goal**: obtain conditions under which $\|G - I\| \leq \delta$ for some $\delta \in (0, 1)$.

Observe that this implies a norm equivalence on $\mathbb{P}_\Lambda(\Gamma) \otimes V$

$$(1 - \delta)\|v\|_{L^2_\rho(\Gamma;V)}^2 \leq \|v\|_{M,V}^2, \quad v \leq (1 + \delta)\|v\|_{L^2_\rho(\Gamma;V)}^2, \quad \forall v \in \mathbb{P}_\Lambda \otimes V$$

(analogous to the Restricted Isometry Property (RIP) in compressed sensing, see [Candès-Tao ’06, Rahout-Ward ’12, ...])
Matrix Chernoff’s bound (for i.i.d. random matrices) [J. Tropp, FoCM 2011]

Let $X_1, \ldots, X_M \in \mathbb{R}^{d \times d}$ be i.i.d. spd random matrices s.t. $\lambda_{\text{max}}(X_i) \leq R$ almost surely. Let $\mu_{\text{min}} = \lambda_{\text{min}}(\mathbb{E}[X_i]), \mu_{\text{max}} = \lambda_{\text{max}}(\mathbb{E}[X_i])$ and $\bar{X} = \frac{1}{M} \sum_{i=1}^{M} X_i$. Then

$$P(\lambda_{\text{max}}(\bar{X}) \geq (1 + \delta)\mu_{\text{max}}) \leq d \exp \left\{ - \frac{M\mu_{\text{max}}\bar{\beta}_\delta}{R} \right\}, \quad \delta \geq 0$$

$$P(\lambda_{\text{min}}(\bar{X}) \leq (1 - \delta)\mu_{\text{min}}) \leq d \exp \left\{ - \frac{M\mu_{\text{max}}\beta_\delta}{R} \right\}, \quad \delta \in [0, 1],$$

with $\bar{\beta}_\delta = (1 + \delta) \log(1 + \delta) - \delta$ and $\beta_\delta = \delta + (1 - \delta) \log(1 - \delta)$.
Concentration of measure result

Theorem [Cohen-Davenport-Leviatan '13]

Introduce the event

$$\Omega^M_+(\delta) := \{ \| G - I \| \leq \delta \} = \{ (1 - \delta) \| v \|_{L^2_\rho(\Gamma, v)}^2 \leq \| v \|_{M, v}^2 \leq (1 + \delta) \| v \|_{L^2_\rho(\Gamma, v)}^2, \forall v \in \mathbb{P}_\Lambda \}. $$

For any $\delta, \gamma > 0$ and $M$ satisfying

$$K(\Lambda) \leq \frac{\beta_\delta}{1 + \gamma \log M}, \quad \beta_\delta = \delta + (1 - \delta) \log(1 - \delta) \quad (2)$$

we have that $P(\Omega^M_+(\delta)) \geq 1 - 2M^{-\gamma}$.

Hence on $\Omega^M_+(\delta)$ the random discrete $L^2$ projection is stable and

$$\text{cond}(D^T D) = \text{cond}(G) \leq \frac{1 + \delta}{1 - \delta}.$$
Concentration of measure result

**Theorem** [Cohen-Davenport-Leviatan ’13]

Introduce the event

\[ \Omega^M_+(\delta) := \{ \| G - I \| \leq \delta \} \]

\[ = \{ (1 - \delta)\| v \|_{L^2_{\rho}(\Gamma, \nu)}^2 \leq \| v \|_{M, \nu}^2 \leq (1 + \delta)\| v \|_{L^2_{\rho}(\Gamma, \nu)}^2, \forall v \in \mathbb{P}_\Lambda \}. \]

For any \( \delta, \gamma > 0 \) and \( M \) satisfying

\[ K(\Lambda) \leq \frac{\beta_\delta}{1 + \gamma \log M}, \quad \beta_\delta = \delta + (1 - \delta) \log(1 - \delta) \quad (2) \]

we have that \( P(\Omega^M_+(\delta)) \geq 1 - 2M^{-\gamma} \).

Hence on \( \Omega^M_+(\delta) \) the random discrete \( L^2 \) projection is stable and

\[ \text{cond}(D^T D) = \text{cond}(G) \leq \frac{1 + \delta}{1 - \delta} \]
Convergence in Probability

From the stability of the random projection one can derive optimality results either in expectation or probability.

**Theorem [Chkifa-Cohen-Miglierati-N.-Tempone '14], [Miglierati-N.-Tempone '15]**

For any $\alpha, \delta \in (0, 1)$, under the condition $\frac{M}{\log M + \log(2/\alpha)} \geq \frac{K(\Lambda)}{\beta \delta}$, it holds with probability greater that $1 - \alpha$

$$\| u - \Pi_M^M u \|_{L^2(\Gamma, \mathcal{V})} \leq (1 + \sqrt{\frac{1}{1 - \delta}}) \inf_{\nu \in \mathcal{P}_\Lambda \otimes \mathcal{V}} \| u - \nu \|_{L^\infty(\Gamma, \mathcal{V})}$$

