Sparse Quadrature Algorithms for Bayesian Inverse Problems

Claudia Schillings, Christoph Schwab

Pro*Doc Retreat Disentis 2013

Numerical Analysis and Scientific Computing
Disentis - 15 August, 2013

SAM
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

erc
European Research Council
Established by the European Commission

research supported by the European Research Council under FP7 Grant AdG247277

Outline

Bayesian Inverse Problems (Stuart 2010)

Find the unknown data $u \in X$ from noisy observations

$$\delta = \mathcal{G}(u) + \eta,$$

- X separable Banach space
- $G : X \mapsto \mathcal{X}$ the forward map

Abstract Operator Equation

$$\text{Given } u \in X, \text{ find } q \in \mathcal{X} : \quad A(u; q) = f$$

with $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$, $\mathcal{X}$, $\mathcal{Y}$ reflexive Banach spaces, $\mathfrak{a}(v, w) :=_{\mathcal{Y}} \langle w, Av \rangle_{\mathcal{Y}'}$ $\forall v \in \mathcal{X}, w \in \mathcal{Y}$ corresponding bilinear form

- $\mathcal{O} : \mathcal{X} \mapsto \mathbb{R}^K$ bounded, linear observation operator
- $\mathcal{G} : X \mapsto \mathbb{R}^K$ uncertainty-to-observation map, $\mathcal{G} = \mathcal{O} \circ G$
- $\eta \in \mathbb{R}^K$ the observational noise ($\eta \sim \mathcal{N}(0, \Gamma)$)

Bayesian Inverse Problems (Stuart 2010)

Find the unknown data $u \in X$ from noisy observations

$$\delta = \mathcal{G}(u) + \eta,$$

- X separable Banach space
- $G : X \mapsto \mathcal{X}$ the forward map
- $\mathcal{O} : \mathcal{X} \mapsto \mathbb{R}^K$ bounded, linear observation operator
- $\mathcal{G} : X \mapsto \mathbb{R}^K$ uncertainty-to-observation map, $\mathcal{G} = \mathcal{O} \circ G$
- $\eta \in \mathbb{R}^K$ the observational noise ($\eta \sim \mathcal{N}(0, \Gamma)$)

Least squares potential $\Phi : X \times \mathbb{R}^K \to \mathbb{R}$

$$\Phi(u; \delta) := \frac{1}{2}\Big((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\Big)$$

Reformulation of the forward problem with unknown stochastic input data as an infinite dimensional, parametric deterministic problem

Bayesian Inverse Problems (Stuart 2010)

Parametric representation of the unknown u

$$u = u(\boldsymbol{y}) := \langle u \rangle + \sum_{j \in \mathbb{J}} y_j \psi_j \in X$$

- $y = (y_j)_{j \in \mathbb{J}}$ i.i.d sequence of real-valued random variables $y_j \sim \mathcal{U}[-1,1]$
- $\langle u \rangle, \psi_j \in X$
- $\mathbb{J}$ finite or countably infinite index set

Prior measure on the uncertain input data

$$\mu_0(dy) := \bigotimes_{j \in \mathbb{J}} \frac{1}{2} \lambda_1(dy_j) \ .$$

- $(U, \mathcal{B}) = \left([-1,1]^{\mathbb{J}}, \ \bigotimes_{j \in \mathbb{J}} \mathcal{B}^1[-1,1] \right)$ measurable space

(p, ϵ) Analyticity

$(p, \epsilon) : 1$ (well-posedness)

For each $\boldsymbol{y} \in U$, there exists a unique realization $u(\boldsymbol{y}) \in X$ and a unique solution $q(\boldsymbol{y}) \in \mathcal{X}$ of the forward problem. This solution satisfies the a-priori estimate

$$\forall \boldsymbol{y} \in U : \quad \|q(\boldsymbol{y})\|_{\mathcal{X}} \leq C_0(\boldsymbol{y}) \,,$$

where $U \ni \boldsymbol{y} \mapsto C_0(\boldsymbol{y}) \in L^1(U; \mu_0)$.

$(p, \epsilon) : 2$ (analyticity)

There exist $0 < p < 1$ and $b = (b_j)_{j \in \mathbb{J}} \in \ell^p(\mathbb{J})$ such that for $0 < \epsilon < 1$, there exist $C_\epsilon > 0$ and $\rho = (\rho_j)_{j \in \mathbb{J}}$ of poly-radii $\rho_j > 1$ such that

$$\sum_{j \in \mathbb{J}} \rho_j b_j \leq 1 - \epsilon \,,$$

and $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y}) \in \mathcal{X}$ admits an analytic continuation to the open polyellipse $\mathcal{E}_\rho := \prod_{j \in \mathbb{J}} \mathcal{E}_{\rho_j} \subset \mathbb{C}^{\mathbb{J}}$ with

$$\forall \boldsymbol{z} \in \mathcal{E}_\rho : \quad \|q(\boldsymbol{z})\|_{\mathcal{X}} \leq C_\epsilon(\boldsymbol{y}) \,.$$

Sparsity of the Forward Solution

Theorem (Chkifa, Cohen, DeVore and Schwab)

Assume that the parametric forward solution map $q(\boldsymbol{y})$ admits a (p,ϵ)-analytic extension to the poly-ellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$.

- The Legendre series converges unconditionally,

$$q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} q_\nu^P P_\nu(\boldsymbol{y}) \quad \text{in } L^\infty(U, \mu_0; \mathcal{X})$$

  with Legendre polynomials $P_k(1) = 1$, $\|P_k\|_{L^\infty(-1,1)} = 1$, $\quad k = 0, 1, ...$

- There exists a p-summable, monotone envelope $\boldsymbol{q} = \{\boldsymbol{q}_\nu\}_{\nu \in \mathcal{F}}$, i.e. $\boldsymbol{q}_\nu := \sup_{\mu \geq \nu} \|q_\nu^P\|_{\mathcal{X}}$ with $C(p, \boldsymbol{q}) := \|\boldsymbol{q}\|_{\ell^p(\mathcal{F})} < \infty$.
  and monotone $\Lambda_N^P \subset \mathcal{F}$ corresponding to the N largest terms of $\boldsymbol{q}$ with

$$\sup_{\boldsymbol{y} \in U} \left\| q(\boldsymbol{y}) - \sum_{\nu \in \Lambda_N^P} q_\nu^P P_\nu(\boldsymbol{y}) \right\|_{\mathcal{X}} \leq C(p, \boldsymbol{q}) N^{-(1/p-1)} .$$

(p,ϵ) Analyticity of Affine Parametric Operator Families

Affine Parametric Operator Families

$$A(\boldsymbol{y}) = A_0 + \sum_{j\in\mathbb{J}} y_j A_j \in \mathcal{L}(\mathcal{X},\mathcal{Y}') \;.$$

Assumption A1 There exists $\mu > 0$ such that

$$\inf_{0\neq v\in\mathcal{X}} \sup_{0\neq w\in\mathcal{Y}} \frac{\mathfrak{a}_0(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0 \;, \inf_{0\neq w\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}} \frac{\mathfrak{a}_0(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0$$

Assumption A2 There exists a constant $0 < \kappa < 1$

$$\sum_{j\in\mathbb{J}} b_j \leq \kappa < 1 \;, \quad \text{where} \quad b_j := \|A_0^{-1} A_j\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')}$$

Assumption A3 For some $0 < p < 1$

$$\|b\|^p_{\ell^p(\mathbb{J})} = \sum_{j\in\mathbb{J}} b_j^p < \infty$$

(p, ϵ) Analyticity of Affine Parametric Operator Families

Theorem (Cohen, DeVore and Schwab 2010)

Under Assumption **A1 - A3**, for every realization $\boldsymbol{y} \in U$ of the parameters, $A(\boldsymbol{y})$ is boundedly invertible, uniformly with respect to the parameter sequence $\boldsymbol{y} \in U$.

