Scaling Limits in Computational Bayesian Inversion

Claudia Schillings, Christoph Schwab

Warwick Centre for Predictive Modelling - 23. April 2015

THE UNIVERSITY OF WARWICK

SAM

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

research supported by the European Research Council under FP7 Grant AdG247277

erc
European Research Council
Established by the European Commission

Outline

Inverse Problem

Physical Model

$$G(u) \to \delta$$

- u parameter vector / parameter function
- G the forward map modelling the physical process
- δ result / observations

Forward Problem

Find the output δ for given parameters u

$\to$ **well-posed**

Inverse Problem

Find the parameters u from (noisy) observations δ

$\to$ **ill-posed**

Inverse Problem

Physical Model

$$\mathcal{G}(u) \to \delta$$

- u parameter vector / parameter function
- G the forward map modelling the physical process
- δ result / observations
- $\mathcal{G}$ forward response operator

Forward Problem

Find the output δ for given parameters u

$\rightarrow$ **well-posed**

Inverse Problem

Find the parameters u from (noisy) observations δ

$\rightarrow$ **ill-posed**

Inverse Problem

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

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

Deterministic optimization problem

$$\min_u \frac{1}{2} \|\delta - \mathcal{G}(u)\|^2 + R(u)$$

- $\|\delta - \mathcal{G}(u)\|$ potential / data misfit
- R regularization term

Inverse Problem

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

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

Deterministic optimization problem

$$\min_u \frac{1}{2} \|\delta - \mathcal{G}(u)\|^2 + R(u)$$

- **Large-scale, deterministic optimization problem**
- **No quantification of the uncertainty in the unknown $\mathbf{u}$**
- **Proper choice of the regularization term $\mathbf{R}$**

Inverse Problem

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

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

Bayesian inverse problem

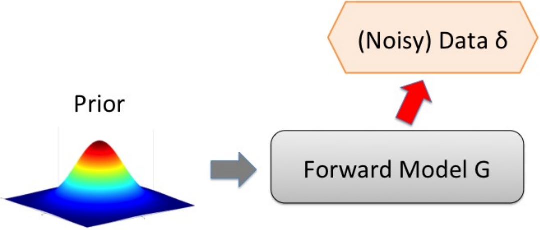
- u, η, δ random variables / fields
- Goal of computation: moments of system quantities under the posterior w.r. to noisy data δ

Inverse Problem

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

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

Bayesian inverse problem

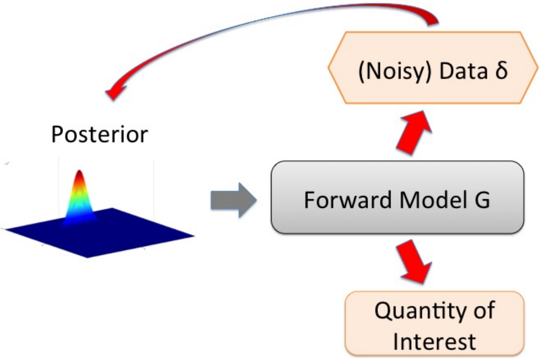


Inverse Problem

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

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

Bayesian inverse problem



Inverse Problem

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

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

Bayesian inverse problem

- **Quantification of uncertainty in $\mathbf{u}$ and system quantities**
- **Well-posedness of the inverse problem**
- **Incorporation of prior knowledge on the uncertain data $\mathbf{u}$**
- **Need of efficient approximations of the posterior**

Inverse Problem

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

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

Bayesian inverse problem

- MCMC methods
- MAP estimators
- Ad hoc algorithms (EnKF, SMC, GP,...)
- Sparse deterministic approximations of posterior expectations

Bayesian Inverse Problems (Stuart 2010)

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

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

Goal: Efficient estimation of system quantities from noisy observations

- Infinite-dimensional parameter space
- Fast convergence by exploiting sparsity of the underlying forward problem
- Suitable for application to a broad class of forward problems

Bayesian Inverse Problems (Stuart 2010)

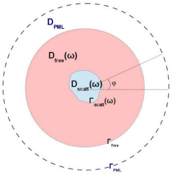
Find the unknown data $u \in X$ from noisy observations

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

Goal: Efficient estimation of system quantities from noisy observations

- Infinite-dimensional parameter space
- Fast convergence by exploiting sparsity of the underlying forward problem
- Suitable for application to a broad class of forward problems

UQ in Nano Optics



Bayesian Inverse Problems (Stuart 2010)

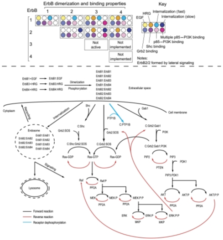
Find the unknown data $u \in X$ from noisy observations

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

Goal: Efficient estimation of system quantities from noisy observations

UQ in Biochemical Networks

- Infinite-dimensional parameter space
- Fast convergence by exploiting sparsity of the underlying forward problem
- Suitable for application to a broad class of forward problems



Source: Chen et al.

