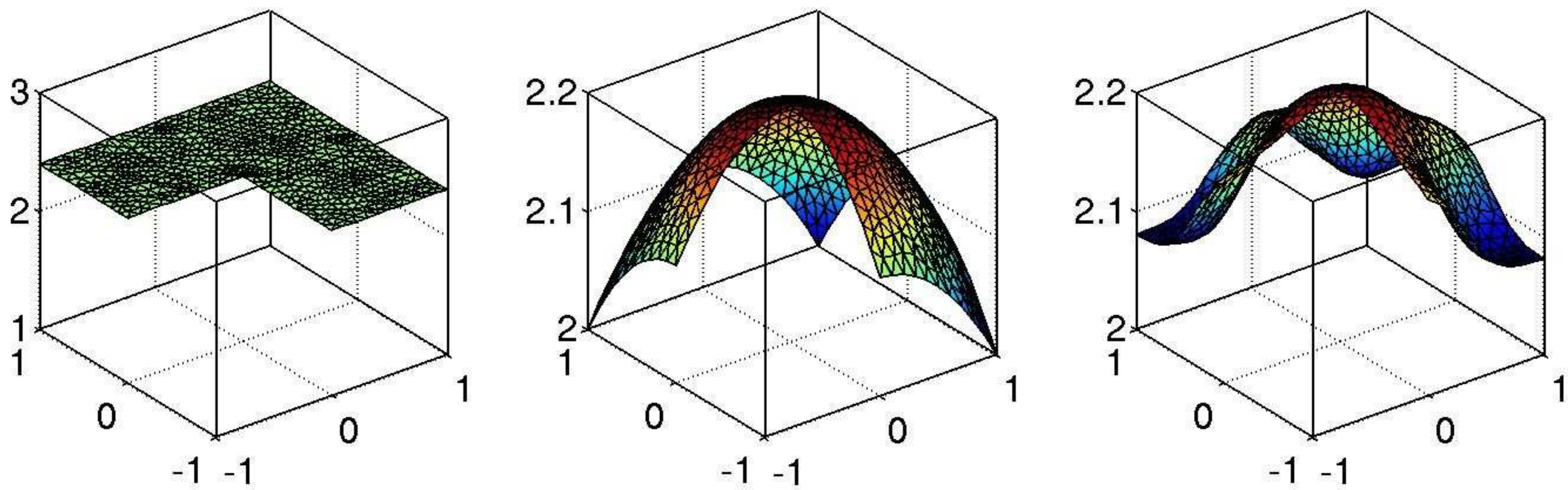
Technische Universität Braunschweig

DFG Deutsche Forschungsgemeinschaft

Institut für Wissenschaftliches Rechnen

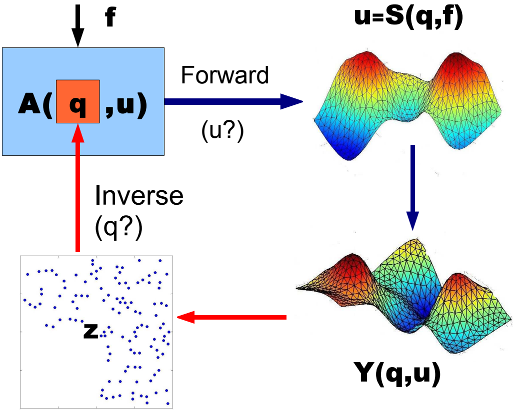


Bayesian Identification for non-Gaussian Parameters

B. Rosić,T. El-Moselhy, A. Litvinenko, O. Pajonk, H. G. Matthies

Email: bojana.rosic@tu-bs.de, ISUME 2011, Prague, Czech Republik, 3.05.2011.

Inverse Problem: **Find parameter** q **given measurement data** z



Ill-posed problem: issues of existence, uniqueness and stability

- Additional information to data z: q_f (apriori information, forecast)

What is q_f?