For the parametric bilinear form $\mathfrak{a}(\boldsymbol{y};\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, there holds the uniform inf-sup conditions with $\mu = (1-\kappa)\mu_0$,

$$\forall \boldsymbol{y} \in U : \qquad \inf_{0 \neq v \in \mathcal{X}} \sup_{0 \neq w \in \mathcal{Y}} \frac{\mathfrak{a}(\boldsymbol{y}; v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu \,, \; \inf_{0 \neq w \in \mathcal{Y}} \sup_{0 \neq v \in \mathcal{X}} \frac{\mathfrak{a}(\boldsymbol{y}; v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu \,.$$

The forward map $q : U \to \mathcal{X},\ q := G(u)$ and the uncertainty-to-observation map $\mathcal{G} : U \to \mathbb{R}^K$ are globally Lipschitz and (p, ϵ)-analytic with $0 < p < 1$ as in Assumption **A3**.

Examples

Stationary Elliptic Diffusion Problem

$$A_1(u;q) := -\nabla \cdot \big(u \nabla q\big) = f \quad \text{in} \quad D, \qquad q = 0 \quad \text{in} \quad \partial D$$

with $\mathcal{X} = \mathcal{Y} = V = H_0^1(D)$.

Time Dependent Diffusion

$$A_2(\boldsymbol{y}) := (\partial_t + A_1(\boldsymbol{y}), \iota_0)$$

where ι_0 denotes the time $t = 0$ trace,

$\mathcal{X} = L^2(0,T;V) \cap H^1(0,T;V^*)$, $\mathcal{Y} = L^2(0,T;V) \times H$.

Bayesian Inverse Problem

Theorem (Schwab and Stuart 2011)

Assume that $\mathcal{G}(u)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j}$ is bounded and continuous.

Then $\mu^\delta(d\boldsymbol{y})$, the distribution of $y \in U$ given δ, is absolutely continuous with respect to $\mu_0(d\boldsymbol{y})$, ie.

$$\frac{d\mu^\delta}{d\mu_0}(y) = \frac{1}{Z}\Theta(y)$$

with the parametric Bayesian posterior Θ given by

$$\Theta(\boldsymbol{y}) = \exp\big(-\Phi(u;\delta)\big)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j},$$

and the normalization constant

$$Z = \int_U \Theta(y)\mu_0(d\boldsymbol{y}) \ .$$

Bayesian Inverse Problem

Expectation of a *Quantity of Interest* $\phi : X \to S$

$$\mathbb{E}^{\mu^\delta}[\phi(u)] = Z^{-1} \int_U \exp\big(-\Phi(u;\delta)\big)\phi(u)\Big|_{u=\langle u\rangle + \sum_{j\in\mathbb{J}} y_j\psi_j} \mu_0(dy) =: Z'/Z$$

with $Z = \int_{y\in U} \exp(-\frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\right))\mu_0(dy)$.

- Reformulation of the forward problem with unknown stochastic input data as an infinite dimensional, parametric deterministic problem
- Parametric, deterministic representation of the derivative of the posterior measure with respect to the prior μ_0
- Approximation of Z' and Z to compute the expectation of QoI under the posterior given data δ

Efficient algorithm to approximate the conditional expectations given the data with dimension-independent rates of convergence

Sparsity of the Posterior Density

Theorem (C.S. and Ch. Schwab 2013)

Assume that the forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is (p, ϵ)-analytic for some $0 < p < 1$.
Then the Bayesian posterior density $\Theta(\boldsymbol{y})$ is, as a function of the parameter $\boldsymbol{y}$, likewise (p, ϵ)-analytic, with the same p and the same ϵ.

N-term Approximation Results

$$\sup_{\boldsymbol{y} \in U} \left\| \Theta(\boldsymbol{y}) - \sum_{\nu \in \Lambda_N^P} \Theta_\nu^P P_\nu(\boldsymbol{y}) \right\|_{\mathcal{X}} \leq N^{-s} \|\boldsymbol{\theta}^P\|_{\ell_m^p(\mathcal{F})}, \quad s := \frac{1}{p} - 1 \ .$$

Adaptive Smolyak quadrature algorithm with convergence rates depending only on the summability of the parametric operator

Univariate Quadrature

Univariate quadrature operators of the form

$$Q^k(\mathrm{g}) = \sum_{i=0}^{n_k} w_i^k \cdot \mathrm{g}(z_i^k)$$

with $\mathrm{g} : [-1,1] \mapsto \mathcal{S}$ for some Banach space $\mathcal{S}$

- $(Q^k)_{k\geq 0}$ sequence of univariate quadrature formulas
- $(z_j^k)_{j=0}^{n_k} \subset [-1,1]$ with $z_j^k \in [-1,1]\,, \forall j,k$ and $z_0^k = 0\,, \forall k$ quadrature points
- $w_j^k\,, 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0$ quadrature weights

Assumption 1

(i) $(I - Q^k)(g_k) = 0\,, \quad \forall g_k \in \mathbb{P}_k = \mathrm{span}\{y^j : j \in \mathbb{N}_0, j \leq k\}$
with $I(g_k) = \int_{[-1,1]} g_k(y)\lambda_1(dy)$

(ii) $w_j^k > 0\,, \qquad 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0.$

Univariate Quadrature

Univariate quadrature operators of the form

$$Q^k(\mathrm{g}) = \sum_{i=0}^{n_k} w_i^k \cdot \mathrm{g}(z_i^k)$$

with $\mathrm{g} : [-1,1] \mapsto \mathcal{S}$ for some Banach space $\mathcal{S}$

- $(Q^k)_{k\geq 0}$ sequence of univariate quadrature formulas
- $(z_j^k)_{j=0}^{n_k} \subset [-1,1]$ with $z_j^k \in [-1,1]\,, \forall j,k$ and $z_0^k = 0\,, \forall k$ quadrature points
- $w_j^k\,, 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0$ quadrature weights

Univariate quadrature difference operator

$$\Delta_j = Q^j - Q^{j-1}, \qquad j \geq 0$$

with $Q^{-1} = 0$ and $z_0^0 = 0, w_0^0 = 1$

Univariate Quadrature

Univariate quadrature operators of the form

$$Q^k(\mathrm{g}) = \sum_{i=0}^{n_k} w_i^k \cdot \mathrm{g}(z_i^k)$$

with $\mathrm{g} : [-1,1] \mapsto \mathcal{S}$ for some Banach space $\mathcal{S}$