Bayesian Inverse Problems (Stuart 2010)

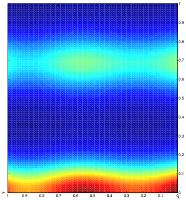
Find the unknown data $u \in X$ from noisy observations

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

Goal: Efficient estimation of system quantities from noisy observations

- Infinite-dimensional parameter space
- Fast convergence by exploiting sparsity of the underlying forward problem
- Suitable for application to a broad class of forward problems

UQ in Groundwater Flow



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

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

with $A(u;\cdot) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$, $F : X \mapsto \mathcal{Y}'$, $\mathcal{X}$, $\mathcal{Y}$ 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

Model Parametric Elliptic Problem

$$\text{Given } u \in X, \text{ find } q \in \mathcal{X}: \quad -\nabla \cdot (u \nabla q) = f \quad \text{in} \quad D, \qquad q = 0 \text{ in } \partial D$$

with weak formulation $\int_D u(x)\nabla q(x) \cdot \nabla w(x) dx = {}_{\mathcal{X}}\langle w, f\rangle_{\mathcal{X}'}$ for all $w \in \mathcal{X}$,
$\mathcal{X} = \mathcal{Y} = H_0^1(D)$, $D \subset \mathbb{R}^d$ bounded Lipschitz domain with Lipschitz boundary ∂D

- $\mathcal{O} : \mathcal{X} \mapsto \mathbb{R}^K$ 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 $\Phi : X \times \mathbb{R}^K \to \mathbb{R}$

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

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

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

- $\boldsymbol{y} = (y_j)_{j \in \mathbb{J}}$ iid 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

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

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

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 $\boldsymbol{y}\in U$ given δ, is absolutely continuous with respect to $\mu_0(d\boldsymbol{y})$, and

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

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

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

Bayesian Inverse Problem

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

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

with $Z = \int_U \exp(-\frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\right))\mu_0(d\mathbf{y})$.

- 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

Bayesian Inverse Problem

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

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

with $Z = \int_U \exp(-\frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\right))\mu_0(d\mathbf{y})$.

Exploiting sparsity in the parametric operator equation

- Parameters belonging to a specified sparsity class
- Analytic regularity of the parametric, deterministic Bayesian posterior
- Parametric, deterministic Bayesian posterior belongs to the same sparsity class

$\to$ **Sparsity of generalized pce + dimension-independent convergence rates for Smolyak integration algorithms**

Model Parametric Elliptic Problem

$$\int_D u(x)\nabla q(x)\cdot\nabla w(x)dx = \ _{\mathcal{X}}\langle w,f\rangle_{\mathcal{X}'} \quad \text{for all} \quad w\in\mathcal{X}$$

with $D\subset\mathbb{R}^d$ bounded Lipschitz domain with Lipschitz boundary ∂D, $\mathcal{X}=H_0^1(D), f\in L^2(D)$.

Structural Assumptions

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

with $\langle u\rangle \in L^\infty(D)$ and $(\psi_j)_{j\in\mathbb{J}}\subset L^\infty(D), \boldsymbol{y}=(y_1,y_2,\dots)\in U=[-1,1]^{\mathbb{J}}$

Assumption 1 (UEA)

$$0<u_{\min}\le u(x,y)\le u_{\max}<\infty \quad \forall x\in D, y\in U$$

with $0<u_{\min}\le u_{\max}<\infty$.