- classical Bayesian approach: $q_f := \pi_f$ apriori pdf

$$\pi_a(q|z) = \text{const}\ \pi_f(q)\pi(z|q) = \text{const}\ \pi_f(q)L(q)$$

- Markov Chain Monte Carlo methods (MCMC) [Gamerman 2006]
- spectral stochastic FEM +MCMC [Kučerová at all 2010, Marzouk 2009]
- collocation methods [Christen & Fox 2010]

-drawback: requires a complete statistical description of the problem

- Probability space $(\Omega, \mathcal{B}, \mathbb{P})$
- the space of RVs with finite variance $\mathcal{S} := L_2(\Omega)$ (stochastic space)
- the Hilbert space $\mathcal{Q}$ (deterministic space)

  $\mathcal{Q}$ -valued RVs form a space $\mathscr{Q} := \mathcal{Q} \otimes \mathcal{S}$

True measurement

- Linear measurement $\check{y} = Y(q, u) \in \mathcal{Y}$ is polluted by noise ϵ :

$$z = \check{y} + \epsilon, \quad \epsilon \sim N(0, \boldsymbol{C}_\epsilon) \quad \Rightarrow \quad z \in \mathscr{Y}_0 \subseteq \mathscr{Y} := \mathcal{Y} \otimes \mathcal{S}$$

Apriori information

$$q_f : \Omega \to \mathcal{Q}, \quad q_f \in \mathscr{Q}_f \subset \mathscr{Q}$$

- already defined: $z \in \mathscr{Y}_0, \quad q_f \in \mathscr{Q}_f$
- given linear mapping $H : \mathscr{Q} \to \mathscr{Y}$, predict observation

$$y = Hq_f, \quad y \in \mathscr{Q}_0 = H^*(\mathscr{Y}_0)$$

Theorem

In the setting just described, the random variable $q_a \in \mathscr{Q}$ — "a" stands for "assimilated" or "analysis" — is the orthogonal (min. variance) projection of q onto the subspace $\mathscr{Q}_f + \mathscr{Q}_0$:

$$q_a(\omega) = q_f(\omega) + K(z(\omega) - y(\omega)), \quad K := C_{q_f y}\,(C_y + C_\epsilon)^{-1}$$

with q_f being the orthogonal projection onto $\mathscr{Q}_f$ and K the "Kalman gain" operator [Luenberger 1969, Rosić at all 2011, Pajonk at all 2011].

- doesn't assume Gaussian statistics; in linear case reduces to Kalman fillter [Evensen 2009]

“Projection of Projection ”

- the orthogonal projector $\hat{P} : \mathscr{Q} \to \hat{\mathscr{Q}}, \quad \hat{P}^* = \hat{P}$

$$\hat{\mathscr{Q}} := \mathcal{Q}_N \otimes \mathcal{S}_J$$

- project onto $\hat{\mathscr{Q}}$

$$\begin{aligned}\hat{q}_a(\omega) = \hat{P}q_a(\omega) &= \hat{P}(q_f(\omega) + K(z(\omega) - y(\omega))) \\ &= \hat{P}q_f(\omega) + \hat{P}K(z(\omega) - \hat{y}(\omega)) \\ &= \hat{q}_f(\omega) + K(\hat{z}(\omega) - \hat{y}(\omega)),\end{aligned}$$

where $\hat{y}(\omega) = H\hat{P}q_f(\omega) = H\hat{q}_f(\omega)$

- Darcy Law

$$-\operatorname{div}(\kappa(x,\omega)\nabla u(x,\omega)) = f(x,\omega),$$
$$u(x,\omega) = 0.$$

- Conductivity is for simplicity assumed to be scalar field with apriori distribution (via maximum entropy principle)

$$\kappa_f(x) := \exp(q_f(x)), \quad q_f(x) \sim N(\mu_{q_f}, \sigma^2_{q_f})$$

- Covariance function

$$\mathrm{Cov}_{q_f}(x,y) = \sigma^2_{q_f} \exp(-|x-y|/l_c)$$

- following conditions hold:

$$\kappa_f(x,\omega) > 0, \quad \|\kappa_f\|_{L_\infty(\mathcal{G}\times\Omega)} < \infty, \quad \|1/\kappa_f\|_{L_\infty(\mathcal{G}\times\Omega)} < \infty.$$

- The solution space:

$$\mathscr{U} := \mathcal{U} \otimes \mathcal{S}, \quad \mathcal{U} := \mathring{H}^1(\mathcal{G}) = \{u \in H^1(\mathcal{G}) \mid u = 0 \text{ on } \partial\mathcal{G}\}$$

- Euqilibrium equation:

$$\mathbf{a}(v,u) := \mathbb{E}\left(\mathrm{a}(\omega)(v(\cdot,\omega), u(\cdot,\omega))\right) = \mathbb{E}\left(\langle \ell(\omega), v(\cdot,\omega)\rangle\right) =: \langle\langle \boldsymbol{\ell}, v \rangle\rangle.$$

$$\mathrm{a}(\omega)(v,u) := \int_{\mathcal{G}} \nabla v(x) \cdot (\kappa_f(x,\omega)\nabla u(x)) \; \mathrm{d}x,$$

$$\langle \ell(\omega), v\rangle := \int_{\mathcal{G}} v(x) f(x,\omega) \, \mathrm{d}x, \quad \forall v \in \mathcal{U},$$

- The well-possednes via Lax-Milgram theorem.