- $(Q^k)_{k\geq 0}$ sequence of univariate quadrature formulas
- $(z_j^k)_{j=0}^{n_k} \subset [-1,1]$ with $z_j^k \in [-1,1]\,, \forall j,k$ and $z_0^k = 0\,, \forall k$ quadrature points
- $w_j^k\,, 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0$ quadrature weights

Univariate quadrature operator rewritten as telescoping sum

$$Q^k = \sum_{j=0}^{k} \Delta_j$$

with $\mathcal{Z}^k = \{z_j^k : 0 \leq j \leq n_k\} \subset [-1,1]$ set of points corresponding to Q^k

Tensorization

Tensorized multivariate operators

$$\mathcal{Q}_\nu = \bigotimes_{j\geq 1} Q^{\nu_j}\,, \qquad \Delta_\nu = \bigotimes_{j\geq 1} \Delta_{\nu_j}$$

with associated set of multivariate points $\mathcal{Z}^\nu = \times_{j\geq 1}\mathcal{Z}^{\nu_j} \in U$

- If $\nu = 0_\mathcal{F}$, then $\Delta_\nu g = Q^\nu g = g(z_{0_\mathcal{F}}) = g(0_\mathcal{F})$
- If $0_\mathcal{F} \neq \nu \in \mathcal{F}$, with $\hat{\nu} = (\nu_j)_{j\neq i}$

$$Q^\nu g = Q^{\nu_i}(t \mapsto \bigotimes_{j\geq 1} Q^{\hat{\nu}_j} g_t)\,, \qquad i \in \mathbb{I}_\nu$$

and

$$\Delta_\nu g = \Delta_{\nu_i}(t \mapsto \bigotimes_{j\geq 1} \Delta_{\hat{\nu}_j} g_t)\,, \qquad i \in \mathbb{I}_\nu\,,$$

for $g \in \mathcal{Z}$, g_t is the function defined on $\mathcal{Z}^\mathbb{N}$ by
$g_t(\hat{y}) = g(y), y = (\ldots, y_{i-1}, t, y_{i+1}, \ldots)\,, i > 1$ and $y = (t, y_2, \ldots)\,, i = 1$

Sparse Quadrature Operator

For any finite monotone set $\Lambda \subset \mathcal{F}$, the quadrature operator is defined by

$$\mathcal{Q}_\Lambda = \sum_{\nu \in \Lambda} \Delta_\nu = \sum_{\nu \in \Lambda} \bigotimes_{j \geq 1} \Delta_{\nu_j}$$

with associated collocation grid

$$\mathcal{Z}_\Lambda = \cup_{\nu \in \Lambda} \mathcal{Z}^\nu$$

Theorem

For any monotone index set $\Lambda_N \subset \mathcal{F}$, the sparse quadrature $\mathcal{Q}_{\Lambda_N}$ is exact for any polynomial $g \in \mathbb{P}_{\Lambda_N}$, i.e. it holds

$$\mathcal{Q}_{\Lambda_N}(g) = I(g), \qquad \forall g \in \mathbb{P}_{\Lambda_N} ,$$

with $\mathbb{P}_{\Lambda_N} = \operatorname{span}\{y^\nu : \nu \in \Lambda_N\}$ and $I(g) = \int_U g(y)\mu_0(dy)$.

Convergence Rates for Adaptive Smolyak Integration

Theorem

Assume that the forward solution map $U \ni \mathbf{y} \mapsto q(\mathbf{y})$ is (p, ϵ)-analytic for some $0 < p < 1$.

Then there exist two sequences $(\Lambda_N^1)_{N \geq 1}$, $(\Lambda_N^2)_{N \geq 1}$ of monotone index sets $\Lambda_N^{1,2} \subset \mathcal{F}$ such that $\#\Lambda_N^{1,2} \leq N$ and

$$|I[\Theta] - \mathcal{Q}_{\Lambda_N^1}[\Theta]| \leq C^1 N^{-s},$$

with $s = 1/p - 1$, $I[\Theta] = \int_U \Theta(y)\mu_0(d\mathbf{y})$ and,

$$\|I[\Psi] - \mathcal{Q}_{\Lambda_N^2}[\Psi]\|_{\mathcal{X}} \leq C^2 N^{-s}, \qquad s = \frac{1}{p} - 1.$$

with $s = 1/p - 1$, $I[\Psi] = \int_U \Psi(y)\mu_0(d\mathbf{y})$, $C^1, C^2 > 0$ independent of N.

C.S. and Ch. Schwab. Sparsity in Bayesian Inversion of Parametric Operator Equations, 2013.

Convergence Rates for Adaptive Smolyak Integration

Sketch of proof

- Relating the quadrature error with the Legendre coefficients

$$|I(\Theta) - \mathcal{Q}_\Lambda(\Theta)| \leq 2 \cdot \sum_{\nu \notin \Lambda} \gamma_\nu |\theta_\nu^P|$$

  and

$$\|I(\Psi) - \mathcal{Q}_\Lambda(\Psi)\|_{\mathcal{X}} \leq 2 \cdot \sum_{\nu \notin \Lambda} \gamma_\nu \|\psi_\nu^P\|_{\mathcal{X}}$$

  for any monotone set $\Lambda \subset \mathcal{F}$, where $\gamma_\nu := \prod_{j \in \mathbb{J}} (1 + \nu_j)^2$.

- $(\gamma_\nu |\theta_\nu^P|)_{\nu \in \mathcal{F}} \in l_m^p(\mathcal{F})$ and $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{X}})_{\nu \in \mathcal{F}} \in l_m^p(\mathcal{F})$.

$\Rightarrow \exists$ sequence $(\Lambda_N)_{N \geq 1}$ of monotone sets $\Lambda_N \subset \mathcal{F}$, $\#\Lambda_N \leq N$, such that the Smolyak quadrature converges with order $1/p - 1$.

Adaptive Construction of the Monotone Index Set

Successive identification of the N largest contributions

$$|\Delta_\nu(\Theta)| = |\bigotimes_{j\geq 1} \Delta_{\nu_j}(\Theta)|, \quad \nu \in \mathcal{F}$$

$\longrightarrow$ A. Chkifa, A. Cohen and Ch. Schwab. High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs, 2012.

