

Parametric Problems, Stochastics, and Identification

Hermann G. Matthies^a

B. Rosić^{ab}, O. Pajonk^{ac}, A. Litvinenko^a

^aInstitute of Scientific Computing, TU Braunschweig

^bUniversity of Kragujevac

^cSPT Group, Hamburg

wire@tu-bs.de

<http://www.wire.tu-bs.de>

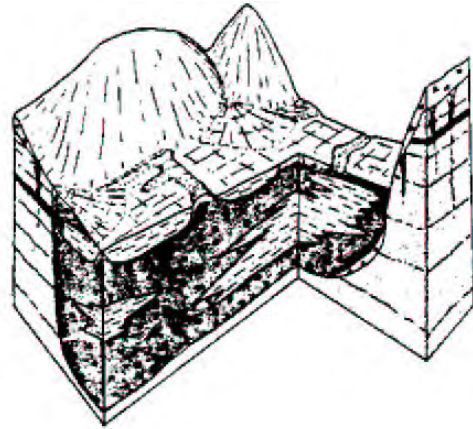


Overview

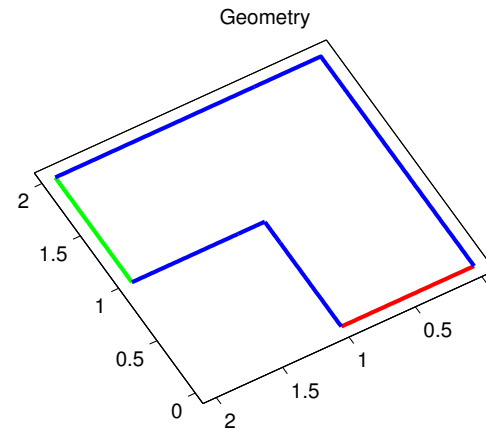
2

1. Parameter identification
2. Parametric forward problem
3. Bayesian updating, inverse problems
4. Tensor approximation
5. Bayesian computation
6. Examples

To fix ideas: example problem



Aquifer



2D Model

Simple stationary model of groundwater flow with stochastic data

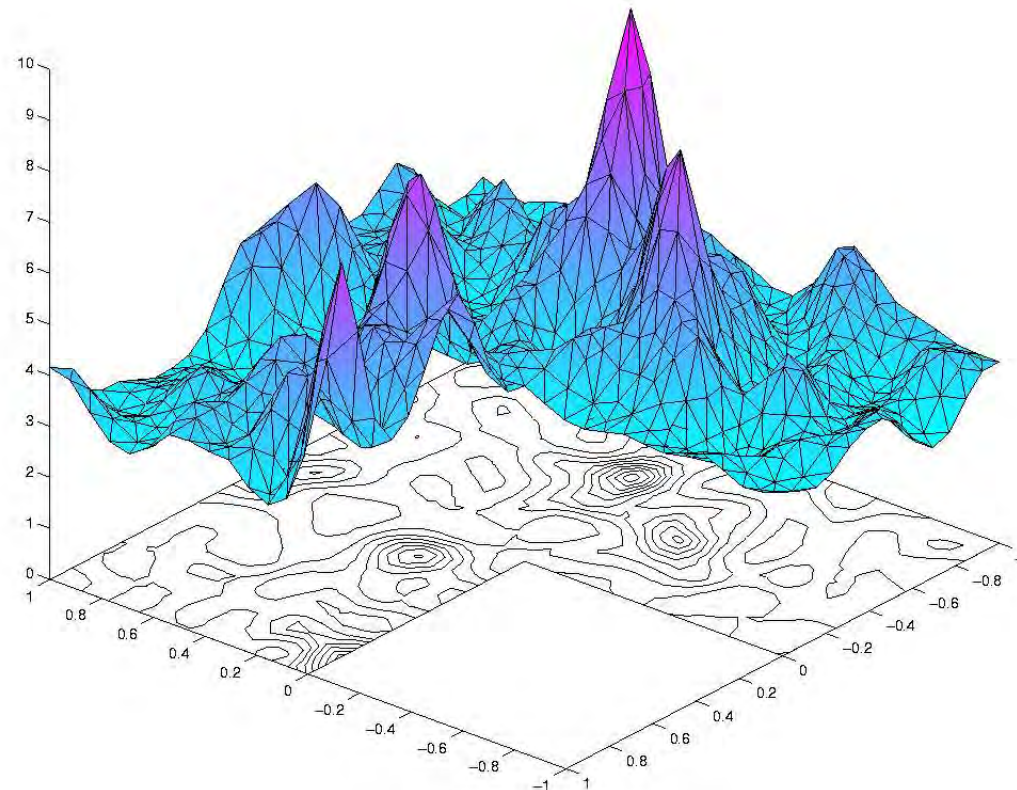
$$-\nabla_x \cdot (\kappa(x, \omega) \nabla_x u(x, \omega)) = f(x, \omega) \quad \& \text{ b.c.}, \quad x \in \mathcal{G} \subset \mathbb{R}^d$$

$$-\kappa(x, \omega) \nabla_x u(x, \omega) = g(x, \omega), \quad x \in \Gamma \subset \partial \mathcal{G}, \quad \omega \in \Omega.$$

Parameter $q(x, \omega) = \log \kappa(x, \omega)$ is **uncertain**,
the **stochastic** conductivity κ , as well as f and g — sinks and sources.