**Proof:** Under the above condition $P(\Omega_M^M(\delta)) \geq 1 - \alpha$. Given any draw in $\Omega_M^M(\delta)$, we have for any $\nu \in \mathcal{P}_\Lambda$

$$\| u - \Pi_M^M u \|_{L^2(\Gamma, \mathcal{V})} \leq \| u - \nu \|_{L^2(\Gamma, \mathcal{V})} + \| \nu - \Pi_M^M u \|_{L^2(\Gamma, \mathcal{V})}$$

$$\leq \| u - \nu \|_{L^2(\Gamma, \mathcal{V})} + \sqrt{(1 - \delta)^{-1}} \| \nu - \Pi_M^M u \|_{M, \mathcal{V}}$$

$$\leq \| u - \nu \|_{L^2(\Gamma, \mathcal{V})} + \sqrt{(1 - \delta)^{-1}} \| u - \nu \|_{M, \mathcal{V}}$$

$$\leq (1 + \sqrt{(1 - \delta)^{-1}}) \| u - \nu \|_{L^\infty(\Gamma, \mathcal{V})}$$
Convergence in Probability

From the stability of the random projection one can derive optimality results either in expectation or probability.

**Theorem** [Chkifa-Cohen-Migliorati-N.-Tempone ’14], [Migliorati-N.-Tempone ’15]

For any \( \alpha, \delta \in (0, 1) \), under the condition
\[
\frac{M}{ \log M + \log(2/\alpha) } \geq \frac{K(\Lambda)}{\beta \delta},
\]

it holds with probability greater that \( 1 - \alpha \)

\[
\| u - \Pi^M \Lambda u \|_{L^2_{\rho}(\Gamma, V)} \leq (1 + \sqrt{\frac{1}{1 - \delta}}) \inf_{v \in \mathbb{P}_\Lambda \otimes V} \| u - v \|_{L^\infty(\Gamma, V)}.
\]

**Proof:** Under the above condition \( P(\Omega^M_+(\delta)) \geq 1 - \alpha \). Given any draw in \( \Omega^M_+(\delta) \), we have for any \( v \in \mathbb{P}_\Lambda \)

\[
\| u - \Pi^M \Lambda u \|_{L^2_{\rho}(\Gamma, V)} \leq \| u - v \|_{L^2_{\rho}(\Gamma, V)} + \| v - \Pi^M \Lambda u \|_{L^2_{\rho}(\Gamma, V)}
\]

\[
\leq \| u - v \|_{L^2_{\rho}(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \| v - \Pi^M \Lambda u \|_{M, V}
\]

\[
\leq \| u - v \|_{L^2_{\rho}(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \| u - v \|_{M, V}
\]

\[
\leq (1 + \sqrt{(1 - \delta)^{-1}}) \| u - v \|_{L^\infty(\Gamma, V)}.
\]
Convergence in Probability

From the stability of the random projection one can derive optimality results either in expectation or probability.

Theorem [Chkifa-Cohen-Migliorati-N.-Tempone ‘14], [Migliorati-N.-Tempone ‘15]

For any $\alpha, \delta \in (0, 1)$, under the condition $\frac{M}{\log M + \log(2/\alpha)} \geq \frac{K(\Lambda)}{\beta \delta}$, it holds with probability greater that $1 - \alpha$

$$\|u - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)} \leq (1 + \sqrt{\frac{1}{1 - \delta}}) \inf_{\nu \in \mathcal{P}_\Lambda \otimes V} \|u - \nu\|_{L^\infty(\Gamma, V)}$$

Proof: Under the above condition $P(\Omega^M_+(\delta)) \geq 1 - \alpha$. Given any draw in $\Omega^M_+(\delta)$, we have for any $\nu \in \mathcal{P}_\Lambda$

$$\|u - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)} \leq \|u - \nu\|_{L^2_\rho(\Gamma, V)} + \|\nu - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)}$$

$$\leq \|u - \nu\|_{L^2_\rho(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \|\nu - \Pi^M_{\Lambda} u\|_{M, V}$$

$$\leq \|u - \nu\|_{L^2_\rho(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \|u - \nu\|_{M, V}$$

$$\leq (1 + \sqrt{(1 - \delta)^{-1}}) \|u - \nu\|_{L^\infty(\Gamma, V)}.$$
Convergence in Probability

From the stability of the random projection one can derive optimality results either in expectation or probability.

Theorem [Chkifa-Cohen-Migliorati-N.-Tempone '14], [Migliorati-N.-Tempone '15]

For any \( \alpha, \delta \in (0, 1) \), under the condition \( \frac{M}{\log M + \log(2/\alpha)} \geq \frac{K(\Lambda)}{\beta \delta} \), it holds with probability greater that \( 1 - \alpha \)

\[
\| u - \Pi^M_\Lambda u \|_{L^2_\rho(\Gamma, \mathcal{V})} \leq (1 + \sqrt{\frac{1}{1 - \delta}}) \inf_{v \in \mathcal{P}_\Lambda \otimes \mathcal{V}} \| u - v \|_{L^\infty(\Gamma, \mathcal{V})}
\]