Set of reduced neighbors

$$\mathcal{N}(\Lambda) := \{\nu \notin \Lambda : \nu - e_j \in \Lambda, \forall j \in \mathbb{I}_\nu \text{ and } \nu_j = 0\,, \forall j > j(\Lambda) + 1\}$$

with $j(\Lambda) = \max\{j : \nu_j > 0$ for some $\nu \in \Lambda\}$, $\mathbb{I}_\nu = \{j \in \mathbb{N} : \nu_j \neq 0\} \subset \mathbb{N}$

Adaptive Construction of the Monotone Index Set

function ASG
  Set $\Lambda_1 = \{0\}$, $k = 1$ and compute $\Delta_0(\Theta)$.
  Determine the set of reduced neighbors $\mathcal{N}(\Lambda_1)$.
  Compute $\Delta_\nu(\Theta), \forall \nu \in \mathcal{N}(\Lambda_1)$.
  while $\sum_{\nu \in \mathcal{N}(\Lambda_k)} |\Delta_\nu(\Theta)| > tol$ **do**
    Select $\nu \in \mathcal{N}(\Lambda_k)$ with largest $|\Delta_\nu|$ and set $\Lambda_{k+1} = \Lambda_k \cup \{\nu\}$.
    Determine the set of reduced neighbors $\mathcal{N}(\Lambda_{k+1})$.
    Compute $\Delta_\nu(\Theta), \forall \nu \in \mathcal{N}(\Lambda_{k+1})$.
    Set $k = k + 1$.
  end while
end function

T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature, *Computing*, 2003

Numerical Experiments

Model parametric parabolic problem

$$\begin{aligned} \partial_t q(t,x) - \mathrm{div}(u(x)\nabla q(t,x)) &= 100 \cdot tx && (t,x) \in T \times D\,, \\ q(0,x) &= 0 && x \in D\,, \\ q(t,0) = q(t,1) &= 0 && t \in T \end{aligned}$$

with

$$u(x,y) = \langle u \rangle + \sum_{j=1}^{64} y_j \psi_j\,, \text{where } \langle u \rangle = 1 \text{ and } \psi_j = \alpha_j \chi_{D_j}$$

where $D_j = [(j-1)\frac{1}{64}, j\frac{1}{64}]$, $y = (y_j)_{j=1,\ldots,64}$ and $\alpha_j = \frac{0.9}{j^\zeta}$, $\zeta = 2, 3, 4$.

- Finite element method using continuous, piecewise linear ansatz functions in space, backward Euler scheme in time
- Uniform mesh with meshwidth $h_T = h_D = 2^{-11}$
- LAPACK's DPTSV routine

Numerical Experiments

Find the unknown data u for given (noisy) data δ,

$$\delta = \mathcal{G}(u) + \eta \,,$$

Expectation of interest Z'/Z

$$
\begin{aligned}
Z' &= \int_U \exp\big(-\Phi(u;\delta)\big)\phi(u)\Big|_{u=\langle u\rangle+\sum_{j=1}^{64} y_j\psi_j}\mu_0(dy) \\
Z &= \int_U \exp\big(-\Phi(u;\delta)\big)\Big|_{u=\langle u\rangle+\sum_{j=1}^{64} y_j\psi_j}\mu_0(dy)
\end{aligned}
$$

- Observation operator $\mathcal{O}$ consists of system responses at K observation points in $T \times D$ at $t_i = \frac{i}{2^{N_{K,T}}}, i = 1, \ldots, 2^{N_{K,T}} - 1$, $x_j = \frac{j}{2^{N_{K,D}}}, k = 1, \ldots, 2^{N_{K,D}} - 1$, $o_k(\cdot,\cdot) = \delta(\cdot - t_k)\delta(\cdot - x_k)$ with $K = 1,\ N_{K,D} = 1, N_{K,T} = 1,\ K = 3,\ N_{K,D} = 2, N_{K,T} = 1,\ K = 9,\ N_{K,D} = 2, N_{K,T} = 2$
- $\mathcal{G} : X \to \mathbb{R}^K$, with $K = 1, 3, 9$, $\phi(u) = G(u)$
- $\eta = (\eta_j)_{j=1,\ldots,K}$ iid with $\eta_j \sim \mathcal{N}(0, 1)$, $\eta_j \sim \mathcal{N}(0, 0.5^2)$ and $\eta_j \sim \mathcal{N}(0, 0.1^2)$

Numerical Experiments

Quadrature points

- Clenshaw-Curtis (CC)

$$\begin{aligned} z_j^k &= -cos\left(\frac{\pi j}{n_k - 1}\right), j = 0, \dots, n_k - 1, \text{if } n_k > 1 \text{ and} \\ z_0^k &= 0, \text{if } n_k = 1 \end{aligned}$$

  with $n_0 = 1$ and $n_k = 2^k + 1$, for $k \geq 1$

- $\mathfrak{R}$-Leja sequence (RL)

Numerical Experiments

Quadrature points

- Clenshaw-Curtis (CC)
- $\mathfrak{R}$-Leja sequence (RL)
  projection on $[-1,1]$ of a Leja sequence for the complex unit disk initiated at i

$$
\begin{aligned}
z_0^k &= 0\,, z_1^k = 1\,, z_2^k = -1\,, \text{if } j = 0,1,2 \text{ and} \\
z_j^k &= \mathfrak{R}(\hat{z}), \text{ with } \hat{z} = \underset{|z|\le 1}{\operatorname{argmax}} \prod_{l=1}^{j-1} |z - z_l^k|\,, j = 3,\ldots,n_k, \text{if } j \text{ odd}\,, \\
z_j^k &= -z_{j-1}^k\,, j = 3,\ldots,n_k, \text{if } j \text{ even}\,,
\end{aligned}
$$

with $n_k = 2\cdot k + 1$, for $k \ge 0$

J.-P. Calvi and M. Phung Van. On the Lebesgue constant of Leja sequences for the unit disk and its applications to multivariate interpolation *Journal of Approximation Theory*, 2011.

J.-P. Calvi and M. Phung Van. Lagrange interpolation at real projections of Leja sequences for the unit disk *Proceedings of the American Mathematical Society*, 2012.

A. Chkifa. On the Lebesgue constant of Leja sequences for the unit disk *Journal of Approximation Theory*, 2013.

Leja quadrature points

Proposition

Let $\mathcal{Q}_\Lambda^{RL}$ denote the sparse quadrature operator for any monotone set Λ based on the univariate quadrature formulas associated with the $\mathfrak{R}$-Leja sequence.

If the forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is (p, ϵ)-analytic for some $0 < p < 1$ and $\epsilon > 0$, then $(\gamma_\nu |\theta_\nu^P|)_{\nu \in \mathcal{F}} \in l_m^p(\mathcal{F})$ and $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{S}})_{\nu \in \mathcal{F}} \in l_m^p(\mathcal{F})$.

Furthermore, there exist two sequences $(\Lambda_N^{RL,1})_{N \geq 1}$, $(\Lambda_N^{RL,2}))_{N \geq 1}$ of monotone index sets $\Lambda_N^{RL,i} \subset \mathcal{F}$ such that $\#\Lambda_N^{RL,i} \leq N$, $i = 1, 2$, and such that, for some $C^1, C^2 > 0$ independent of N, with $s = \frac{1}{p} - 1$,

$$|I[\Theta] - \mathcal{Q}_{\Lambda_N^{RL,1}}[\Theta]| \leq C^1 N^{-s} ,$$

and

$$\|I[\Psi] - \mathcal{Q}_{\Lambda_N^{RL,2}}[\Psi[\|_{\mathcal{S}} \leq C^2 N^{-s} .$$

Leja quadrature points

Sketch of proof

Univariate polynomial interpolation operator

$$\mathcal{I}_{RL}^{k}(g) = \sum_{i=0}^{n_k} \mathrm{g}\left(z_i^k\right) \cdot l_i^k \,,$$

with $\mathrm{g} : U \mapsto \mathcal{S}$, $l_i^k(y) := \prod_{i=0, i\neq j}^{n_k} \frac{y - z_i}{z_j - z_i}$ the Lagrange polynomials.