Realisation of $\kappa(x, \omega)$

A sample realization



Mathematical setup

Consider operator equation, physical **system** modelled by A :

$$A(u) = f \quad u \in \mathcal{U}, f \in \mathcal{F},$$

$$\Leftrightarrow \forall v \in \mathcal{U} : \quad \langle A(u), v \rangle = \langle f, v \rangle,$$

\mathcal{U} — space of **states**, $\mathcal{F} = \mathcal{U}^*$ — dual space of **actions** / **forcings**.

Solution operator: $u = U(f)$, inverse of A .

Operator depends on **parameters** $q \in \mathcal{Q}$,
hence state u is also function of q :

$$A(u; q) = f(q) \quad \Rightarrow \quad u = U(f; q).$$

Measurement operator Y with values in \mathcal{Y} :

$$y = Y(q; u) = Y(q, U(f; q)).$$

Forward parametric problem

Parametric elements: operator $A(\cdot; q)$, rhs $f(q)$, state $u(q)$, $\rightarrow r(q)$.

Goal are representations of $r(q) \in \mathcal{W}$, i.e. $r : \mathcal{Q} \rightarrow \mathcal{W}$.

Help from **inner product** $\langle \cdot | \cdot \rangle_{\mathcal{R}}$ on subspace $\mathcal{R} \subset \mathbb{R}^{\mathcal{Q}}$.

In case \mathcal{Q} is a measure / probability space, $\mathcal{R} = L_2$.

To each parametric element corresponds **linear map**

$$R : \mathcal{W} \ni \hat{r} \mapsto \langle \hat{r} | r(\cdot) \rangle_{\mathcal{R}} \in \mathcal{R}.$$

Key is self-adjoint positive map $C = R^*R : \mathcal{W} \rightarrow \mathcal{W}$.

Spectral factorisation of C leads to **Karhunen-Loève** representation, a **tensor product** rep., corresponds to **SVD** of R (a.k.a. POD).

Each **factorisation** $C = B^*B$ leads to a tensor representation,
(ex.: smoothed white noise)

a 1–1 correspondence between factorisations and representations.

Setting for the identification process

General idea:

We observe / measure a system, whose structure we know in principle.

The system behaviour depends on some quantities (parameters),
which we do not know \Rightarrow uncertainty.

We model (uncertainty in) our knowledge in a Bayesian setting:
as a probability distribution on the parameters.

We start with what we know a priori, then perform a measurement.
This gives new information, to update our knowledge (identification).

Update in probabilistic setting works with conditional probabilities
 \Rightarrow Bayes's theorem.

Repeated measurements lead to better identification.

Inverse problem

For given f , measurement y is just a function of q .
This function is usually **not invertible** \Rightarrow **ill-posed** problem,
measurement y does **not** contain **enough information**.

In **Bayesian** framework state of knowledge **modelled** in a probabilistic way,
parameters q are **uncertain**, and **assumed** as **random**.

Bayesian setting allows **updating / sharpening** of **information**
about q when measurement is performed.

The problem of updating **distribution**—state of knowledge of q
becomes **well-posed**.

Can be applied **successively**, each new measurement y and
forcing f —may also be uncertain—will provide **new information**.

Model with uncertainties

For simplicity assume that \mathcal{Q} is a Hilbert space, and $q(\omega)$ has **finite** variance — $\|q\|_{\mathcal{Q}} \in \mathcal{S} := L_2(\Omega)$, so that

$$q \in L_2(\Omega, \mathcal{Q}) \cong \mathcal{Q} \otimes L_2(\Omega) = \mathcal{Q} \otimes \mathcal{S} =: \mathcal{Q}.$$

System model is now

$$A(u(\omega); q(\omega)) = f(\omega) \quad \text{a.s. in } \omega \in \Omega,$$

state $u = u(\omega)$ becomes \mathcal{U} -valued **random variable** (RV), element of a **tensor** space $\mathcal{U} = \mathcal{U} \otimes \mathcal{S}$.

As **variational** statement:

$$\forall v \in \mathcal{U} : \quad \mathbb{E} (\langle A(u(\cdot); q(\cdot)), v \rangle) = \mathbb{E} (\langle f(\cdot), v \rangle) =: \langle\langle f, v \rangle\rangle.$$

Leads to **well-posed** stochastic PDE (SPDE).

Representation of randomness

Parameters q modelled as \mathcal{Q} -valued (a vector space) **RVs** on some probability space $(\Omega, \mathbb{P}, \mathfrak{A})$, with expectation operator $\mathbb{E}(q) = \bar{q}$.

RVs $q : \Omega \rightarrow \mathcal{Q}$ (and $u(q)$) may be **represented** in the following ways:

Samples: the best known representation, i.e. $q(\omega_1), \dots, q(\omega_N), \dots$

Distribution of q . This is the **push-forward** measure $q_*\mathbb{P}$ on \mathcal{Q} .

Moments of q , like $\mathbb{E}(q \otimes \dots \otimes q)$ (mean, covariance, ...).

Functional/Spectral: Functions of other (**known**) RVs, like Wiener's polynomial chaos, i.e. $q(\omega) = q(\theta_1(\omega), \dots, \theta_M(\omega), \dots) =: q(\boldsymbol{\theta})$.

Sampling and **functional** representation work with **vectors**,
allows **linear algebra** in computation.

