A Note on Sparse, Adaptive Smolyak Quadratures for Bayesian Inverse Problems

C. Schillings

Sparse, Adaptive Smolyak Quadratures for Bayesian Inverse Problems

Claudia Schillings, SAM ETHZ
(joint work with Christoph Schwab, SAM ETHZ)

Mathematisches Forschungsinstitut Oberwolfach

Numerical Methods for PDE Constrained Optimization with Uncertain Data

Based on the parametric deterministic formulation of Bayesian inverse problems with unknown input parameter from infinite dimensional, separable Banach spaces proposed in [10], we develop a practical computational algorithm whose convergence rates are provably higher than those of Monte-Carlo (MC) and Markov-Chain Monte-Carlo methods, in terms of the number of solutions of the forward problem. The focus is on linear, elliptic PDE with unknown diffusion coefficient, however, the derived convergence results are not limited to linear, elliptic PDEs: analogous results hold for forward maps of a rather wide range of mathematical models.

A basic problem in Bayesian inverse problems consists of determining the unknown diffusion coefficient $u \in X$ from given noisy observation data $\delta = \mathcal{O}(G(u)) + \eta$ (with $\eta \in \mathbb{R}^K$ representing the observation noise, $\mathcal{O} : R \mapsto \mathbb{R}^K$ bounded, linear observation operator and $G : X \to R$ forward response map from some separable Banach space X of unknown parameters into a separable Banach space R of responses) in order to compute the expectation of a quantity of interest. The proposed approach relies on a reformulation of the forward problem with unknown stochastic input data as an infinite dimensional, parametric deterministic problem. Therefore, the unknown diffusion coefficient u is assumed to admit a *parametric*

representation of the form

$$u = \bar{a} + \sum_{j \in \mathbb{J}} y_j \psi_j$$

where $y = (y_j)_{j \in \mathbb{J}}$ is an i.i.d sequence of real-valued random variables $y_j \sim \mathcal{U}[-1/2, 1/2]$, ie. the prior is given by $\mu_0(dy) := \bigotimes_{j \in \mathbb{J}} \lambda_1(dy_j)$, $\bar{a}$, $\psi_j \in X$ and $\mathbb{J}$ denotes a finite or countably infinite index set, ie. either $\mathbb{J} = \{1, 2, ..., J\}$ or $\mathbb{J} = \mathbb{N}$.

Under appropriate assumptions on the forward and observation model and the prior measure, the posterior distribution on u is absolutely continuous with respect to the prior, see [10]. The density of the posterior with respect to the prior is a Radon-Nikodym derivative that is given by an infinite dimensional version of Bayes rule. Based on this result, we are interested in computing the expectation of a *prediction function* $\phi : X \to S$ for a *Quantity of Interest (QoI)*. The expectation of QoI under the posterior (given data δ) is given by $\mathbb{E}_{\mu^\delta}[\phi(u)] := Z'/Z \in S$ with

$$Z' = \int_{[-1/2, 1/2]^{\mathbb{J}}} \Psi(y)\mu_0(dy) \quad Z = \int_{[-1/2, 1/2]^{\mathbb{J}}} \Theta(y)\mu_0(dy) \,,$$

$$\Theta(y) = \exp(-\tfrac{1}{2}|\delta - \mathcal{O}(G(u))|_\Gamma^2 \tfrac{1}{2}\Big|_{u = \bar{a} + \sum_{j \in \mathbb{J}} y_j \psi_j}, \quad \Psi(y) = \Theta(y)\phi(u)\Big|_{u = \bar{a} + \sum_{j \in \mathbb{J}} y_j \psi_j}.$$

In [10], joint analyticity of the posterior density as a function of the parameter vector $y \in U$ is proven. In particular, the estimates of the size of domains of analytic continuation which were obtained in [10] allowed to prove rates on so-called *sparse, monotone N-term polynomial chaos approximations* of the posterior density $\Psi(y)$. The resulting approximation rates are independent of the dimension, ie. of the number of active coordinates y_j in the quadrature approximations of Z and Z', and will be the basis for the presented proofs of (dimension independent) convergence rates of the adaptive Smolyak quadrature algorithms. 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 difference operators $\Delta_\nu = \bigotimes_{j \geq 1} \Delta_{\nu_j}$, $\Delta_j = Q^j - Q^{j-1}$ and $(Q^k)_{k \geq 0}$ sequence of univariate quadrature formulas, see [9] for details on the construction of the tensorized multivariate quadrature formulas.