Assumption 2

$$\sum_{j\in\mathbb{J}}\|\psi_j\|_{L^\infty(D)} \le \frac{\kappa}{1+\kappa}\langle u\rangle_{\min}$$

with $\langle u\rangle_{\min} = \min_{x\in D}\langle u\rangle(x) > 0$ and $\kappa>0$

Sparse Polynomial Chaos Approximations

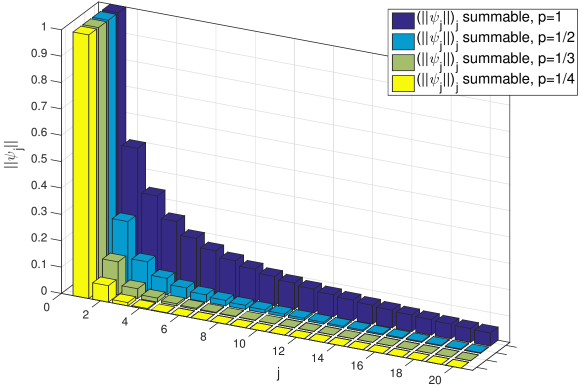
Sparsity in the unknown coefficient function u,

i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$ for $0 < p < 1$,

Sparse Polynomial Chaos Approximations

Sparsity in the unknown coefficient function u,

i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$ for $0 < p < 1$,



Sparse Polynomial Chaos Approximations

Sparsity in the unknown coefficient function u,

i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$ for $0 < p < 1$,

implies p-sparsity in the solution's Taylor expansion,

$\forall y \in U : \quad q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} \tau_\nu \boldsymbol{y}^\nu \ , \quad \tau_\nu := \frac{1}{\nu!} \partial_{\boldsymbol{y}}^\nu q(\boldsymbol{y}) \ |_{\boldsymbol{y}=0} \ , \quad \mathcal{F} = \{\nu \in \mathbb{N}_0^{\mathbb{N}} : |\nu| < \infty\}$:

Sparse Polynomial Chaos Approximations

Sparsity in the unknown coefficient function u,

i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$ for $0 < p < 1$,

implies p-sparsity in the solution's Taylor expansion,

$\forall y \in U : \quad q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} \tau_\nu \boldsymbol{y}^\nu \ , \quad \tau_\nu := \frac{1}{\nu!} \partial_{\boldsymbol{y}}^\nu q(\boldsymbol{y}) \mid_{\boldsymbol{y}=0} \ , \quad \mathcal{F} = \{\nu \in \mathbb{N}_0^{\mathbb{N}} : |\nu| < \infty\}$:

There exists a p-summable, monotone envelope $\mathbf{t} = \{\mathbf{t}_\nu\}_{\nu \in \mathcal{F}}$

i.e. $\mathbf{t}_\nu := \sup_{\mu \geq \nu} \|\tau_\nu\|_{\mathcal{X}}$ with $C(p, \boldsymbol{q}) := \|\mathbf{t}\|_{\ell^p(\mathcal{F})} < \infty$.

Sparse Polynomial Chaos Approximations

Sparsity in the unknown coefficient function u,

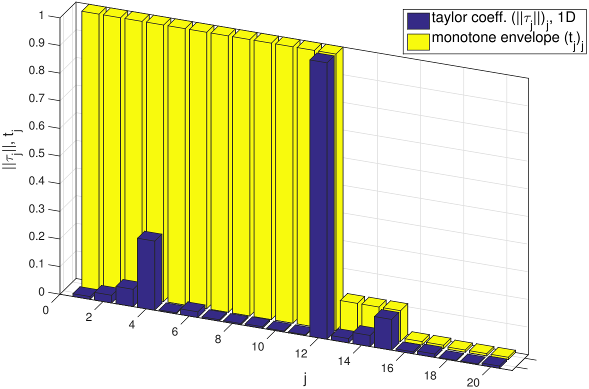
i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|^p_{L^\infty(D)} < \infty$ for $0 < p < 1$,

imp

$\forall y \in$ … $\infty\}$:

The

i.e. …



Sparse Polynomial Chaos Approximations

Sparsity in the unknown coefficient function u,

i.e. if $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}^p < \infty$ for $0 < p < 1$,

implies p-sparsity in the solution's Taylor expansion,

$\forall y \in U: \quad q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} \tau_\nu \boldsymbol{y}^\nu \ , \quad \tau_\nu := \frac{1}{\nu!} \partial_{\boldsymbol{y}}^\nu q(\boldsymbol{y}) \mid_{\boldsymbol{y}=0} \ , \quad \mathcal{F} = \{\nu \in \mathbb{N}_0^{\mathbb{N}} : |\nu| < \infty\}$:

There exists a p-summable, monotone envelope $\mathbf{t} = \{\mathbf{t}_\nu\}_{\nu \in \mathcal{F}}$

i.e. $\mathbf{t}_\nu := \sup_{\mu \geq \nu} \|\tau_\nu\|_{\mathcal{X}}$ with $C(p, \boldsymbol{q}) := \|\mathbf{t}\|_{\ell^p(\mathcal{F})} < \infty$.