- Finite element discretisation: $u(x,\omega) = \sum_{n=1}^{N} u_n(\omega)\phi_n(x)$

$$\boldsymbol{A}(\omega)[\boldsymbol{u}(\omega)] = \boldsymbol{f}(\omega)$$

- Wiener's polynomial chaos expansion: $u_n(\boldsymbol{\theta}) = \sum_{\alpha\in\mathcal{J}} u_n^{\alpha} H_{\alpha}(\boldsymbol{\theta}(\omega))$

$$\mathbb{E}\left([\boldsymbol{f}(\boldsymbol{\theta}) - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{u}(\boldsymbol{\theta})]H_{\beta}(\boldsymbol{\theta})\right) = 0.$$

- The Karhunen-Loève expansion (KLE) of stiffness and rhs

$$\mathbf{Au} := \Big(\sum_{j=0}^{\infty} \boldsymbol{A}_j \otimes \boldsymbol{\Delta}^j\Big)\Big(\sum_{\alpha\in\mathcal{J}} \boldsymbol{u}^{\alpha} \otimes \boldsymbol{e}^{\alpha}\Big) = \Big(\sum_{\alpha\in\mathcal{J}} \boldsymbol{f}_{\alpha} \otimes \boldsymbol{e}^{\alpha}\Big) =: \mathbf{f},$$

where $\boldsymbol{\Delta}^j = \mathbb{E}(H_{\alpha}\xi_j H_{\beta})$, $\kappa_f = \sum_{j=1}^{M} \kappa_f^j \xi_j$ and $|\mathcal{J}| = R$.

- The sparse tensor Galerkin methods [Zander at all 2010]

- Measure some functional of the solution u in finitely many patches L:

$$\hat{\mathcal{G}} := \{x_1, ..., x_L\} \subset \mathcal{G}, \quad L := |\hat{\mathcal{G}}|.$$

- The average hydraulic head:

$$y(u,\omega) := [..., y(x_j), ...] \in \mathbb{R}^L, \quad y(x_j) = \int_{\mathcal{G}_j} u(x,\omega) \mathrm{d}x,$$

$$\check{\boldsymbol{y}} = [y(x_1,\check{\omega}), ..., y(x_L,\check{\omega})]^T$$

- Observation:

$$\boldsymbol{z} := \check{\boldsymbol{y}} + \epsilon, \quad \epsilon \sim N(0, \boldsymbol{C}_\epsilon)$$

- κ_f is cone in the vector space of RVs (not subspace)

- project: $\kappa_f = \sum_{\alpha \in \mathcal{J}} \kappa_f^{(\alpha)} H_\alpha(\boldsymbol{\theta}(\omega))$ (similar for z and y)

- map to a Lie algebra:

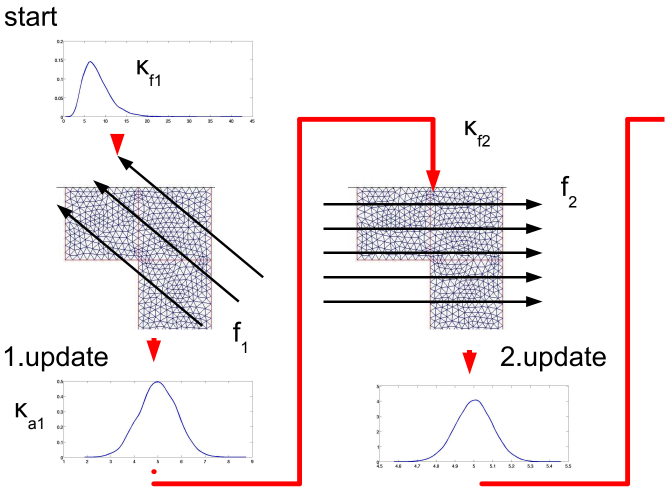
$$q_f(x, \omega) = \log \kappa_f = \sum_{\alpha \in \mathcal{J}} q_f^{(\alpha)} H_\alpha(\boldsymbol{\theta}(\omega)) = \boldsymbol{Q}_f \mathbf{H}, \quad \boldsymbol{Q}_f \in \mathbb{R}^{N \times R}$$