Then, it can be proven (under appropriate assumptions on the univariate quadrature formulas and on the forward model) that sparsity in the unknown coefficient function u, i.e. if $\sum_{j=1}^\infty \|\psi_j\|_{L^\infty(D)}^\sigma < \infty$ for $0 < \sigma < 1$, implies the existence of two sequences $(\Lambda_N^1)_{N \geq 1}, (\Lambda_N^2)_{N \geq 1}$ of monotone sets $\Lambda_N^{1,2} \subset \mathcal{F}$ such that with $\#\Lambda_N^{1,2} \leq N$

$$|Z - \mathcal{Q}_{\Lambda_N^1}(\Theta)| \leq C_Z N^{-s} \,, \qquad s = \frac{1}{\sigma} - 1 \,,$$

and

$$\|Z' - \mathcal{Q}_{\Lambda_N^2}(\Psi)\|_{V^{(m)}} \leq C_{Z'} N^{-s} \,, \qquad s = \frac{1}{\sigma} - 1 \,.$$

The construction of the monotone index set $(\Lambda_N^{1,2})_{N\geq 1}$ is based on a greedy-type strategy which attempts to control the global approximation error by locally collecting indices of the current set of reduced neighbors with the largest error contributions. We will present numerical experiments based on the following model parametric elliptic boundary value problem

$$-\mathrm{div}(u\nabla p) = f \quad \text{in } D := [0,1]\,,\ p = 0 \quad \text{in } \partial D\,,$$

with $f(x) = 100 \cdot x$ and diffusion coefficient $u(x,y) = \bar{a} + \sum_{j=1}^{64} y_j \psi_j$, where $\bar{a} = 1$ and $\psi_j = \alpha_j \chi_{D_j}$ with $D_j = [(j-1)\frac{1}{64}, j\frac{1}{64}]$, $y = (y_j)_{j=1,\dots,64}$ and $\alpha_j = \frac{1.8}{j^\zeta}$, $\zeta = 2, 3, 4$ in order to numerically verify the theoretical results. Exemplarily, the approximation error of the normalization constant Z for three values of the parameter ζ controlling the sparsity of the unknown input data is shown in Figure 1.
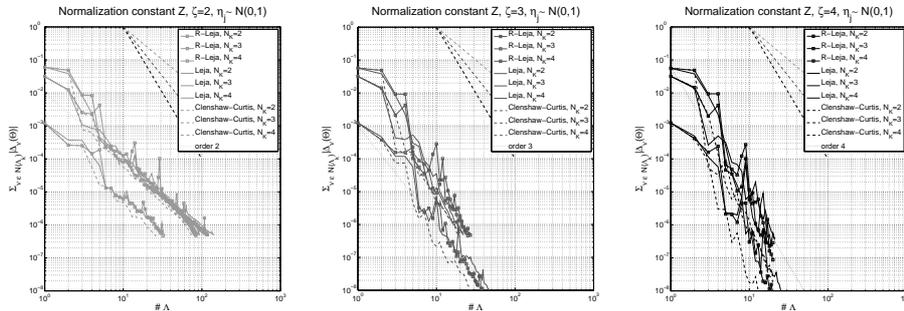

FIGURE 1.   Comparison of the error curves of the normalization constant Z with respect to $\#\Lambda_N$ based on the sequences with Clenshaw-Curtis, symmetrized Leja and R-Leja quadrature points with number of observations $K = 2^{N_K} - 1$, $N_K = 2,3,4$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