Proof: Under the above condition \( P(\Omega^M_+(\delta)) \geq 1 - \alpha \). Given any draw in \( \Omega^M_+(\delta) \), we have for any \( v \in \mathcal{P}_\Lambda \)

\[
\| u - \Pi^M_\Lambda u \|_{L^2_\rho(\Gamma, \mathcal{V})} \leq \| u - v \|_{L^2_\rho(\Gamma, \mathcal{V})} + \| v - \Pi^M_\Lambda u \|_{L^2_\rho(\Gamma, \mathcal{V})}
\]

\[
\leq \| u - v \|_{L^2_\rho(\Gamma, \mathcal{V})} + \sqrt{(1 - \delta)^{-1}} \| v - \Pi^M_\Lambda u \|_{M, \mathcal{V}}
\]

\[
\leq (1 + \sqrt{(1 - \delta)^{-1}}) \| u - v \|_{L^\infty(\Gamma, \mathcal{V})}
\]

\[
\leq (1 + \sqrt{(1 - \delta)^{-1}}) \| u - v \|_{L^\infty(\Gamma, \mathcal{V})}
\]
Convergence in Probability

From the stability of the random projection one can derive optimality results either in expectation or probability.

Theorem [Chkifa-Cohen-Migliorati-N.-Tempone '14], [Migliorati-N.-Tempone '15]

For any $\alpha, \delta \in (0, 1)$, under the condition $\frac{M \log M + \log(2/\alpha)}{\beta \delta} \geq K(\Lambda)$, it holds with probability greater than $1 - \alpha$

$$\|u - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)} \leq (1 + \sqrt{\frac{1}{1 - \delta}}) \inf_{v \in \mathbb{P}_{\Lambda} \otimes V} \|u - v\|_{L^\infty(\Gamma, V)}$$

Proof: Under the above condition $P(\Omega^M_+(\delta)) \geq 1 - \alpha$. Given any draw in $\Omega^M_+(\delta)$, we have for any $v \in \mathbb{P}_{\Lambda}$

$$\|u - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)} \leq \|u - v\|_{L^2_\rho(\Gamma, V)} + \|v - \Pi^M_{\Lambda} u\|_{L^2_\rho(\Gamma, V)}$$

$$\leq \|u - v\|_{L^2_\rho(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \|v - \Pi^M_{\Lambda} u\|_{M, V}$$

$$\leq \|u - v\|_{L^2_\rho(\Gamma, V)} + \sqrt{(1 - \delta)^{-1}} \|u - v\|_{M, V}$$

$$\leq (1 + \sqrt{(1 - \delta)^{-1}}) \|u - v\|_{L^\infty(\Gamma, V)}.$$
Convergence in expectation

assume \( \| u \|_{L^\infty(\Gamma, \mathcal{V})} \leq \tau \) and define the truncation operator

\[
T_\tau : V \rightarrow V, \quad T_\tau(v) = \begin{cases} v & \text{if } \| v \|_V \leq \tau \\ \frac{\tau}{\| v \|_V} v, & \text{if } \| v \|_V > \tau \end{cases}
\]

Theorem [Cohen-Davenport-Leviatan ’13], [Chkifa-Cohen-Migliorati-N.-Tempone ’14]

For any \( \delta \in (0, 1) \) and any \( \gamma > 0 \), under the condition \( \frac{M}{\log M} \geq (1 + \gamma) \frac{K(\Lambda)}{\beta_\delta} \), it holds

\[
\mathbb{E}(\| u - T_\tau \circ \Pi^M u \|_{L^2_\rho(\Gamma, \mathcal{V})}^2) \leq C \inf_{v \in \mathcal{P} \otimes \mathcal{V}} \| u - v \|_{L^2_\rho(\Gamma, \mathcal{V})}^2 + 8 \tau^2 M^{-\gamma}
\]

with \( C = 1 + \frac{4\beta_\delta}{(1+\gamma)\log M} \xrightarrow{M \to \infty} 1. \)
Case of noisy observations

Let us consider the case of a QoI $\varphi(y) = Q(u(y))$ and noisy observations

$$\varphi^{(k)} = \varphi(y_k) + \eta_k$$

with $\eta_k$ i.i.d. and

$$\mathbb{E}[\eta_k|y_k] = \bar{\eta}(y_k) \in L^2_\rho(\Gamma) \quad \text{(offset)}$$

$$\sup_{y_k \in \Gamma} \text{Var}(\eta_k|y_k) = \sigma^2 < \infty \quad \text{(variance)}$$

The offset could model any deterministic source of error due e.g. to numerical discretization.

The fluctuations $\tilde{\eta}_k = \eta_k - \bar{\eta}(y_k)$ model random measurement errors.

We will also consider the case of bounded noise

$$|\tilde{\eta}_k| \leq \tilde{\eta}_{\text{max}}, \quad \|\tilde{\eta}\|_{L^\infty(\Gamma)} < \infty.$$
Case of noisy observations

Let us consider the case of a QoI $\varphi(y) = Q(u(y))$ and noisy observations

$$\varphi^{(k)} = \varphi(y_k) + \eta_k$$

with $\eta_k$ i.i.d. and

$$\mathbb{E}[\eta_k|y_k] = \bar{\eta}(y_k) \in L^2_\rho(\Gamma)$$  \hspace{1cm} (offset)

$$\sup_{y_k \in \Gamma} \text{Var}(\eta_k|y_k) = \sigma^2 < \infty$$  \hspace{1cm} (variance)

The offset could model any deterministic source of error due e.g. to numerical discretization.