Computational approaches

Representation determines algorithms:

- **Distributions** → Kolmogorov / Fokker-Planck equations.
Needs **new** software, deterministic solver $u = S(f, q)$ **not** used.
- **Moments** → New (sometimes difficult) equations.
Needs **new** software, deterministic solver **mostly not** used.
- **Sampling** → Domain of **direct integration** methods;
(quasi) Monte Carlo, sparse (Smolyak) grids, etc.
Obviously non-intrusive; software **interface** → **solve**.
- **Functional / Spectral** →
 - (1) **Interpolation / collocation**. Based on samples of solution, **non-intrusive**, **solve** interface.
 - (2) **Galerkin** at first sight **intrusive**, but with **quadrature** is also **non-intrusive**, **precond. residual** interface. Allows **greedy** rank-1

Conditional probability and expectation

With state $u \in \mathcal{U} = \mathcal{U} \otimes \mathcal{S}$ a RV, the quantity to be measured

$$y(\omega) = Y(q(\omega), u(\omega)) \in \mathcal{Y} := \mathcal{Y} \otimes \mathcal{S}$$

is also **uncertain**, a **random variable**.

A **new** measurement z is performed, composed from the “true” value $y \in \mathcal{Y}$ and a **random** error ϵ : $z(\omega) = y + \epsilon(\omega) \in \mathcal{Y}$.

Classically, **Bayes’s theorem** gives **conditional probability**

$$\mathbb{P}(I_q | M_z) = \frac{\mathbb{P}(M_z | I_q) \mathbb{P}(I_q)}{\mathbb{P}(M_z)}$$

expectation with this posterior measure is **conditional expectation**.

Kolmogorov starts from **conditional expectation** $\mathbb{E}(\cdot | M_z)$,
from this **conditional probability** via $\mathbb{P}(I_q | M_z) = \mathbb{E}(\chi_{I_q} | M_z)$.

Update

The conditional expectation is defined as orthogonal projection onto the closed subspace $L_2(\Omega, \mathbb{P}, \sigma(z))$:

$$\mathbb{E}(q|\sigma(z)) := P_{\mathcal{Q}_\infty} q = \operatorname{argmin}_{\tilde{q} \in L_2(\Omega, \mathbb{P}, \sigma(z))} \|q - \tilde{q}\|_{L_2}^2$$

The subspace $\mathcal{Q}_\infty := L_2(\Omega, \mathbb{P}, \sigma(z))$ represents the available information, estimate minimises $\Phi(\cdot) := \|q - (\cdot)\|^2$ over \mathcal{Q}_∞ . More general loss functions than mean square error are possible.

The update, also called the assimilated value $q_a(\omega) := P_{\mathcal{Q}_\infty} q = \mathbb{E}(q|\sigma(z))$, is a \mathcal{Q} -valued RV and represents new state of knowledge after the measurement.

Reduction of variance—Pythagoras: $\|q\|_{L_2}^2 = \|q - q_a\|_{L_2}^2 + \|q_a\|_{L_2}^2$

Doob-Dynkin: $\mathcal{Q}_\infty = \{ \varphi \in \mathcal{Q} : \varphi = \phi \circ Y, \phi \text{ measurable} \}$

Important points I

The probability measure \mathbb{P} is not the **object of desire**.
It is the **distribution** of q , a measure on \mathcal{Q} —**push forward** of \mathbb{P} :

$$q_*\mathbb{P}(\mathcal{E}) := \mathbb{P}(q^{-1}(\mathcal{E})) \quad \text{for measurable } \mathcal{E} \subseteq \mathcal{Q}.$$

Bayes's original formula **changes** \mathbb{P} , **leaves** q as is.
Kolmogorov's conditional expectation **changes** q , **leaves** \mathbb{P} as is.
In both cases the update is a new $q_*\mathbb{P}$.

\mathbb{P} (a probability measure) is on positive part of **unit sphere**,
whereas q is **free** in a **vector space**.

This will allow the use of (multi-)linear algebra
and **tensor approximations**.

Important points II

Identification process:

- Use **forward problem** $A(u(\omega); q(\omega)) = f(\omega)$ to **forecast** new state $u_f(\omega)$ and measurement $y_f(\omega) = Y(q(\omega), u_f(\omega))$.
- Perform minimisation of **loss function** to obtain **update map / filter**.
- Use innovation in **inverse problem** to update forecast q_f to obtain **assimilated** (updated) q_a with update map.
- All operations in vector space, use **tensor approximations** throughout.

Approximation

Minimisation equivalent to **orthogonality**: find $\phi \in L_0(\mathcal{Y}, \mathcal{Q})$

$$\forall p \in \mathcal{Q}_\infty : \quad \langle\langle D_{q_a} \Phi(q_a(\phi)), p \rangle\rangle_{L_2} = \langle\langle q - q_a, p \rangle\rangle_{L_2} = 0,$$

Approximation of \mathcal{Q}_∞ : take $\mathcal{Q}_n \subset \mathcal{Q}_\infty$

$$\mathcal{Q}_n := \{ \varphi \in \mathcal{Q} : \varphi = \psi_n \circ Y, \psi_n \text{ a } n^{\text{th}} \text{ degree polynomial} \}$$

$$\text{i.e. } \varphi = {}^0H + {}^1HY + \dots + {}^kHY^{\otimes k} + \dots + {}^nHY^{\otimes n},$$

where ${}^kH \in \mathcal{L}_s^k(\mathcal{Y}, \mathcal{Q})$ is **symmetric** and **k -linear**.