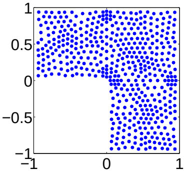
- matrix form of update formula:

$$\boldsymbol{Q}_a = \boldsymbol{Q}_f + \boldsymbol{K}(\boldsymbol{Z} - \boldsymbol{Y}), \quad \boldsymbol{K} \in \mathbb{R}^{N \times L}; \quad \boldsymbol{Z}, \boldsymbol{Y} \in \mathbb{R}^{L \times R}$$

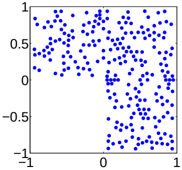
- map back

$$\kappa_a = \exp(q_a(x, \omega))$$

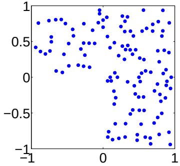

start
Kf1
f1
1.update
Ka1
Kf2
f2
2.update



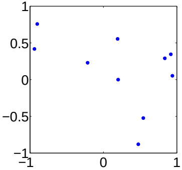
a) 447 measurement patches



b) 239 measurement patches



c) 120 measurement patches



d) 10 measurement patches

Table: Position of measurement points (FEM nodes) used in the experiments

- Right hand side: $f = f_0 \sin(\frac{2\pi}{\lambda} \boldsymbol{x}^T \boldsymbol{d} + \varphi)$

$$\boldsymbol{d} = [\cos\alpha \; \sin\alpha], \quad \alpha \in [-\pi/2, \pi/2], \quad \varphi \in [0, 2\pi]$$

- ' Virtual truth' is taken as

$$\begin{aligned} \text{a) } \kappa &= 2 \\ \text{b) } \kappa &= 2 + 0.3 \cdot (x + y) \\ \text{c) } \kappa &= 2.2 - 0.1 \cdot (x^2 + y^2) \end{aligned}$$

- Apriori information:

$$\mathbb{E}(\kappa) = 2.4, \quad \sigma_\kappa = 0.4$$

order of PCE $p = 3$ and number of KLE modes: $M <= 50$

Experiment	L	ε_p	1st	2nd	3rd	4th
1.	477	0.45	0.08	0.04	0.03	0.03
2.	239	0.45	0.08	0.05	0.05	0.04
3.	120	0.45	0.07	0.05	0.05	0.04
4.	60	0.45	0.07	0.06	0.05	0.05
5.	10	0.45	0.13	0.08	0.07	0.07

Table: "Constant truth": Decay of the relative error ε_a in each experiment

$$\varepsilon_a := \frac{\|\kappa_a - \kappa_t\|_{L_2(\Omega\otimes\mathcal{G})}}{\|\kappa_t\|_{L_2(\Omega\otimes\mathcal{G})}}; \quad \bar{\varepsilon}_a := \frac{|\mathbb{E}(\kappa_a) - \mathbb{E}(\kappa_t)|}{|\mathbb{E}(\kappa_t)|}$$

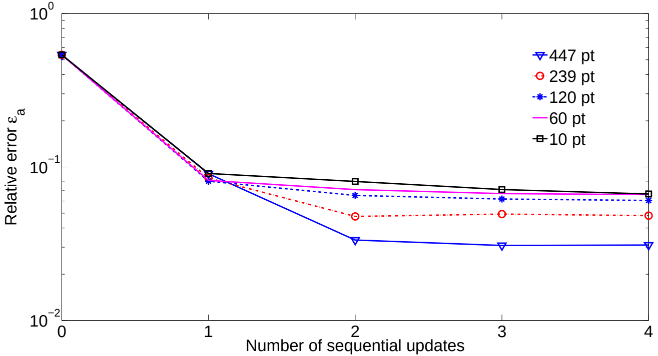


Figure: “Linear truth”, experiment 1 (L=447): Convergence behaviour of the relative error ε_a with respect to the number of sequential updates and measurement points

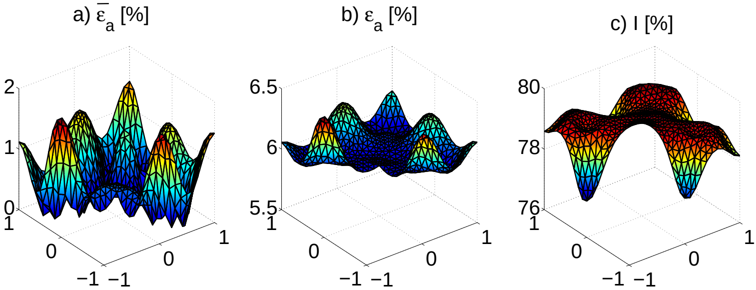


Figure: "Constant truth", experiment 1 (L=447) after 4th update: a) Relative error $\bar{\varepsilon}_a$ (the mean of the posterior compared to the mean of the truth) b) relative error ε_a (the posterior compared to the truth) c) improvement I (the posterior compared to the prior)