- $$(I - Q_{RL}^k)(g_k) = (I - I[\mathcal{I}_{RL}^k])(g_k) = I(g_k - \mathcal{I}_{RL}^k(g_k)) = 0$$

  $\forall g_k \in \mathbb{P}_k = \mathrm{span}\{y^j : j \in \mathbb{N}_0, j \leq k\}$.

Leja quadrature points

Sketch of proof

Univariate polynomial interpolation operator

$$\mathcal{I}^k_{RL}(g) = \sum_{i=0}^{n_k} \mathrm{g}\left(z_i^k\right) \cdot l_i^k \,,$$

with $\mathrm{g} : U \mapsto \mathcal{S}$, $l_i^k(y) := \prod_{i=0, i\neq j}^{n_k} \frac{y - z_i}{z_j - z_i}$ the Lagrange polynomials.

- $$(I - Q^k_{RL})(g_k) = 0\,, \quad \forall g_k \in \mathbb{P}_k$$

- $$\begin{aligned} \|Q^k_{RL}\| &= \sup_{0\neq g \in C(U;\mathcal{S})} \frac{\|Q^k_{RL}(\mathrm{g})\|_{\mathcal{S}}}{\|\mathrm{g}\|_{L^\infty(U;\mathcal{S})}} \\ &\leq \sup_{0\neq g \in C(U;\mathcal{S})} \frac{\|\mathcal{I}^k_{RL}(\mathrm{g})\|_{L^\infty(U;\mathcal{S})}}{\|\mathrm{g}\|_{L^\infty(U;\mathcal{S})}} \leq 3(k+1)^2 \log(k+1) \end{aligned}$$

Leja quadrature points

Sketch of proof

Univariate polynomial interpolation operator

$$\mathcal{I}_{RL}^k(g) = \sum_{i=0}^{n_k} \mathrm{g}\left(z_i^k\right) \cdot l_i^k \,,$$

with $\mathrm{g} : U \mapsto \mathcal{S}$, $l_i^k(y) := \prod_{i=0, i\neq j}^{n_k} \frac{y-z_i}{z_j-z_i}$ the Lagrange polynomials.

- $$(I - Q_{RL}^k)(g_k) = 0\,, \quad \forall g_k \in \mathbb{P}_k$$
- $$\|Q_{RL}^k\| \leq 3(k+1)^2 \log(k+1)$$
- Relating the quadrature error with the Legendre coefficients θ_ν^P of g