The fluctuations $\tilde{\eta}_k = \eta_k - \bar{\eta}(y_k)$ model random measurement errors.

We will also consider the case of bounded noise

$$|\tilde{\eta}_k| \leq \tilde{\eta}_{\text{max}}, \quad \|\tilde{\eta}\|_{L^\infty(\Gamma)} < \infty.$$
Case of noisy observations

Let us consider the case of a QoI $\varphi(y) = Q(u(y))$ and noisy observations

$$\varphi^{(k)} = \varphi(y_k) + \eta_k$$

with $\eta_k$ i.i.d. and

$$\mathbb{E}[\eta_k | y_k] = \bar{\eta}(y_k) \in L^2_\rho(\Gamma) \quad \text{(offset)}$$

$$\sup_{y_k \in \Gamma} \text{Var}(\eta_k | y_k) = \sigma^2 < \infty \quad \text{(variance)}$$

The offset could model any deterministic source of error due e.g. to numerical discretization.

The fluctuations $\tilde{\eta}_k = \eta_k - \bar{\eta}(y_k)$ model random measurement errors.

We will also consider the case of bounded noise

$$|\tilde{\eta}_k| \leq \tilde{\eta}_{\text{max}}, \quad \|\tilde{\eta}\|_{L^\infty(\Gamma)} < \infty.$$
Convergence in expectation

Theorem [Chkifa-Cohen-Migliorati-N.-Tempone ’14], [Migliorati-N.-Tempone ’15]

Assume \( \| \varphi \|_{L^\infty(\Gamma)} \leq \tau \). For any \( \delta \in (0,1) \) and any \( \gamma > 0 \), under the condition \( \frac{M}{\log M} \geq (1 + \gamma) \frac{K(\Lambda)}{\beta_\delta} \), it holds

\[
\mathbb{E}(\| \varphi - T_\tau \circ \Pi^M_{\Lambda} \varphi \|^2_{L^2_\rho(\Gamma)}) \leq C_1 \inf_{\nu \in \mathcal{P}_\Lambda} \| \varphi - \nu \|^2_{L^2_\rho(\Gamma)}
\]

\[
= C_1 \inf_{\nu \in \mathcal{P}_\Lambda} \| \varphi - \nu \|^2_{L^2_\rho(\Gamma)} + \frac{2}{(1 - \delta)^2} \left( \frac{\#\Lambda}{M} \sigma^2 + C_2 \| \bar{\eta} \|^2_{L^2_\rho(\Gamma)} \right) + 8\tau^2 M^{-\gamma}
\]

with \( C_1, C_2 \xrightarrow{M \to \infty} 1 \).
Convergence in Probability – bounded noise

Theorem [Migliorati-N.-Tempone ’15]

In the bounded noise case, for any $\alpha, \delta \in (0, 1)$, under the condition
\[
\frac{M}{\log M + \log(3/\alpha)} \geq \frac{K(\Lambda)}{\beta \delta},
\]
it holds with probability greater that $1 - \alpha$

\[
\|\varphi - \Pi^M_\Lambda \varphi\|_{L^2_\rho(\Gamma)}^2 \leq (1 + \frac{2}{1 - \delta}) \inf_{v \in \mathcal{P}_\Lambda} \|\varphi - v\|_{L^\infty_\rho(\Gamma)}^2
\]

best approx. error in $L^\infty$

\[
+ \frac{4(1 + \delta)}{(1 - \delta)^2} \left( 2 \frac{\#\Lambda \log(3M\alpha^{-1})}{M} \tilde{\eta}_{max}^2 + \|\tilde{\eta}\|_{L^\infty_\rho(\Gamma)}^2 \right)
\]

bounded noise

noise offset
Case of uniform random variables in $[-1, 1]$

The discrete $L^2$ projection is stable and optimally convergent under the condition

$$K(\Lambda) := \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) \leq \frac{\beta_\delta}{1 + \gamma \log M}$$

where $\beta_\delta$ is defined in (2). Recall that for Legendre polynomials we have:

$$|\psi_p(y)| \leq \prod_{n=1}^N \sqrt{2p_n + 1}, \text{ for all } y \in [-1, 1]^N.$$
Case of uniform random variables in $[-1, 1]$

The discrete $L^2$ projection is stable and optimally convergent under the condition

$$K(\Lambda) := \sup_{y \in \Gamma} \left( \sum_{p \in \Lambda} |\psi_p(y)|^2 \right) \leq \frac{\beta \delta}{1 + \gamma \log M}$$

where $\beta \delta$ is defined in (2). Recall that for Legendre polynomials we have:

$$|\psi_p(y)| \leq \prod_{n=1}^{N} \sqrt{2p_n + 1}, \text{ for all } y \in [-1, 1]^N.$$

**Theorem [Chkifa-Cohen-Migliorati-Nobile-Tempone ’14]**

For any set $\Lambda \subset \mathbb{N}^N$ monotone it holds $(\# \Lambda) \leq K(\Lambda) \leq (\# \Lambda)^2$.