and monotone $\Lambda_N^T \subset \mathcal{F}$ corresponding to the N largest terms of $\mathbf{t}$ with

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

A. Cohen, R. DeVore and Ch. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs, *Analysis and Applications*, 2010.

Sparse Polynomial Chaos Approximations

Idea of Proof

Holomorphic Extension

$$z \mapsto q(z)\;, \qquad z \in \mathcal{U}_\rho := \otimes_{j\geq 1}\{z_j \leq \rho_j\}$$

holomorph with bound $\|q(z)\|_{\mathcal{X}} \leq C_\delta$ for any positive sequence $\rho = \{\rho_j\}_{j\geq 1}$ such that $\sum_{j\geq 1}\rho_j|\psi_j(x)| \leq \langle u\rangle(x) - \delta,\ x \in D$ for some $\delta > 0$.

Estimate on the Taylor Coefficients

$$\|\tau_\nu\|_{\mathcal{X}} \leq C_\delta \inf\{\rho^{-\nu} : \sum_{j\geq 1}\rho_j|\psi_j(x)| \leq \langle u\rangle(x) - \delta,\ x \in D\}$$

by recursive application of Cauchy's formula

Construction of a particular ρ

$$\{\|\psi_j\|_{L^\infty}\}_{j\geq 1} \in \ell^p(\mathbb{N}) \Rightarrow \{\mathbf{t}_\nu\}_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$$

Stechkin

$$\sup_{\mathbf{y}\in U}\left\|q(\mathbf{y}) - \sum_{\nu\in\Lambda_N^T}\tau_\nu \mathbf{y}^\nu\right\|_{\mathcal{X}} \leq \sum_{\nu\notin\Lambda_N^T}\|\mathbf{t}_\nu\|_{\mathcal{X}} \leq CN^{-\frac{1}{p}+1}$$

Sparsity of the Posterior

Theorem (CIS and Schwab 2013)

Assume that the unknown coefficient function u is p-sparse for $0 < p < 1$. Then, the posterior density's Taylor expansion is p-sparse with the same p.

Sparsity of the Posterior

Theorem (CIS and Schwab 2013)

Assume that the unknown coefficient function u is p-sparse for $0 < p < 1$. Then, the posterior density's Taylor expansion is p-sparse with the same p.

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

Sparsity of the Posterior

Theorem (CIS and Schwab 2013)

Assume that the unknown coefficient function u is p-sparse for $0 < p < 1$. Then, the posterior density's Taylor expansion is p-sparse with the same p.

Examples

- Parametric initial value ODEs (Hansen & Schwab; Vietnam J. Math. 2013)
- Affine-parametric, linear operator equations (CIS & Schwab; 2013)
- Semilinear elliptic PDEs (Hansen & Schwab; Math. Nachr. 2013)
- Elliptic multiscale problems (Hoang & Schwab; Analysis and Applications 2012)
- Nonaffine holomorphic-parametric, nonlinear problems (Cohen, Chkifa & Schwab; 2013)

Sparsity of the Posterior

Theorem (CIS and Schwab 2013)

Assume that the unknown coefficient function u is p-sparse for $0 < p < 1$. Then, the posterior density's Taylor expansion is p-sparse with the same p.

Goal of computation

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

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

with $Z = \int_U \exp(-\frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\right)\mu_0(d\boldsymbol{y})$.

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

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

- $(Q^k)_{k \geq 0}$ sequence of univariate quadrature formulas
- $\Delta_j = Q^j - Q^{j-1}$, $j \geq 0$ univariate quadrature difference operator
- $\mathcal{Q}_\nu = \bigotimes_{j \geq 1} Q^{\nu_j}$, $\Delta_\nu = \bigotimes_{j \geq 1} \Delta_{\nu_j}$ tensorized multivariate operators

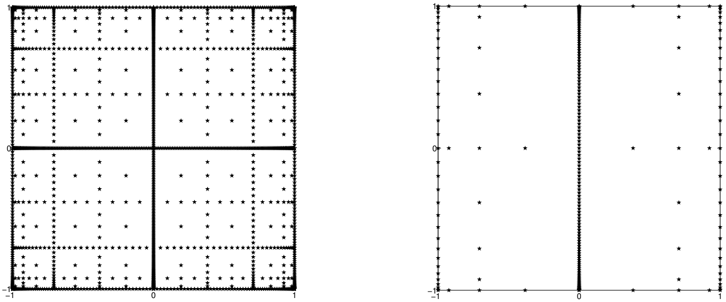
Sparse Quadrature Operator

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

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



Convergence Rates for Adaptive Smolyak Integration

Theorem

Assume that the unknown coefficient function u is p-sparse for $0 < p < 1$.