With $q_a(\phi) = q_a(({}^0H, \dots, {}^kH, \dots, {}^nH)) = \sum_{k=0}^n {}^kH z^{\otimes k} = P_{\mathcal{Q}_n} q$,
orthogonality implies

$$\forall \ell = 0, \dots, n : \quad D_{({}^\ell H)} \Phi(q_a({}^0H, \dots, {}^kH, \dots, {}^nH)) = 0$$

Determining the n -th degree Bayesian update

With the abbreviations

$$\langle p \otimes v^{\otimes k} \rangle := \mathbb{E} (p \otimes v^{\otimes k}) = \int_{\Omega} p(\omega) \otimes v(\omega)^{\otimes k} \mathbb{P}(d\omega),$$

$$\text{and } {}^k H \langle z^{\otimes(\ell+k)} \rangle := \langle z^{\otimes\ell} \otimes ({}^k H z^{\otimes k}) \rangle = \mathbb{E} (z^{\otimes\ell} \otimes ({}^k H z^{\otimes k})),$$

we have for the **unknowns** $({}^0 H, \dots, {}^k H, \dots, {}^n H)$

$$\ell = 0 : {}^0 H \quad \dots + {}^k H \langle z^{\otimes k} \rangle \quad \dots + {}^n H \langle z^{\otimes n} \rangle = \langle q \rangle,$$

$$\ell = 1 : {}^0 H \langle z \rangle \quad \dots + {}^k H \langle z^{\otimes(1+k)} \rangle \dots + {}^n H \langle z^{\otimes(1+n)} \rangle = \langle q \otimes z \rangle,$$

$$\vdots \quad \dots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$\ell = n : {}^0 H \langle z^{\otimes n} \rangle \dots + {}^k H \langle z^{\otimes(n+k)} \rangle \dots + {}^n H \langle z^{\otimes 2n} \rangle = \langle q \otimes z^{\otimes n} \rangle$$

a linear system with **symmetric positive definite**

Hankel operator matrix $(\langle z^{\otimes(\ell+k)} \rangle)_{\ell,k}$.

Bayesian update in components

For **short** $\forall \ell = 0, \dots, n$:

$$\sum_{k=0}^n {}^k H \langle z^{\otimes(\ell+k)} \rangle = \langle q \otimes z^{\otimes \ell} \rangle,$$

For finite dimensional spaces, or after discretisation,
in **components** (or à la Penrose in ‘**symbolic index**’ notation):

let $q = (q^m)$, $z = (z^j)$, and ${}^k H = ({}^k H_{j_1 \dots j_k}^m)$, then:

$\forall \ell = 0, \dots, n$;

$$\langle z^{j_1} \dots z^{j_\ell} \rangle ({}^0 H^m) + \dots + \langle z^{j_1} \dots z^{j_{\ell+1}} \dots z^{j_{\ell+k}} \rangle ({}^k H_{j_{\ell+1} \dots j_{\ell+k}}^m) + \dots + \langle z^{j_1} \dots z^{j_{\ell+1}} \dots z^{j_{\ell+n}} \rangle ({}^n H_{j_{\ell+1} \dots j_{\ell+n}}^m) = \langle q^m z^{j_1} \dots z^{j_\ell} \rangle.$$

(Einstein summation convention used)

matrix does **not** depend on m —it is identically **block diagonal**.

Special cases

For $n = 0$ (**constant** functions) $\Rightarrow q_a = {}^0H = \langle q \rangle$ ($= \mathbb{E}(q)$).

For $n = 1$ the approximation is with **affine** functions

$${}^0H + {}^1H \langle z \rangle = \langle q \rangle$$

$${}^0H \langle z \rangle + {}^1H \langle z \otimes z \rangle = \langle q \otimes z \rangle$$

\Rightarrow (remember that $[\text{cov}_{qz}] = \langle q \otimes z \rangle - \langle q \rangle \otimes \langle z \rangle$)

$${}^0H = \langle q \rangle - {}^1H \langle z \rangle$$

$${}^1H (\langle z \otimes z \rangle - \langle z \rangle \otimes \langle z \rangle) = \langle q \otimes z \rangle - \langle q \rangle \otimes \langle z \rangle$$

$$\Rightarrow {}^1H = [\text{cov}_{qz}] [\text{cov}_{zz}]^{-1} \text{ (Kalman's solution),}$$

$${}^0H = \langle q \rangle - [\text{cov}_{qz}] [\text{cov}_{zz}]^{-1} \langle z \rangle,$$

and **finally**

$$q_a = {}^0H + {}^1H z = \langle q \rangle + [\text{cov}_{qz}] [\text{cov}_{zz}]^{-1} (z - \langle z \rangle).$$

Case with prior information

Here we have **prior information** \mathcal{Q}_f and **prior estimate** $q_f(\omega)$ (forecast) and measurements z **generating** via Y a subspace $\mathcal{Q}_y \subset \mathcal{Q}$.

We now need **projection** onto $\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_y$, with reformulation as an **orthogonal direct** sum:

$$\mathcal{Q}_a = \mathcal{Q}_f + \mathcal{Q}_y = \mathcal{Q}_f \oplus (\mathcal{Q}_y \cap \mathcal{Q}_f^\perp) = \mathcal{Q}_f \oplus \mathcal{Q}_\infty.$$

The **update** / **conditional expectation** / **assimilated** value is the orthogonal projection

$$q_a = q_f + P_{\mathcal{Q}_\infty} q = q_f + q_\infty,$$

where q_∞ is the **innovation**.

Compute q_a by approximating: $\mathcal{Q}_n \subset \mathcal{Q}_\infty$. We now take $n = 1$.