Normalization Constant Z

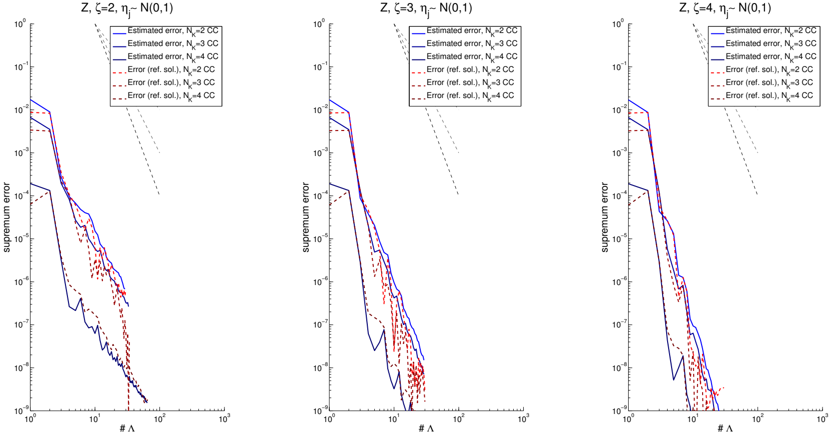

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z

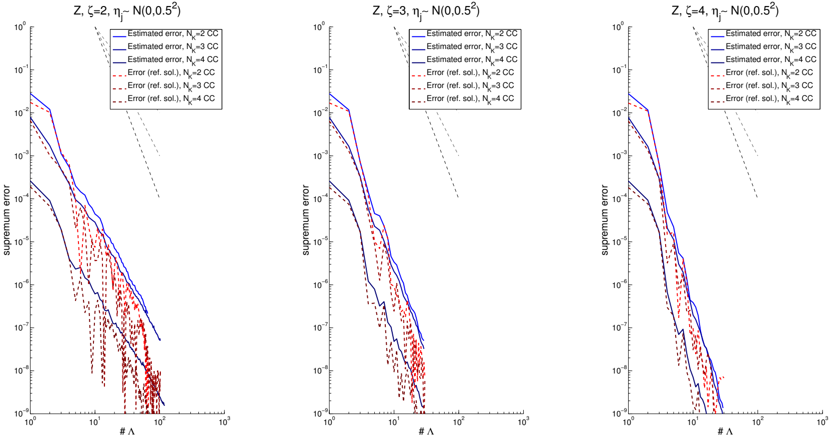

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z

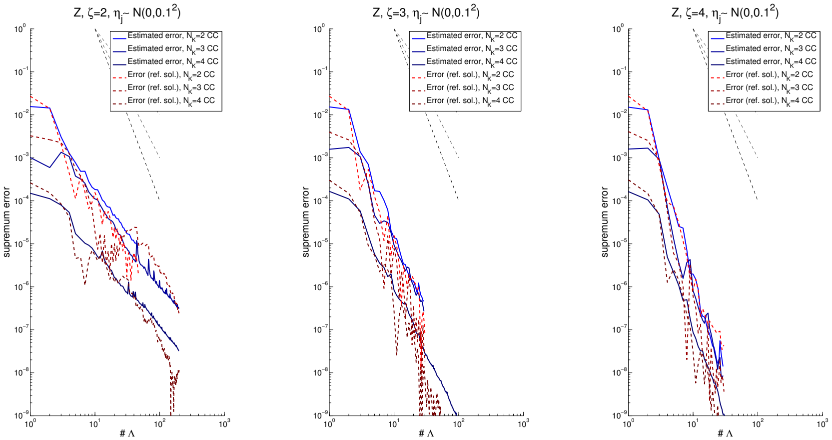

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z

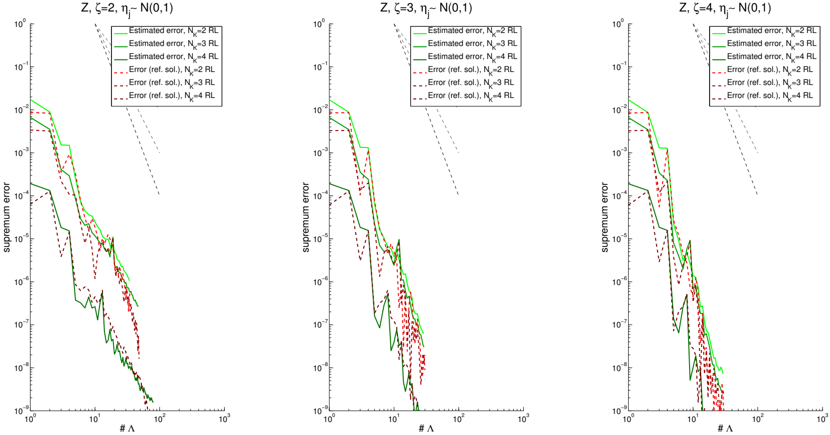

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z

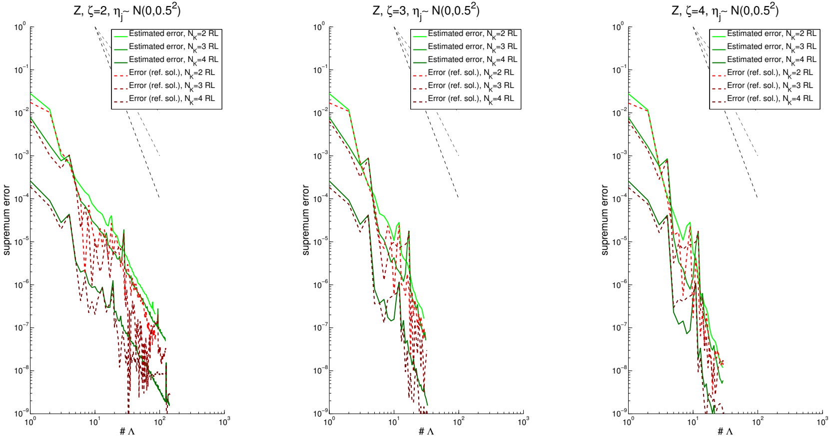

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z

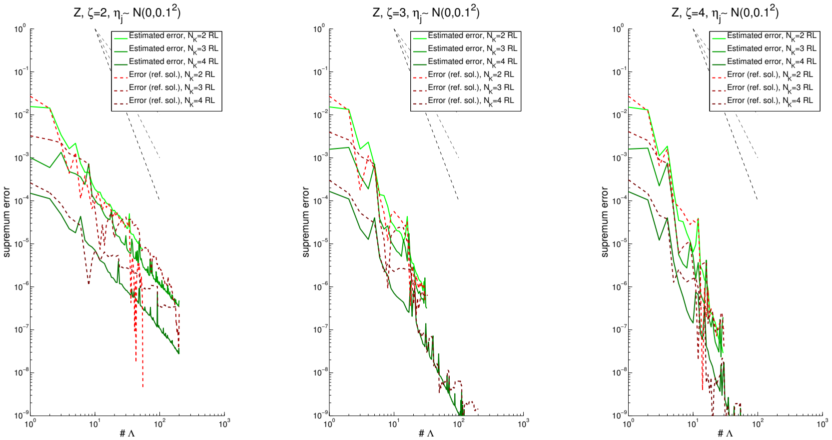

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

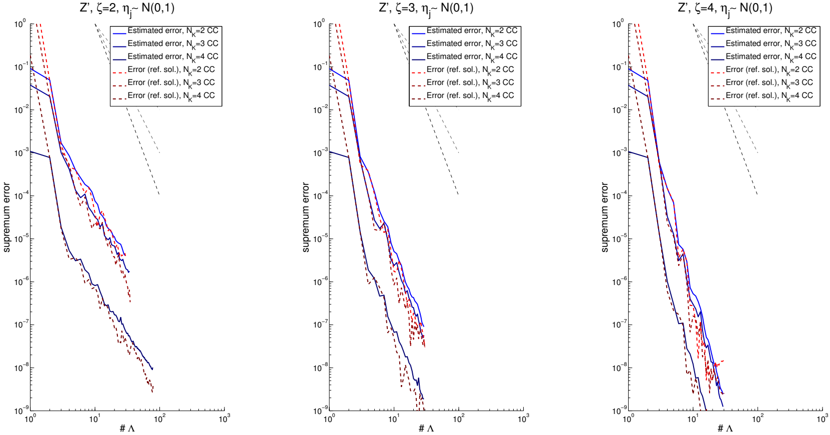

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

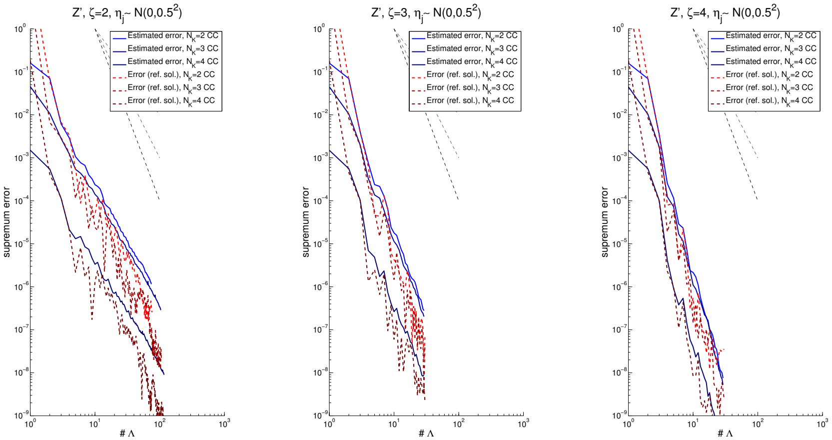

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

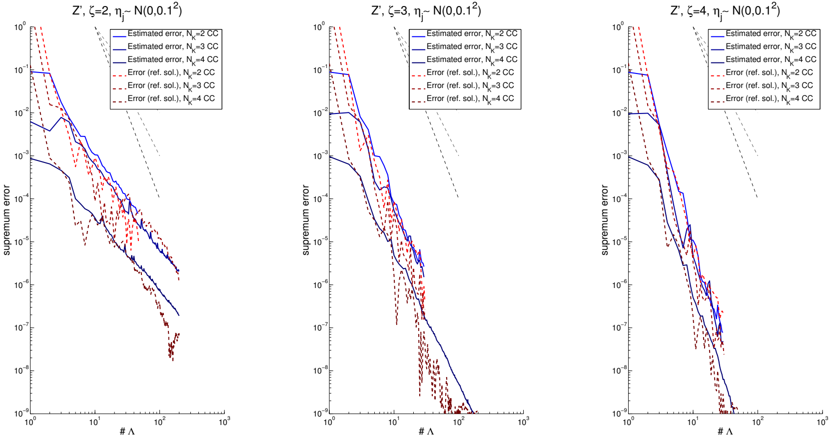

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

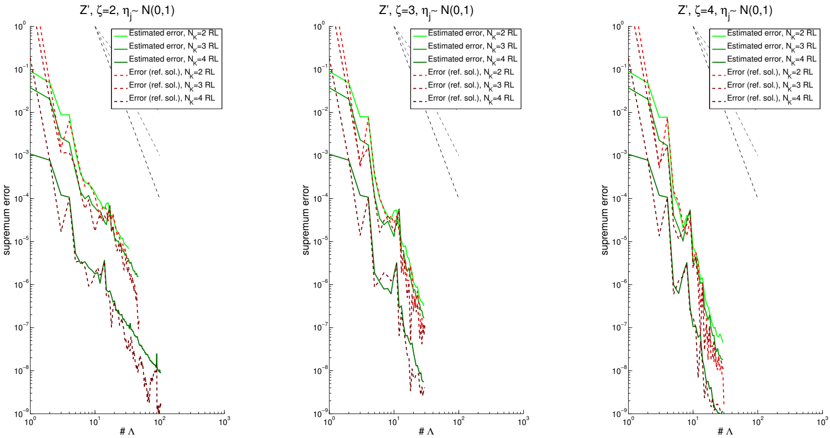

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

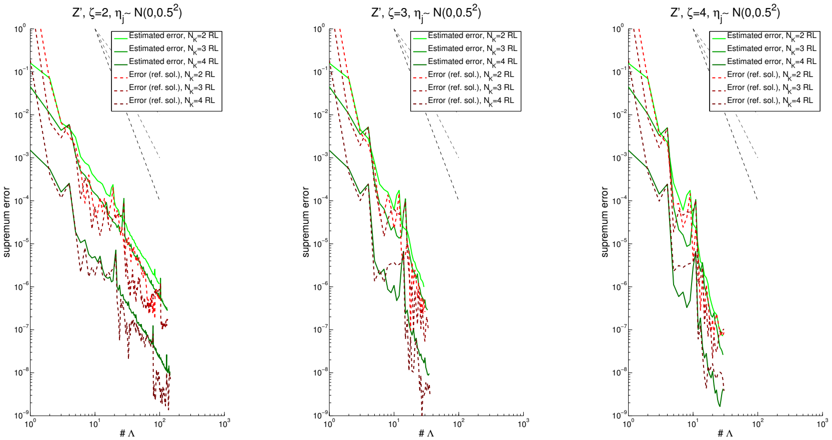

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z'