Then there exists a sequence $(\Lambda_N)_{N\geq 1}$ of monotone index sets $\Lambda_N \subset \mathcal{F}$ such that $\#\Lambda_N \leq N$ and

$$|Z - \mathcal{Q}_{\Lambda_N}[\Theta]| \leq C^1 N^{-\frac{1}{p}+1},$$

with $Z = \int_U \Theta(\boldsymbol{y})\mu_0(d\boldsymbol{y})$, $\Theta(\boldsymbol{y}) = \exp\big(-\Phi(u;\delta)\big)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j}$ and,

$$\|Z' - \mathcal{Q}_{\Lambda_N}[\Psi]\|_{\mathcal{X}} \leq C^2 N^{-\frac{1}{p}+1},$$

with $Z' = \int_U \Psi(\boldsymbol{y})\mu_0(d\boldsymbol{y})$, $\Psi(\boldsymbol{y}) = \Theta(\boldsymbol{y})\phi(u)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j}$ $C^1, C^2 > 0$ independent of N.

Remark: SAME index sets Λ_N for BOTH, Z' and Z.

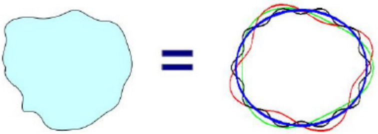
ClS and Schwab Sparsity in Bayesian Inversion of Parametric Operator Equations, 2013.

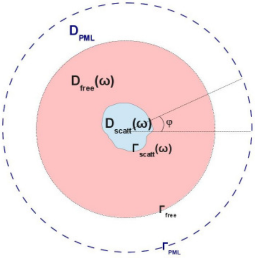
Uncertainty Quantification in Nano Optics

Goal: Quantification of the influence of defects in fabrication process on the optical response of nano structures

- Propagation of plane wave and its interaction with scatterer described by Helmholtz equation (2D).
- Stochastic shape of the scatterer

$$0 < \rho_{\min} \leq \rho(\omega, \phi) \leq \rho_{\max}, \quad \omega \in \Omega,\ \phi \in [0, 2\pi)$$





Collaboration with Ralf Hiptmair, Laura Scarabosio

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

Uncertainty Quantification in Nano Optics

Forward Problem

$$
\begin{aligned}
-\Delta u - k_1^2 u &= 0 \quad \text{in } D_1(\omega) \\
-\Delta u - k_2^2 u &= 0 \quad \text{in } D_2(\omega) \\
u|_+ - u|_- &= 0 \quad \text{on } \Gamma(\omega) \\
\left.\frac{\partial u}{\partial \mathbf{n}}\right|_+ - \mu \left.\frac{\partial u}{\partial \mathbf{n}}\right|_- &= 0 \quad \text{on } \Gamma(\omega)
\end{aligned}
$$

$$
\lim_{|x|\to\infty} \sqrt{|x|}\left(\frac{\partial}{\partial |x|} - ik_1\right)(u(\omega) - u_i)(x) = 0, \qquad u_i(x) = e^{ik_1 \cdot x}
$$

Parametrization of the shape

$$
r(\omega,\varphi) = r_0(\varphi) + \sum_{k=1}^{64} \frac{1}{k^{\zeta}} y_{2k}(\omega)\cos(k\varphi) + \frac{1}{k^{\zeta}} y_{2k+1}(\omega)\sin(k\varphi), \quad \varphi \in [0, 2\pi)
$$

Numerical Results of Sparse Interpolation

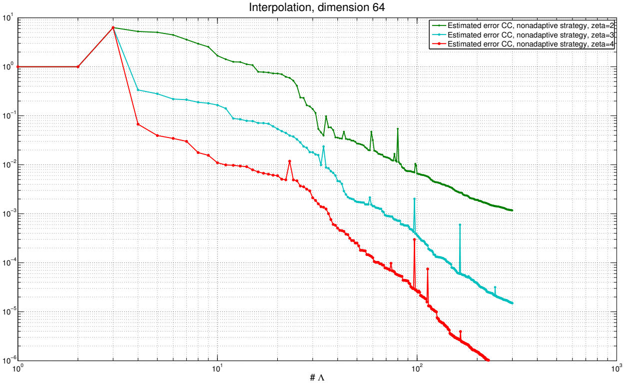


Figure: Estimated error curves of the interpolation error w.r. to the cardinality of the index set Λ_N based on the sequences CC with nonadaptive refinement, uniform distribution of the parameters, dimension of the parameter space 64.