Simplification

The case $n = 1$ —linear functions, projecting onto \mathcal{Q}_1 —is well known:

this is the **linear minimum variance** estimate \hat{q}_a .

Theorem: (Generalisation of **Gauss-Markov**)

$$\hat{q}_a(\omega) = q_f(\omega) + {}^1H(z(\omega) - y_f(\omega)),$$

where the linear **Kalman** gain operator ${}^1H : \mathcal{Y} \rightarrow \mathcal{Q}$ is

$${}^1H := [\text{cov}_{qz}][\text{cov}_{zz}]^{-1} = [\text{cov}_{qy}][\text{cov}_{yy} + \text{cov}_{\epsilon\epsilon}]^{-1}.$$

(The **normal Kalman** filter is a **special case**.)

Or in tensor space $q \in \mathcal{Q} = \mathcal{Q} \otimes \mathcal{S}$:

$$\hat{q}_a = q_f + ({}^1H \otimes I)(z - y_f).$$

Deterministic model, discretisation, solution

Remember operator equation: $A(u) = f \quad u \in \mathcal{U}, f \in \mathcal{F}$.

Solution is usually by first **discretisation**

$$A(u) = f \quad u \in \mathcal{U}_N \subset \mathcal{U}, f \in \mathcal{F}_N = \mathcal{U}_N^* \subset \mathcal{F},$$

and then **(iterative)** numerical **solution** process

$$u_{k+1} = S(u_k), \quad \lim_{k \rightarrow \infty} u_k = u.$$

S evaluates (pre-conditioned) **residua** $f - A(u_k)$.

Similarly for model with **uncertainty**:

$$A(u(\omega); q(\omega)) = f(\omega),$$

assume $\{v_j\}_{j=1}^N$ a basis in \mathcal{U}_N , then the approx. solution in $\mathcal{U}_N \otimes \mathcal{S}$

$$u(\omega) = \sum_{j=1}^N u_j(\omega) v_j.$$

Discretisation by functional approximation

Choose subspace $\mathcal{S}_B \subset \mathcal{S}$ with basis $\{X_\beta\}_{\beta=1}^B$,
make **ansatz** for each $u_j(\omega) \approx \sum_{\beta} u_j^\beta X_\beta(\omega)$, giving

$$\mathbf{u}(\omega) = \sum_{j,\beta} u_j^\beta X_\beta(\omega) \mathbf{v}_j = \sum_{j,\beta} u_j^\beta X_\beta(\omega) \otimes \mathbf{v}_j.$$

Solution is in **tensor product** $\mathcal{U}_{N,B} := \mathcal{U}_N \otimes \mathcal{S}_B \subset \mathcal{U} \otimes \mathcal{S} = \mathcal{U}$.

State $\mathbf{u}(\omega)$ represented by **tensor** $\mathbf{u} := \mathbf{u}_N^B := \{u_j^\beta\}_{j=1,\dots,N}^{\beta=1,\dots,B}$,
(β is usually multi-index)

similarly for all other quantities, **fully discrete** forward model
is obtained by **weighting** residual with Ξ_α with ansatz inserted:

$$\forall \alpha : \left\langle \Xi_\alpha(\omega), \mathbf{f}(\omega) - \mathbf{A} \left(\sum_{j,\beta} u_j^\beta X_\beta(\omega) \mathbf{v}_j; \mathbf{q}(\omega) \right) \right\rangle_{\mathcal{S}} = 0.$$

Stochastic forward problem

⇒ generally **coupled** system of equations for $\mathbf{u} = \{u_j^\beta\}$:

$$\mathbf{A}(\mathbf{u}; \mathbf{q}) = \mathbf{f}, \quad \mathbf{y} = \mathbf{Y}(\mathbf{q}; \mathbf{u}).$$

- If $\Xi_\alpha(\cdot) = \delta(\cdot - \omega_\alpha)$, system **decouples** → **collocation / interpolation**; may use for each ω_α **original** solver \mathcal{S} (obviously **non-intrusive**).
- If $\Xi_\alpha(\cdot) = X_\alpha(\cdot)$ → **Bubnov-Galerkin** conditions; with **numerical integration** uses also **original** solver \mathcal{S} and is also **non-intrusive**.
- In **greedy** rank-one **update** tensor solver one uses Bubnov-Galerkin conditions (proper gener. decomp. (**PGD**)/ succ. rank-1 upd. (**SR1U**)/ alt. least squ. (**ALS**)), also possible by **non-intrusive** use of **original** \mathcal{S} .

For update: ${}^1\mathbf{H} = {}^1\mathbf{H} \otimes \mathbf{I}$ computed **analytically** ($X_\beta =$ Hermite basis)
 $[\text{cov}_{qy}] = \sum_{\alpha>0} \alpha! \mathbf{q}^\alpha (\mathbf{y}^\alpha)^T; \quad [\text{cov}_{yy}] = \sum_{\alpha>0} \alpha! \mathbf{y}^\alpha (\mathbf{y}^\alpha)^T.$

Important points III

Update formulation in **vector spaces**.

This makes tensor representation possible .

Parametric problems lead to **tensor** (or separated) representations.

Sparse approximation by **low-rank** representation.

Possible for **forward** problem (progressive or iterative).

Possible for **inverse** problem.