Hence, the discrete $L^2$ projection over $P_{\Lambda}$ is stable and optimally convergent in expectation under the (sufficient) condition

$$\frac{1 + \gamma}{\beta \delta} (\# \Lambda)^2 \leq \frac{M}{\log M}$$
Case of uniform random variables in $[-1, 1]$

- For specific index sets $\Lambda$ the condition can be improved.
- For instance for the Total Degree polynomial space of degree $w$ the bound $K(\Lambda) \leq (\#\Lambda)^2$ is very conservative

The bound for $K(\Lambda)$ heavily depends on the underlying distribution.
For instance,

Chebyshev distribution $\implies K(\Lambda) \leq \min\{(\#\Lambda)^{\frac{\log 3}{\log 2}}, 2^N \#\Lambda\}$

Beta distribution with $\theta_1, \theta_2 \in \mathbb{N}$ $\implies K(\Lambda) \leq (\#\Lambda)^2 \max\{\theta_1, \theta_2\} + 2$
Case of uniform random variables in $[-1, 1]$

- For specific index sets $\Lambda$ the condition can be improved.
- For instance for the Total Degree polynomial space of degree $w$ the bound $K(\Lambda) \leq (\#\Lambda)^2$ is very conservative.

![Graphs showing the bound for different dimensions.]

- The bound for $K(\Lambda)$ heavily depends on the underlying distribution. For instance,

  - **Chebyshev** distribution $\implies K(\Lambda) \leq \min\left\{\left(\frac{\log 3}{\log 2}\right)^{\log 2}, 2^N \#\Lambda\right\}$
  - **Beta** distribution with $\theta_1, \theta_2 \in \mathbb{N} \implies K(\Lambda) \leq (\#\Lambda)^2 \max\{\theta_1, \theta_2\} + 2$
Some numerical examples – 1D function

Condition number of $D^T D$

$M = c \cdot \#\Lambda$

$M = c \cdot (\#\Lambda)^2$
Some numerical examples – 1D function

Approximation of the meromorphic function $\phi(y) = \frac{1}{1+0.5y}$

$M = c \cdot \#\Lambda$

error with respect to polynomial degree.

$M = c \cdot (\#\Lambda)^2$
Some numerical examples – 1D function

Approximation of the meromorphic function \( \phi(y) = \frac{1}{1+0.5y} \)

c=2, \( \alpha =1 \)
c=20, \( \alpha =1 \)
c=1, \( \alpha =2 \)
c=3, \( \alpha =2 \)

error with respect to total number of sampling points.
Some numerical examples

Condition number of $D^T D$ – multiD – Total Degree poly. space

\[ M = c \cdot \# \Lambda \]

\[ M = c \cdot (\# \Lambda)^2 \]
Elliptic PDE with random inclusions

We derived the theoretical bound

$$\mathbb{E}(\|u - T_\tau \circ \Pi^M u\|_{L^2(\Gamma, H^1_0(D))}^2) \leq c_1 e^{-c_2 NM^{1+2N}}$$
Cantilever beam

- linear elasticity equations
- Young modulus uncertain in each brick:
  \[ E_i = e^{7+Y_i}, \quad \text{in } \Omega_i, \quad Y_i \sim U([-1, 1]), \text{ iid} \]
- Uncertainty analysis on maximum vertical displacement.

\[ \Gamma_{wall} \]

\[ \Omega_1 \quad \Omega_2 \quad \Omega_3 \quad \Omega_4 \quad \Omega_5 \quad \Omega_6 \quad \Omega_7 \]

\[ 2 \]

\[ 0.5 \]

\[ 1 \]

\[ 7 \]

\[ \text{Condition number Total Degree, } N=7, M=c \cdot \#\Lambda \]

\[ \text{Error QOI}_6(u), \text{ Total Degree, } N=7, M=c \cdot \#\Lambda \]

F. Nobile (EPFL)
Improvements on the quadratic relation

Improvements can be obtained by sampling from a different distribution \( \hat{\rho} \). Let us consider the weighted least squares approx.

\[
\hat{u}_{\Lambda, M} = \arg\min_{v \in \mathbb{P}_\Lambda(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} \frac{\rho(y^{(k)})}{\hat{\rho}(y^{(k)})} \| u^{(k)} - v(y^{(k)}) \|_V^2
\]

where the sample \( \{y^k\}_k \) is drawn from the distribution \( \hat{\rho}(y)dy \).

- \( \rho(y) = \hat{\rho}(y) = \) Chebyshev distribution in \([-1, 1]^N\), then the relation
  \[ M \propto \min\{2^N(\#\Lambda), (\#\Lambda)^{\log(3)/\log(2)}\} \]
  is enough to guarantee optimal convergence [Chkifa-Cohen-Migliorati-N-Tempone '14]
- \( \rho(y) = \) uniform and \( \hat{\rho}(y) = \) Chebyshev distribution in \([-1, 1]^N\), then, the relation
  \[ M \propto 2^N(\#\Lambda) \]
  guarantees optimal convergence [Rauhut-Ward '12]. However, the constant depends on \( N \) [Yan-Guo-Xiu '12].
- \( \rho(y) = \) Gaussian: still unclear. Numerically, the situation seems to be worse. Improvements suggested in [Tang-Zhou '14].
Improvements on the quadratic relation

Improvements can be obtained by sampling from a different distribution \( \hat{\rho} \). Let us consider the weighted least squares approx.

\[
\hat{u}_{\Lambda,M} = \arg\min_{v \in \mathbb{P}_\Lambda(\Gamma) \otimes \mathbb{V}} \frac{1}{M} \sum_{k=1}^{M} \frac{\rho(y^{(k)})}{\hat{\rho}(y^{(k)})} \| u^{(k)} - v(y^{(k)}) \|^2_{\mathbb{V}}
\]

where the sample \( \{y^k\}_k \) is drawn from the distribution \( \hat{\rho}(y)dy \).