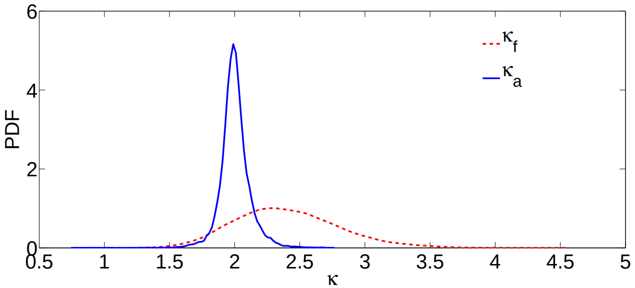


Figure: **"Constant truth", experiment 3 (L=120)**: Posterior probability density function κ_a compared to the prior κ_f for a single point in domain

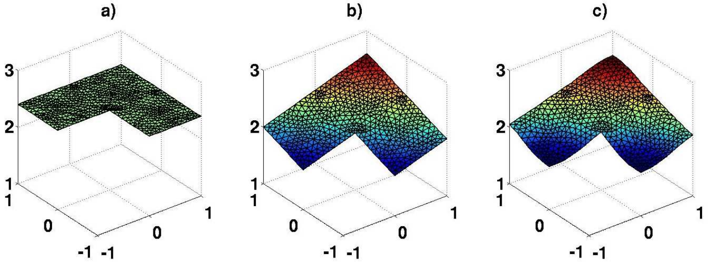


Figure: **“Linear truth”, experiment 1 (L=447) after 1th update**: a) mean of the prior, $\bar{\kappa}_f$ b) truth, κ c) mean of the posterior, $\bar{\kappa}_a$

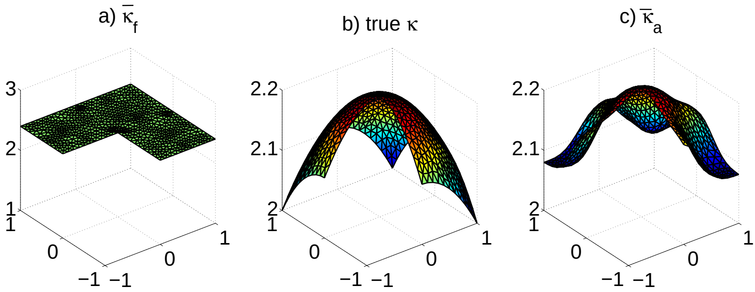


Figure: “Quadratic truth”, experiment 1 (L=447) after 4th update: a) mean of the prior, $\bar{\kappa}_f$ b) truth, κ c) mean of the posterior, $\bar{\kappa}_a$

Conclusion

- The ill-posed problem is regularized by introduction of apriori information
- the update of the prior is a projection of the minimum variance estimator from linear Bayesian updating onto the polynomial chaos basis
- for the mean and variance the estimation is of the Kalman type.
- The estimation is purely deterministic without need for any kind of sampling procedures
- The presented linear Bayesian update does not need any linearity in the forward model, and it can readily update non-Gaussian uncertainties.

Thank you for your attention! Any Questions?

LInear **B**ay**E**sian di**R**ec**T** pol**Y**nomial chaos update

References

1. Gamerman, D. and Lopes, H. F. , Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman and Hall, 2006
2. Kučerová, A. and Matthies, H. G., Uncertainty Updating in the Description of Heterogeneous Materials, Technische Mechanik, Vol. 30, pp. 211–225, 2010
3. Marzouk, Y. M. and Najm, H. N. ,Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, J. Comput. Phys, Vol. 228, 2009
4. Christen, J. A. and Fox, C., MCMC using an approximation, J. Comput. Graph. Stat., Vol. 14, pp. 795–810, 2005
5. Luenberger, D. G., Optimization by Vector Space Methods, John Wiley and Sons, Inc., New York, 1969
6. Rosić, B., Litvinenko. A, Pajonk O., Matthies H.G., Direct Bayesian update of polynomial chaos representations, J. Comput. Phys, 2011, submitted
7. Pajonk, O. and Rosić, B. V. and Litvinenko, A. and Matthies, H. G., A Deterministic Filter for non-Gaussian Bayesian