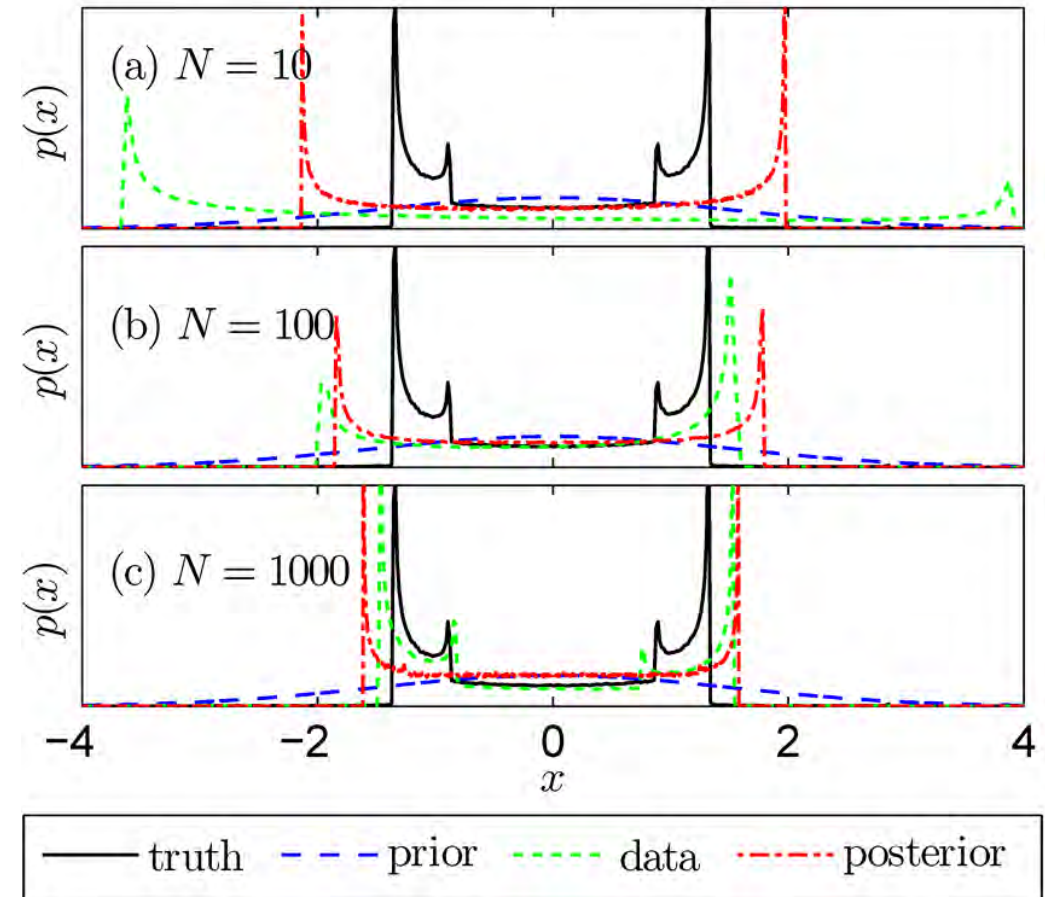
Low-rank approximation can be kept **throughout** update.

Example 1: Identification of multi-modal dist

Setup: Scalar RV x with **non-Gaussian** multi-modal “truth” $p(x)$; Gaussian prior; Gaussian measurement errors.

Aim: Identification of $p(x)$.

10 updates of $N = 10, 100, 1000$ measurements.