Furthermore, we will present numerical results considering a lognormal diffusion coefficient, ie. $\ln(u(x,y)) = \sum_{j=1}^{64} y_j \psi_j$, where $\psi_j = \alpha_j \chi_{D_j}$, indicating the same convergence behavior as in the uniform case, cp. Figure 2.
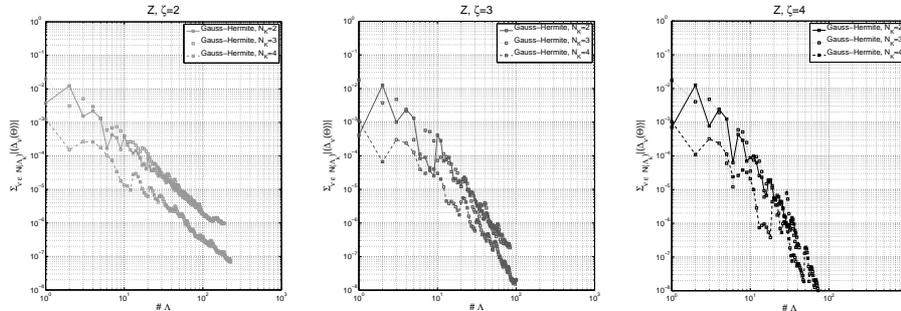
FIGURE 2. Comparison of the error curves of the normalization constant Z with respect to $\#\Lambda_N$ based on the Gauss-Hermite quadrature with number of observations $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.).

REFERENCES

[1] 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 **163** (2011), 608–622.

[2] 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, (to appear) (2012).

[3] A. Chkifa, *On the Lebesgue constant of Leja sequences for the unit disk*, (submitted) (2012).

[4] A. Chkifa, A. Cohen and C. Schwab, *High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs*, Report 2012-22, SAM, ETH Zürich (2012),

[5] A. Cohen, R. DeVore and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs*, Analysis and Applications **9** (2011), 1–37.

[6] T. Gerstner and M. Griebel, *Dimension-adaptive tensor-product quadrature* Computing **71** (2003), 65–87.

[7] M. Hansen and C. Schwab, *Analytic regularity and best N-term approximation of high dimensional parametric initial value problems*, Report 2011-64, SAM, ETH Zürich (2011).

[8] M. Hansen, C. Schillings and C. Schwab, *Sparse approximation algorithms for high dimensional parametric initial value problems*, Proceedings 5th Conf. High Performance Computing, Hanoi (2012).

[9] C. Schillings and C. Schwab, *Sparse, Adaptive Smolyak Algorithms for Bayesian Inverse Problems*, (submitted) (2012).

[10] C. Schwab and A. M. Stuart, *Sparse Deterministic Approximation of Bayesian Inverse Problems*, *Inverse Problems* **28** (2012), 045003.

Recent Research Reports

Nr.	Authors/Title

2012-39   A. Buffa and G. Sangalli and Ch. Schwab
           Exponential convergence of the hp version of isogeometric analysis in 1D

2012-40   D. Schoetzau and C. Schwab and T. Wihler and M. Wirz
           Exponential convergence of hp-DGFEM for elliptic problems in polyhedral domains

2012-41   M. Hansen
           n-term approximation rates and Besov regularity for elliptic PDEs on polyhedral
           domains

2012-42   C. Gittelson and R. Hiptmair
           Dispersion Analysis of Plane Wave Discontinuous Galerkin Methods

2012-43   J. Waldvogel
           Jost Bürgi and the discovery of the logarithms

2013-01   M. Eigel and C. Gittelson and C. Schwab and E. Zander
           Adaptive stochastic Galerkin FEM

2013-02   R. Hiptmair and A. Paganini and M. Lopez-Fernandez
           Fast Convolution Quadrature Based Impedance Boundary Conditions

2013-03   X. Claeys and R. Hiptmair
           Integral Equations on Multi-Screens

2013-04   V. Kazeev and M. Khammash and M. Nip and C. Schwab
           Direct Solution of the Chemical Master Equation using Quantized Tensor Trains

2013-05   R. Kaeppeli and S. Mishra
           Well-balanced schemes for the Euler equations with gravitation