Concentration Effect of the Posterior for $0 < \Gamma \ll 1$

Model parametric elliptic problem

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

with $f(x) = 100 \cdot x$ and

$$u(x,y) = 0.15 + y_1\psi_1 + y_2\psi_2\,,$$

with $J = 2$, $\mathbb{J} = \{1,2\}$, $\psi_1(x) = 0.1\sin(\pi x)$, $\psi_2(x) = 0.025\cos(2\pi x)$ and with $y_j \sim \mathcal{U}[-1,1]$, $j \in \mathbb{J}$.

Concentration Effect of the Posterior for $0 < \Gamma \ll 1$

For given (noisy) data δ,

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

we are interested in the behavior of the posterior

$$\Theta(\boldsymbol{y}) = \exp\big(-\Phi_{\Gamma_{obs}}(u;\delta)\big)\Big|_{u=\sum_{j=1}^{2} y_j\psi_j},$$

with

$$\Phi_{\Gamma_{obs}}(u;\delta) = \frac{1}{2}\Big((\delta - \mathcal{G}(u))^{\top}\Gamma^{-1}(\delta - \mathcal{G}(u))\Big)\ .$$

and variance of the noise $\Gamma = \lambda \cdot id$

- QoI ϕ is the solution of the forward problem at $x = 0.5$.
- Observation operator $\mathcal{O}$ consists of 2 system responses at $x = 0.25$ and $x = 0.75$.

Concentration Effect of the Posterior for $0 < \Gamma \ll 1$

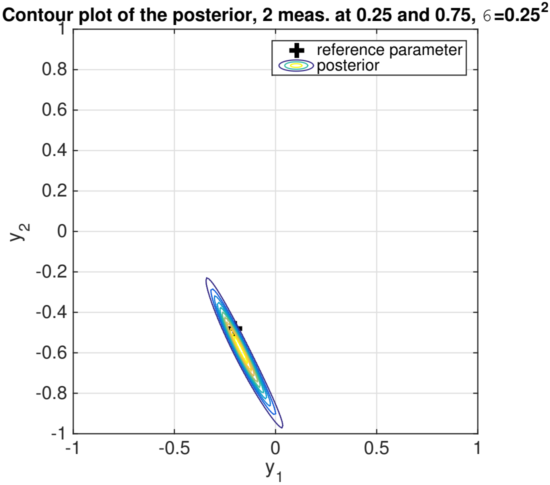


Figure: Contour plot of the posterior with observational noise $\lambda = 0.25^2$.

Concentration Effect of the Posterior for $0 < \Gamma \ll 1$

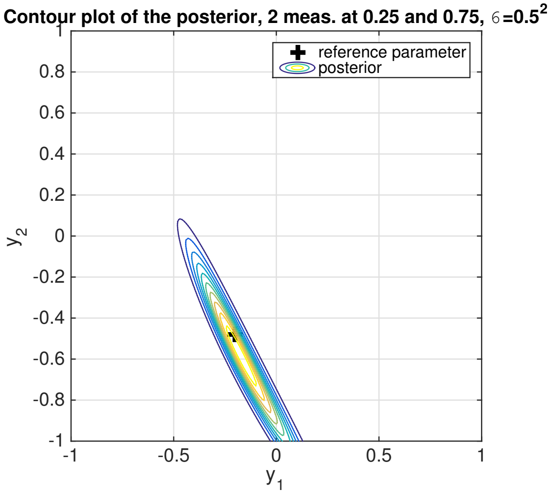


Figure: Contour plot of the posterior with observational noise $\lambda = 0.5^2$.

Concentration Effect of the Posterior for $0 < \Gamma \ll 1$

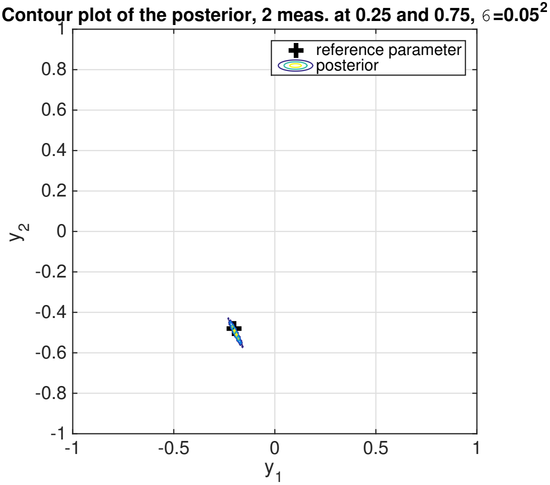


Figure: Contour plot of the posterior with observational noise $\lambda = 0.05^2$.