Example 2: Lorenz-84 chaotic model

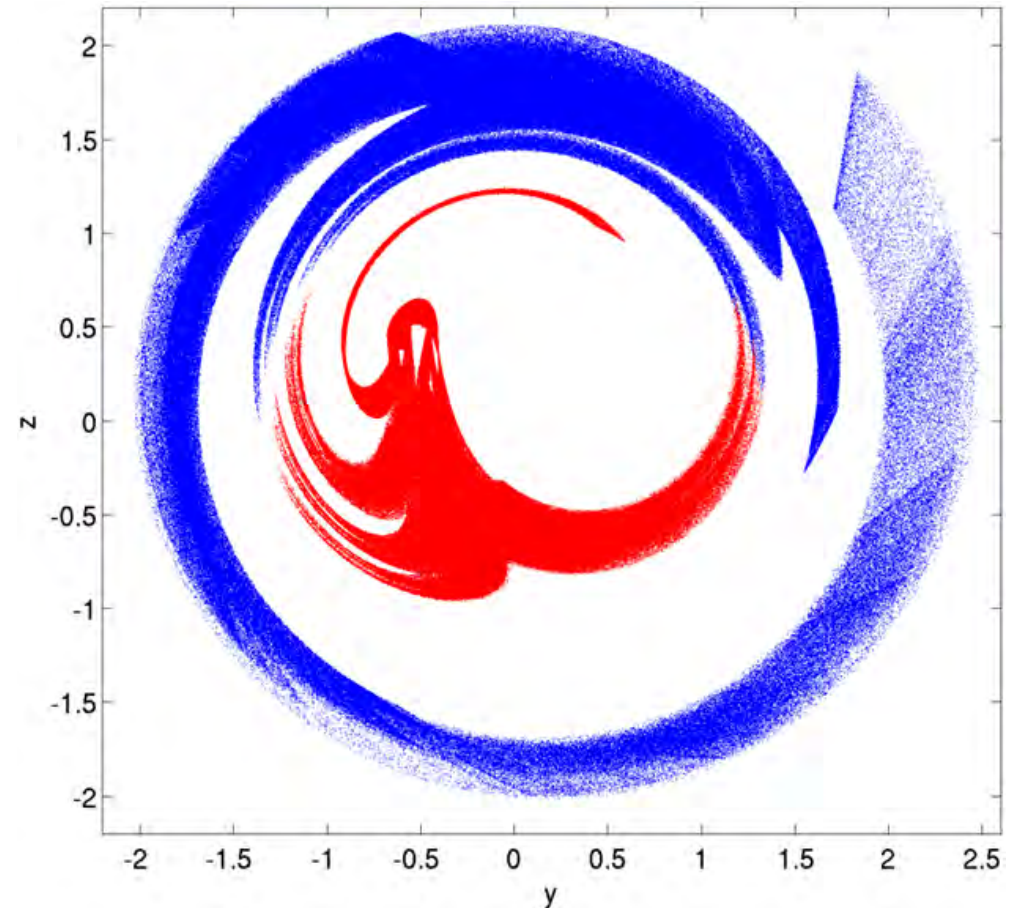
Setup: Non-linear, **chaotic** system

$$\dot{u} = f(u), \quad u = [x, y, z]$$

Small uncertainties in initial conditions u_0 have large impact.

Aim: Sequentially identify state u_t .

Methods: PCE representation and
 PCE updating and
 sampling representation and
 (Ensemble Kalman Filter)
 EnKF updating.

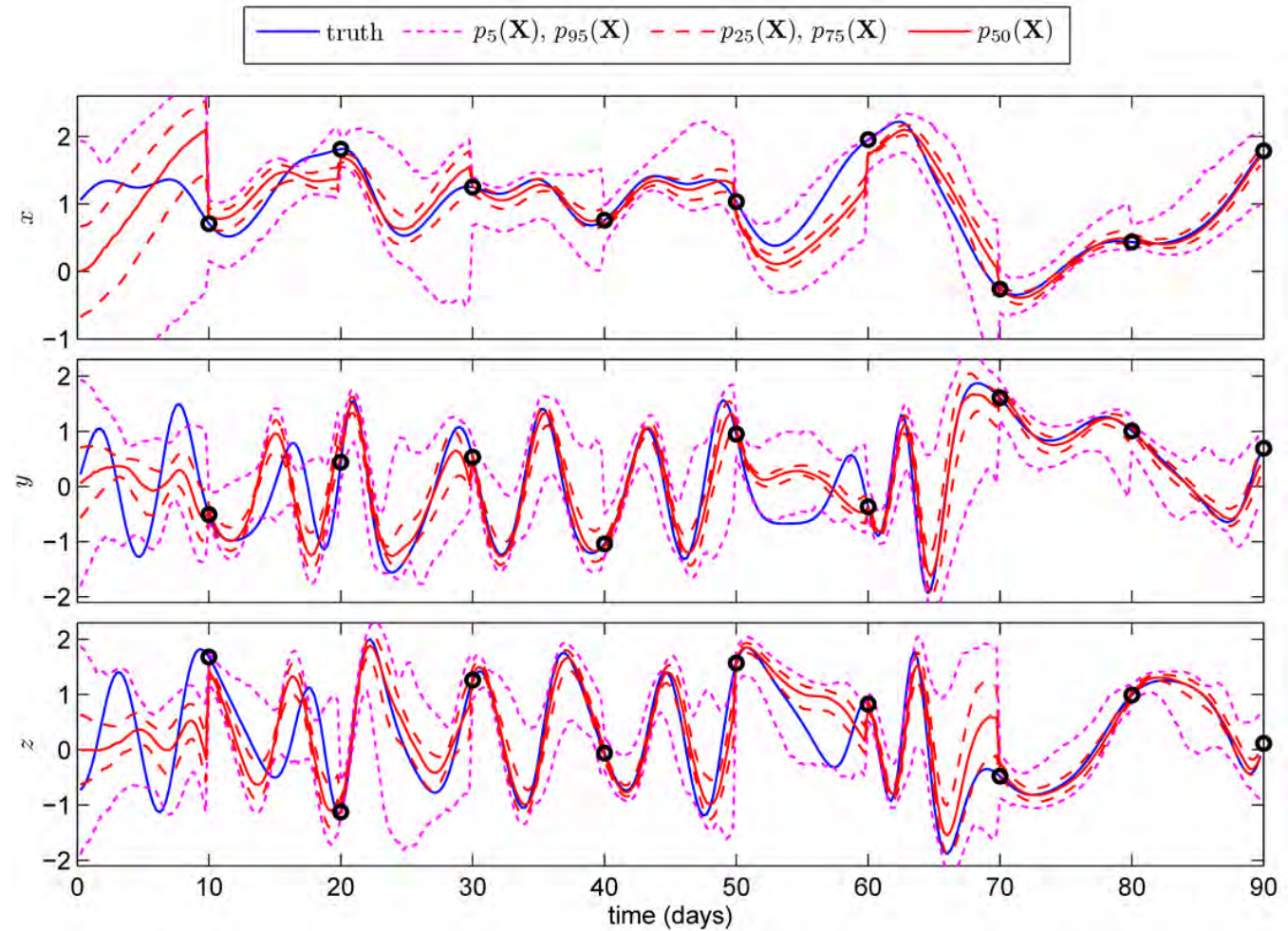


Poincaré cut for $x = 1$.

Example 2: Lorenz-84 PCE representation

PCE: Variance reduction and shift of mean at update points.