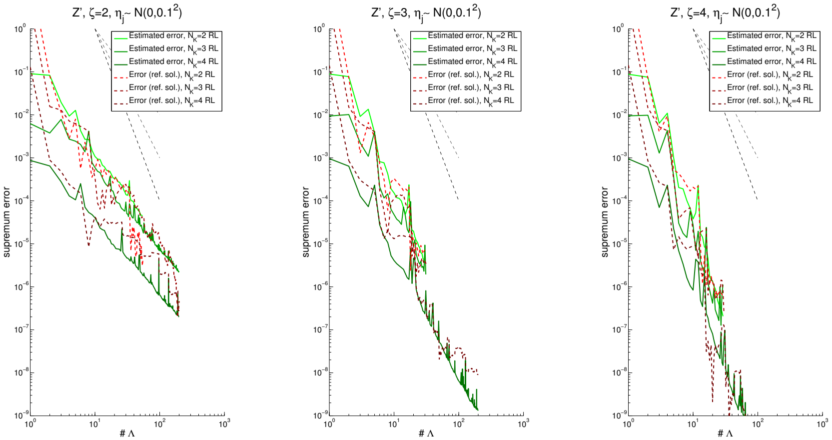

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Numerical Experiments

Model parametric parabolic problem

$$\partial_t q(t,x) - \mathrm{div}(u(x)\nabla q(t,x)) = 100 \cdot tx \qquad (t,x) \in T \times D\,,$$
$$q(0,x) = 0 \qquad x \in D\,,$$
$$q(t,0) = q(t,1) = 0 \qquad t \in T$$

with

$$u(x,y) = \langle u \rangle + \sum_{j=1}^{128} y_j \psi_j\,, \text{where } \langle u \rangle = 1 \text{ and } \psi_j = \alpha_j \chi_{D_j}$$

where $D_j = [(j-1)\frac{1}{128}, j\frac{1}{128}]$, $y = (y_j)_{j=1,\dots,128}$ and $\alpha_j = \frac{0.6}{j^\zeta}\,, \zeta = 2,3,4$.

- Finite element method using continuous, piecewise linear ansatz functions in space, backward Euler scheme in time
- Uniform mesh with meshwidth $h_T = h_D = 2^{-11}$
- LAPACK's DPTSV routine

Normalization Constant Z (128 parameters)

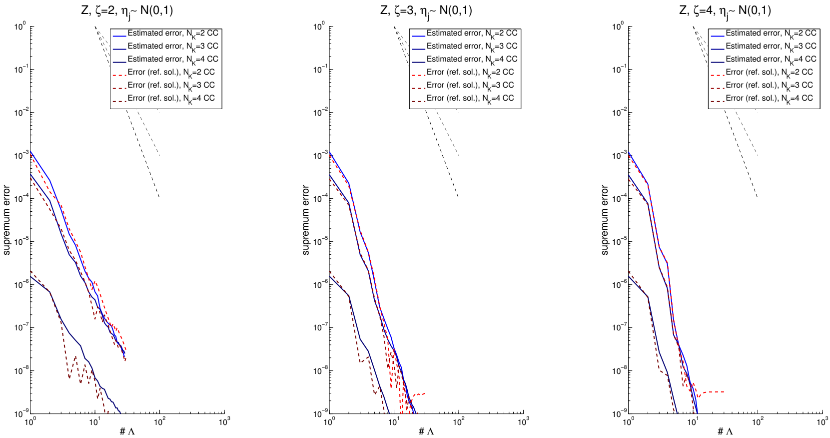

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z (128 parameters)

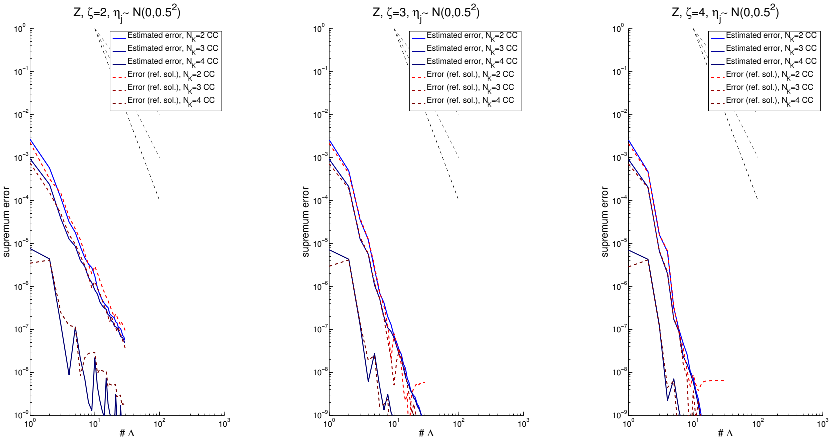

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Normalization Constant Z (128 parameters)