Asymptotic Analysis

Theorem (CIS and Schwab 2014)

Assume that the parameter space is finite (possibly after dimension - truncation) and $\mathcal{G}(\cdot)$, δ are such that the assumptions of Laplace's method hold; in particular, the minimum $\boldsymbol{y}_0$ of

$$S(y) = \frac{1}{2}\Big((\delta - \mathcal{G}(u))^\top(\delta - \mathcal{G}(u))\Big)$$

is nondegenerate.
Then, as $\Gamma \downarrow 0$, the Bayesian estimate admits an asymptotic expansion

$$\mathbb{E}^{\mu^\delta}[\phi] = \frac{Z'_\Gamma}{Z_\Gamma} \sim a_0 + a_1\lambda + a_2\lambda^2 + ...$$

where $a_0 = \phi(\boldsymbol{y}_0)$.

Curvature Rescaling Regularization

Removing the degeneracy in the integrand function

- The maximizer $\boldsymbol{y}_0$ of the posterior measure $\Theta(\boldsymbol{y})$ is computed by minimizing the potential

$$\frac{1}{2}\big((\delta - \mathcal{G}(u))^\top(\delta - \mathcal{G}(u))\big)\Big|_{u=\sum_{j=1}^{2} y_j\psi_j} .$$

  using a trust-region Quasi-Newton approach with SR1 updates.

- Diagonalize the approximated Hessian $H_{SR1} = QMQ^\top$ and regularize the integrand by the curvature rescaling transformation

$$\boldsymbol{y}_0 + \Gamma^{1/2}QM^{-1/2}\boldsymbol{z}\,, \quad \boldsymbol{z} \in \mathbb{R}^J .$$

Numerical Experiments

Curvature rescaling

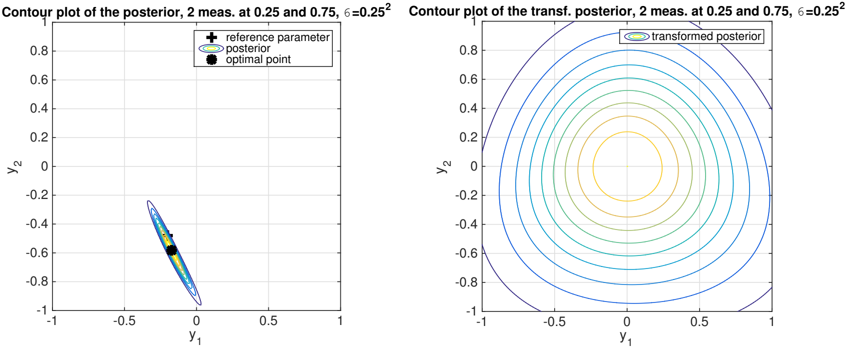


Figure: Contour plot of the posterior with observational noise $\Gamma_{obs} = 0.25^2 \cdot Id$ (left) and contour plot of the transformed posterior (right).

Numerical Experiments

Curvature rescaling

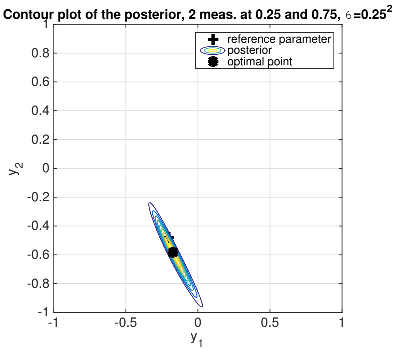


Figure: Contour plot of the posterior with observational noise $\Gamma_{obs} = 0.25^2 \cdot Id$ (left) and and truncated domain of integration of the rescaled Smolyak approach in the original coordinate system (right).