- \( \rho(y) = \hat{\rho}(y) = \text{Chebyshev distribution in } [-1,1]^N \), then the relation \( M \propto \min\{2^N(\#\Lambda), (\#\Lambda)^{\log(3)}/\log(2)\} \) is enough to guarantee optimal convergence [Chkifa-Cohen-Migliorati-N.-Tempone '14]

- \( \rho(y) = \text{uniform and } \hat{\rho}(y) = \text{Chebyshev distribution in } [-1,1]^N \), then, the relation \( M \propto 2^N(\#\Lambda) \) guarantees optimal convergence [Rauhut-Ward '12]. However, the constant depends on \( N \) [Yan-Guo-Xiu '12].

- \( \rho(y) = \text{Gaussian: still unclear. Numerically, the situation seems to be worse. Improvements suggested in } [\text{Tang-Zhou '14}] \)
Improvements on the quadratic relation

Improvements can be obtained by sampling from a different distribution $\hat{\rho}$. Let us consider the weighted least squares approx.

$$\hat{u}_{\Lambda, M} = \arg\min_{v \in \mathcal{P}_\Lambda(\Gamma) \otimes \mathcal{V}} \frac{1}{M} \sum_{k=1}^{M} \frac{\rho(y^{(k)})}{\hat{\rho}(y^{(k)})} \| u^{(k)} - v(y^{(k)}) \|_V^2$$

where the sample $\{y^k\}_k$ is drawn from the distribution $\hat{\rho}(y)dy$.

- $\rho(y) = \hat{\rho}(y) = $ Chebyshev distribution in $[-1, 1]^N$, then the relation $M \propto \min \{2^N(\#\Lambda), (\#\Lambda)^{\log(3)/\log(2)}\}$ is enough to guarantee optimal convergence [Chkifa-Cohen-Migliorati-N.-Tempone '14]

- $\rho(y) =$ uniform and $\hat{\rho}(y) =$ Chebyshev distribution in $[-1, 1]^N$, then, the relation $M \propto 2^N(\#\Lambda)$ guarantees optimal convergence [Rauhut-Ward '12]. However, the constant depends on $N$ [Yan-Guo-Xiu '12].

- $\rho(y) =$ Gaussian: still unclear. Numerically, the situation seems to be worse. Improvements suggested in [Tang-Zhou '14]
Improvements on the quadratic relation

Improvements can be obtained by sampling from a different distribution $\hat{\rho}$. Let us consider the weighted least squares approx.

$$
\hat{u}_{\Lambda,M} = \arg\min_{v \in \mathbb{P}_{\Lambda}(\Gamma) \otimes V} \frac{1}{M} \sum_{k=1}^{M} \frac{\rho(y^{(k)})}{\hat{\rho}(y^{(k)})} \| u^{(k)} - v(y^{(k)}) \|_{V}^{2}
$$

where the sample $\{y^{k}\}_{k}$ is drawn from the distribution $\hat{\rho}(y)dy$.

- $\rho(y) = \hat{\rho}(y) = $ Chebyshev distribution in $[-1, 1]^{N}$, then the relation $M \propto \min\{2^{N}(\#\Lambda), (\#\Lambda)^{\log(3)/\log(2)}\}$ is enough to guarantee optimal convergence [Chkifa-Cohen-Migliorati-N.-Tempone '14]
- $\rho(y) =$ uniform and $\hat{\rho}(y) =$ Chebyshev distribution in $[-1, 1]^{N}$, then, the relation $M \propto 2^{N}(\#\Lambda)$ guarantees optimal convergence [Rauhut-Ward '12]. However, the constant depends on $N$ [Yan-Guo-Xiu '12].
- $\rho(y) =$ Gaussian: still unclear. Numerically, the situation seems to be worse. Improvements suggested in [Tang-Zhou '14]
Numerical example with Chebyshev preconditioning

Expansion in Legendre polynomials ($\rho(y) = \text{uniform}$) and samples from Chebyshev distribution ($\hat{\rho}(y) = \text{Chebyshev}$)

$$u(y) = \left(1 + \frac{0.7}{2N} \sum_{n=1}^{N} y_n\right)^{-1}$$

Condition number $\text{cond}(D^T D)$

$$M = 3 \cdot (\#\Lambda)$$

Error for $u(y)$
Adaptive construction of polynomial spaces

\( \{\Lambda_k\}_{k \geq 0} \) sequence of downward closed multi-index sets, with \( \Lambda_0 = \{0\} \). The sequence is adaptively computed by means of greedy algorithms based on the random discrete \( L^2 \) projection.