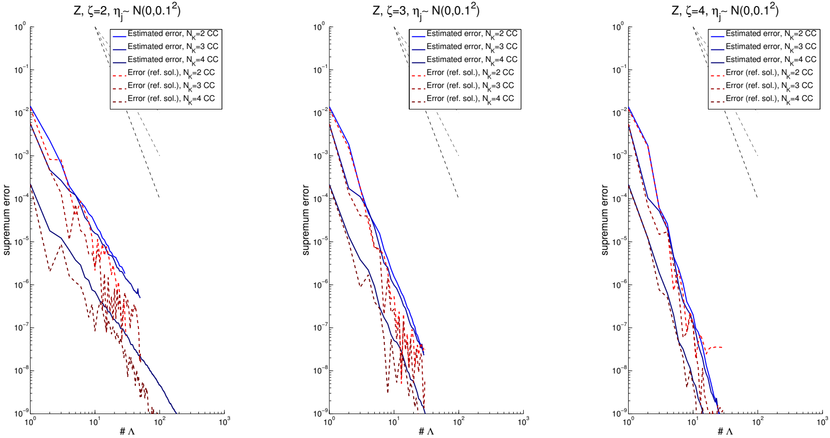

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z' (128 parameters)

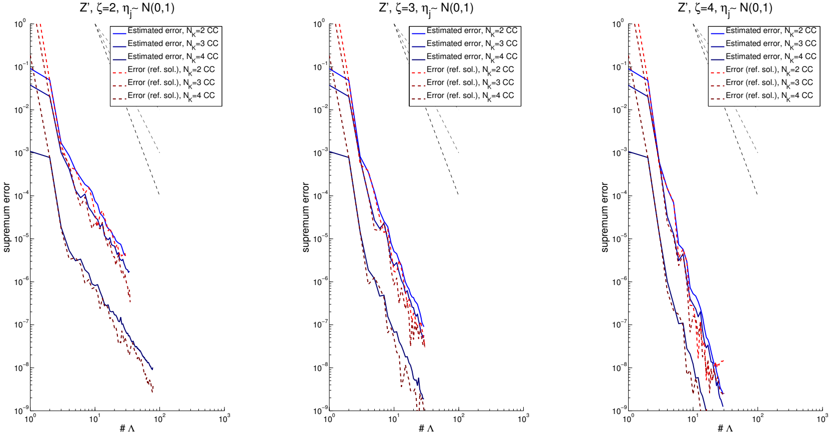

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z' (128 parameters)

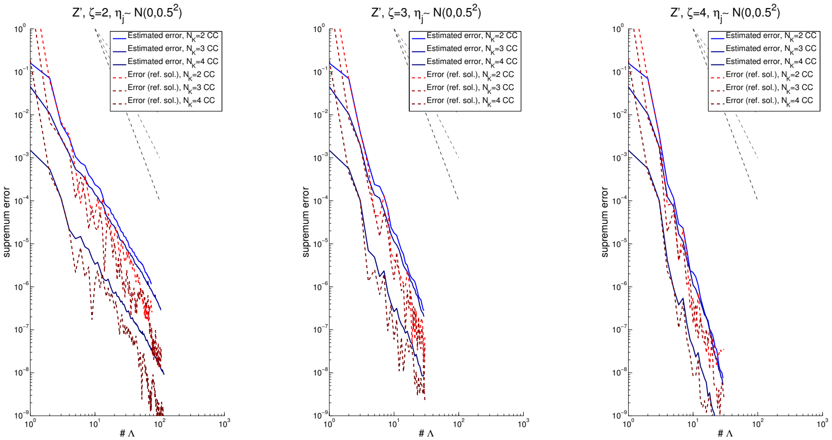

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Quantity Z' (128 parameters)

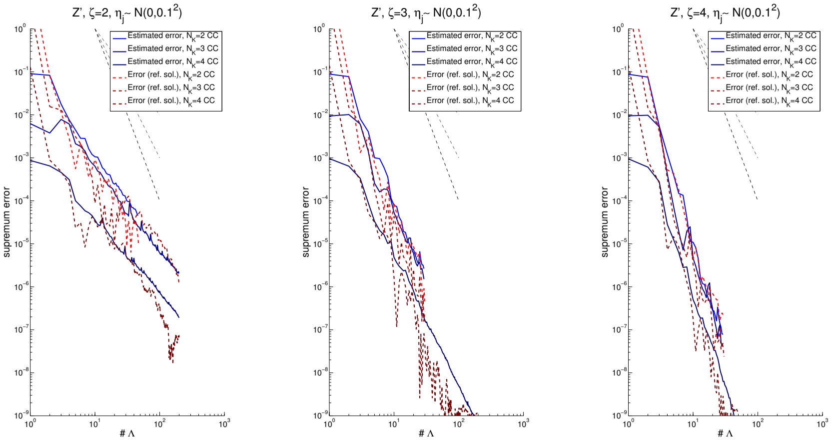

Figure: Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z' with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

Conclusions and Outlook

- New class of sparse, adaptive quadrature methods for Bayesian inverse problems for a broad class of operator equations
- Dimension-independent convergence rates depending only on the summability of the parametric operator
- Numerical confirmation of the predicted convergence rates

Conclusions and Outlook

- New class of sparse, adaptive quadrature methods for Bayesian inverse problems for a broad class of operator equations
- Dimension-independent convergence rates depending only on the summability of the parametric operator
- Numerical confirmation of the predicted convergence rates

- Gaussian priors and lognormal coefficients
- Adaptive control of the discretization error of the forward problem with respect to the expected significance of its contribution to the Bayesian estimate
- Efficient treatment of large sets of data δ

References

- Calvi J-P and Phung Van M 2011 On the Lebesgue constant of Leja sequences for the unit disk and its applications to multivariate interpolation *Journal of Approximation Theory* **163** 608-622
- Calvi J-P and Phung Van M 2012 Lagrange interpolation at real projections of Leja sequences for the unit disk *Proceedings of the American Mathematical Society* **140** 4271–4284
- CHKIFA A 2013 On the Lebesgue constant of Leja sequences for the unit disk *Journal of Approximation Theory* **166** 176-200
- Chkifa A, Cohen A and Schwab C 2013 High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs *Foundations of Computational Mathematics* 1-33
- Cohen A, DeVore R and Schwab C 2011 Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs *Analysis and Applications* **9** 1-37
- Gerstner T and Griebel M 2003 Dimension-adaptive tensor-product quadrature *Computing* **71** 65-87
- Hansen M and Schwab C 2013 Sparse adaptive approximation of high dimensional parametric initial value problems *Vietnam J. Math.* **1** 1-35
- Hansen M, Schillings C and Schwab C 2012 Sparse approximation algorithms for high dimensional parametric initial value problems *Proceedings 5th Conf. High Performance Computing, Hanoi*
- Schillings C and Schwab C 2013 Sparse, adaptive Smolyak algorithms for Bayesian inverse problems *Inverse Problems* **29** 065011
- Schillings C and Schwab C 2013 Sparsity in Bayesian inversion of parametric operator equations *SAM-Report 2013-17*
- Schwab C and Stuart A M 2012 Sparse deterministic approximation of Bayesian inverse problems *Inverse Problems* **28** 045003