Numerical Experiments

Curvature rescaling

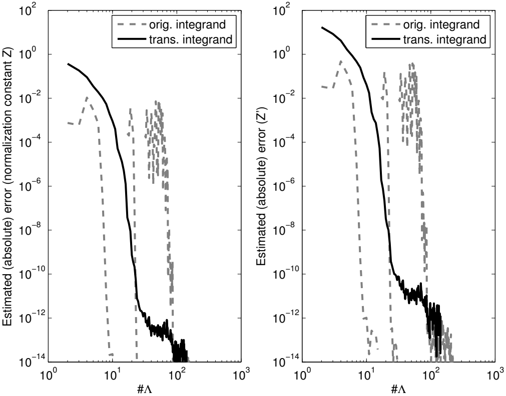


Figure: Comparison of the estimated (absolute) error curves using the Smolyak approach for the original integrand (gray) and the transformed integrand (black) for the computation of $Z_{\Gamma_{obs}}$ (left) and $Z'_{\Gamma_{obs}}$ (right) with observational noise $\Gamma_{obs} = 0.25^2 \cdot Id$.

Numerical Experiments

Curvature rescaling

Extension to lognormal case

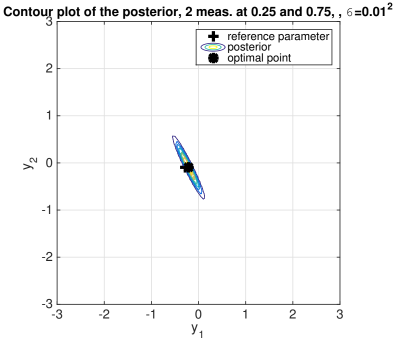
$$\ln(u(x,y)) = 0.15 + y_1\psi_1 + y_2\psi_2\,,$$

with $J = 2$, $\mathbb{J} = \{1, 2\}$, $\psi_1(x) = 0.1\sin(\pi x)$, $\psi_2(x) = 0.025\cos(2\pi x)$ and with $y_j \sim \mathcal{N}(0, 1),\ j \in \mathbb{J}$.

Numerical Experiments

Curvature rescaling

Extension to lognormal case



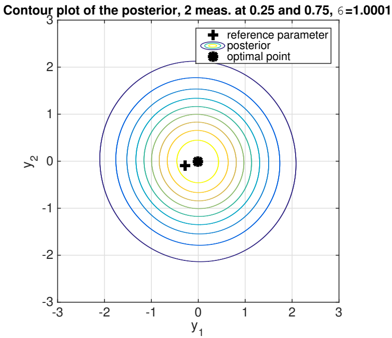


Figure: Contour plot of the posterior density with observational noise $\Gamma_{obs} = 0.01^2 \cdot Id$ (left), and contour plot of the transformed posterior (right).

Numerical Experiments

Curvature rescaling

Extension to lognormal case

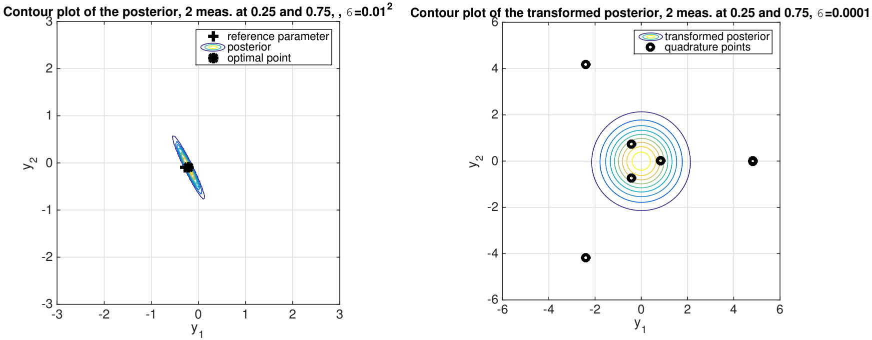


Figure: Contour plot of the posterior density with observational noise $\Gamma_{obs} = 0.01^2 \cdot Id$ (left), and contour plot of the transformed posterior (right).

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

- Efficient treatment of small observation noise variance Γ
- Development of preconditioning techniques to overcome the convergence problems in the case $\Gamma \downarrow 0$
- Combination of optimization and sampling techniques

References

- 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 2014 Sparsity in Bayesian inversion of parametric operator equations *Inverse Problems* **30** 065007
- Schillings C 2013 A Note on Sparse, Adaptive Smolyak Quadratures for Bayesian Inverse Problems *Oberwolfach Report* **10/1** 239-293
- Schillings C and Schwab C 2014 Scaling Limits in Computational Bayesian Inversion *SAM Report 2014-26*
- Schwab C and Stuart A M 2012 Sparse deterministic approximation of Bayesian inverse problems *Inverse Problems* **28** 045003