**Definitions:**

- **Margin** \( M(\Lambda) \) associated to a multi-index set \( \Lambda \):
  \[
  M(\Lambda) = \{ p : p \notin \Lambda \text{ and } \exists j > 0 : p - e_j \in \Lambda \}
  \]

- **Reduced margin** \( R(\Lambda) \) associated to a multi-index set \( \Lambda \):
  \[
  R(\Lambda) = \{ p : p \notin \Lambda \text{ and } \forall j = 1, \ldots, d : p_j \neq 0 \Rightarrow p - e_j \in \Lambda \}
  \]
Adaptive construction of polynomial spaces

\(\{\Lambda_k\}_{k \geq 0}\) sequence of downward closed multi-index sets, with \(\Lambda_0 = \{0\}\). The sequence is adaptively computed by means of greedy algorithms based on the random discrete \(L^2\) projection.

**Definitions:**

- **Margin** \(\mathcal{M}(\Lambda)\) associated to a multi-index set \(\Lambda\):

\[
\mathcal{M}(\Lambda) = \{p : p \not\in \Lambda \text{ and } \exists j > 0 : p - e_j \in \Lambda\}
\]

- **Reduced margin** \(\mathcal{R}(\Lambda)\) associated to a multi-index set \(\Lambda\):

\[
\mathcal{R}(\Lambda) = \{p : p \not\in \Lambda \text{ and } \forall j = 1, \ldots, d : p_j \neq 0 \Rightarrow p - e_j \in \Lambda\}
\]
Adaptive construction of polynomial spaces

\{\Lambda_k\}_{k \geq 0} sequence of downward closed multi-index sets, with \( \Lambda_0 = \{0\} \).

The sequence is adaptively computed by means of greedy algorithms based on the random discrete \( L^2 \) projection.

**Definitions:**

- **Margin** \( \mathcal{M}(\Lambda) \) associated to a multi-index set \( \Lambda \):

  \[ \mathcal{M}(\Lambda) = \{p : p \notin \Lambda \text{ and } \exists j > 0 : p - \mathbf{e}_j \in \Lambda\} \]

- **Reduced margin** \( \mathcal{R}(\Lambda) \) associated to a multi-index set \( \Lambda \):

  \[ \mathcal{R}(\Lambda) = \{p : p \notin \Lambda \text{ and } \forall j = 1, \ldots, d : p_j \neq 0 \Rightarrow p - \mathbf{e}_j \in \Lambda\} \]
The Dörfler marking


Given a multi-index set $\Lambda$, a subset $R \subseteq \mathcal{R}(\Lambda)$, a (continuous) function $e: R \to \mathbb{R}$ and a parameter $\theta \in (0, 1]$, we define a procedure

$$\text{Dörfler} = \text{Dörfler}(R, e, \theta)$$

that computes a set $F \subseteq R \subseteq \mathcal{R}(\Lambda)$ of minimal cardinality such that

$$\sum_{p \in F} e(p)^2 \geq \theta \sum_{p \in R} e(p)^2.$$

In practice, for any $p \in R$, the error indicator $e(p)$ will be either an estimator of the coefficient $c_p$ of the function $u$ expanded over the Legendre basis or the projected residual on the $p$-th Legendre basis function.

This corresponds to choose a fraction $\theta$ of the energy associated with the (estimates of the) coefficients in the set $R$. 
The Dörfler marking


Given a multi-index set $\Lambda$, a subset $R \subseteq \mathcal{R}(\Lambda)$, a (continuous) function $e : R \rightarrow \mathbb{R}$ and a parameter $\theta \in (0, 1]$, we define a procedure

$$\text{Dörfler} = \text{Dörfler}(R, e, \theta)$$

that computes a set $F \subseteq R \subseteq \mathcal{R}(\Lambda)$ of minimal cardinality such that

$$\sum_{p \in F} e(p)^2 \geq \theta \sum_{p \in R} e(p)^2.$$

In practice, for any $p \in R$, the error indicator $e(p)$ will be either an estimator of the coefficient $c_p$ of the function $u$ expanded over the Legendre basis or the projected residual on the $p$-th Legendre basis function.

This corresponds to choose a fraction $\theta$ of the energy associated with the (estimates of the) coefficients in the set $R$. 
Orthogonal Matching Pursuit with Dörfler marking

**Algorithm 1** Orthogonal Matching Pursuit with Dörfler marking

Set $r_0 = u(y)$, $u_0 \equiv 0$ and $\Lambda_0 = \{0\}$,

for $k = 1, \ldots, k_{\text{max}}$ do

$F_1 = \text{Dörfler}(\mathcal{R}(\Lambda_{k-1}), \{(r_{k-1}, \psi_p)_{M,V}\}_p, \theta_1)$

$\tilde{\Lambda}_k = \Lambda_{k-1} \cup F_1$

$u_k = \arg\min_{v \in \mathcal{P}_{\tilde{\Lambda}_k}} \|u - v\|_{M,V}$, \quad $u_k = \sum_{p \in \tilde{\Lambda}_k} c^{(k)}_p \psi_p$

$F_2 = \text{Dörfler}(F_1, \{c^{(k)}_p\}_p, \theta_2)$

$\Lambda_k = \Lambda_{k-1} \cup F_2$

$r_k = u - u_k|_{\Lambda_k}$

end for

$\theta_1 \in (0, 1)$ and $\theta_2 = 1$: Dörfler marking only with the correlations.

$\theta_1 = 1$ and $\theta_2 \in (0, 1)$: Dörfler marking only with the random discrete $L^2$ projection on $\Lambda_{k-1} \cup \mathcal{R}(\Lambda_k)$. 

F. Nobile (EPFL)
Some remarks and open issues

- The first Dörfler marking performs a screening of the reduced margin, to avoid an $L^2$ discrete minimization over a too large polynomial space.

- At each step the correlations $\{|(r_{k-1}, \psi_p)_{M,V}| : p \in \mathcal{R}(\Lambda_k)\}$ are mutually uncoupled and cheap to compute, but might provide only a rough estimate of the coefficients (depending on the choice of $M_k$).

- The second Dörfler marking performs a selection based on the more accurate estimates of the coefficients coming from the $L^2$ projection.

- At each step the adaptive algorithm remains stable and accurate by choosing $M_k \propto (\# \Lambda_k)^2$ (consequence of the theory in the first part).

- The adaptive algorithm generates a sequence $\{\Lambda_k\}_{k \geq 0}$ of only quasi best $N$-term sets.

- Rate of convergence? Choice of $\theta_1, \theta_2$? What if $M_k \propto \# \Lambda_k$?
A numerical test

Approximation of a meromorphic function (16-variables)

\[ \phi(y) = \frac{1}{1 - \gamma \cdot y}, \quad y \sim \mathcal{U}([-1, 1]^{16}) \]

\[ \gamma = 0.3 \times (1, 5 \cdot 10^{-1}, 10^{-1}, 5 \cdot 10^{-2}, \ldots, 5 \cdot 10^{-8}) \]
Outline

1. Introduction – PDEs with random parameters

2. Stochastic polynomial approximation

3. Discrete projection using random evaluations
   - Stability
   - Convergence results in expectation and probability
   - The case of noisy observations

4. Conclusions
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.

- The condition $M \geq C(\#Λ)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.

- The condition $M \sim (\#Λ)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.

- Still open questions on preconditioned least squares or unbounded random variables.

- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinite dimensional approximations are possible with this algorithm.

- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.

- The condition $M \geq C(\#\Lambda)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.

- The condition $M \sim (\#\Lambda)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.

- Still open questions on preconditioned least squares or unbounded random variables.

- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinite dimensional approximations are possible with this algorithm.

- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.
- The condition $M \geq C(\#\Lambda)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.
- The condition $M \sim (\#\Lambda)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.
- Still open questions on preconditioned least squares or unbounded random variables.
- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinite dimensional approximations are possible with this algorithm.
- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.

- The condition $M \geq C(\#\Lambda)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.

- The condition $M \sim (\#\Lambda)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.

- Still open questions on preconditioned least squares or unbounded random variables.

- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinite dimensional approximations are possible with this algorithm.

- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.
- The condition $M \geq C(\#\Lambda)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.
- The condition $M \sim (\#\Lambda)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.
- Still open questions on preconditioned least squares or unbounded random variables.
- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinte dimensional approximations are possible with this algorithm.
- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Conclusions

- We have derived conditions under which the random discrete least squares projection is stable and optimally convergent.
- The condition $M \geq C(\#\Lambda)^2$ for uniform points and Legendre polynomials holds in any dimension and for any “shape” of the polynomial space, opening the possibility of adaptive algorithms.
- The condition $M \sim (\#\Lambda)^2$ seems to be too stringent in high dimension and a linear scaling is often enough, making this technique more attractive for high dimensional problems.
- Still open questions on preconditioned least squares or unbounded random variables.
- We have proposed an adaptive algorithm based on a double Dörfler marking that performs very well. The analysis is still ongoing. Very high/infinite dimensional approximations are possible with this algorithm.
- Other sampling schemes can be used to build the discrete least squares (DLS) projection. In [Migliorati-N. 15] we have shown that DLS with low discrepancy sequences has similar stability conditions as the random DLS, at least for tensor product polynomial spaces.
Thank you for your attention!
Conclusions

References

G. Milgiorati, F. Nobile and R. Tempone,
*Convergence estimates in probability and in expectation for discrete least squares with noisy evaluations at random points*, in preparation.

G. Migliorati and F. Nobile,

A. Chkifa, A. Cohen, G. Migliorati, F. Nobile, and R. Tempone
*Discrete least squares polynomial approximation with random evaluations: application to parametric and stochastic elliptic PDEs*, to appear in M2AN, 2015

G. Migliorati,
*Adaptive polynomial approximation by means of random discrete least squares*, ENUMATH 2013 Proceedings, LNCSE Springer.

G. Migliorati, F. Nobile, E. von Schwerin, and R. Tempone

G. Migliorati, F. Nobile, E. von Schwerin and R. Tempone,
*Approximation of quantities of interest in stochastic PDEs by the random discrete $L^2$ projection on polynomial spaces*, SISC 35(3), 2013