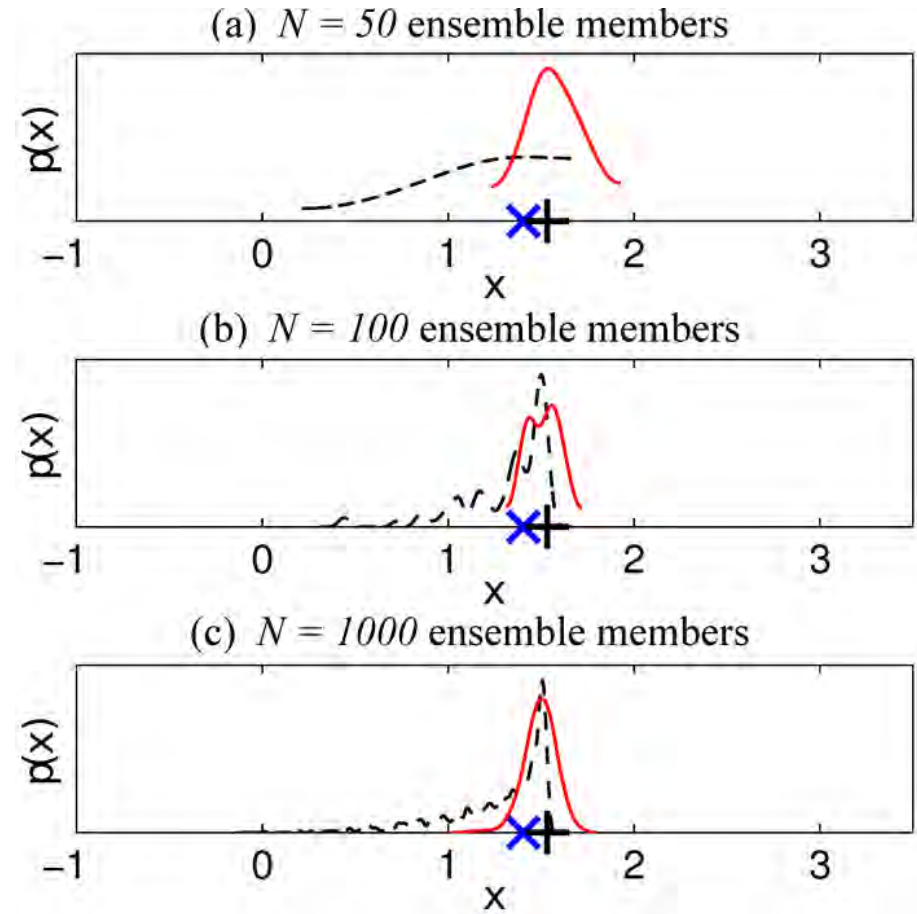
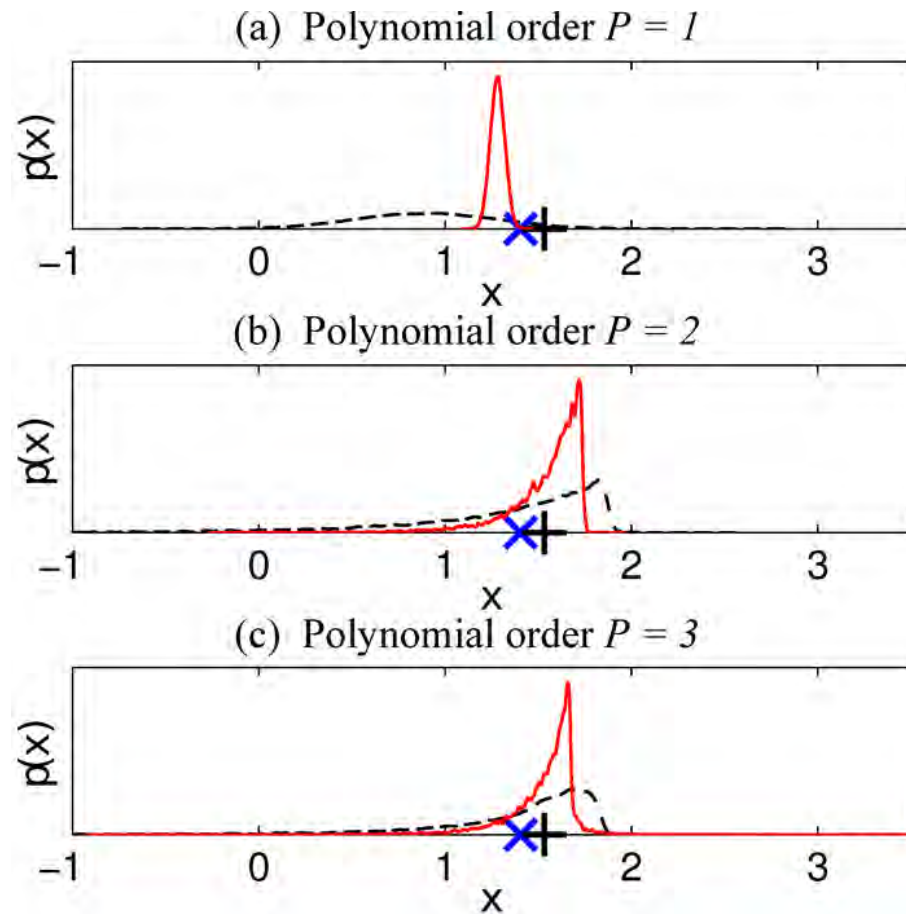
Skewed structure clearly visible, preserved by updates.



Example 2: Lorenz-84 non-Gaussian identification

PCE

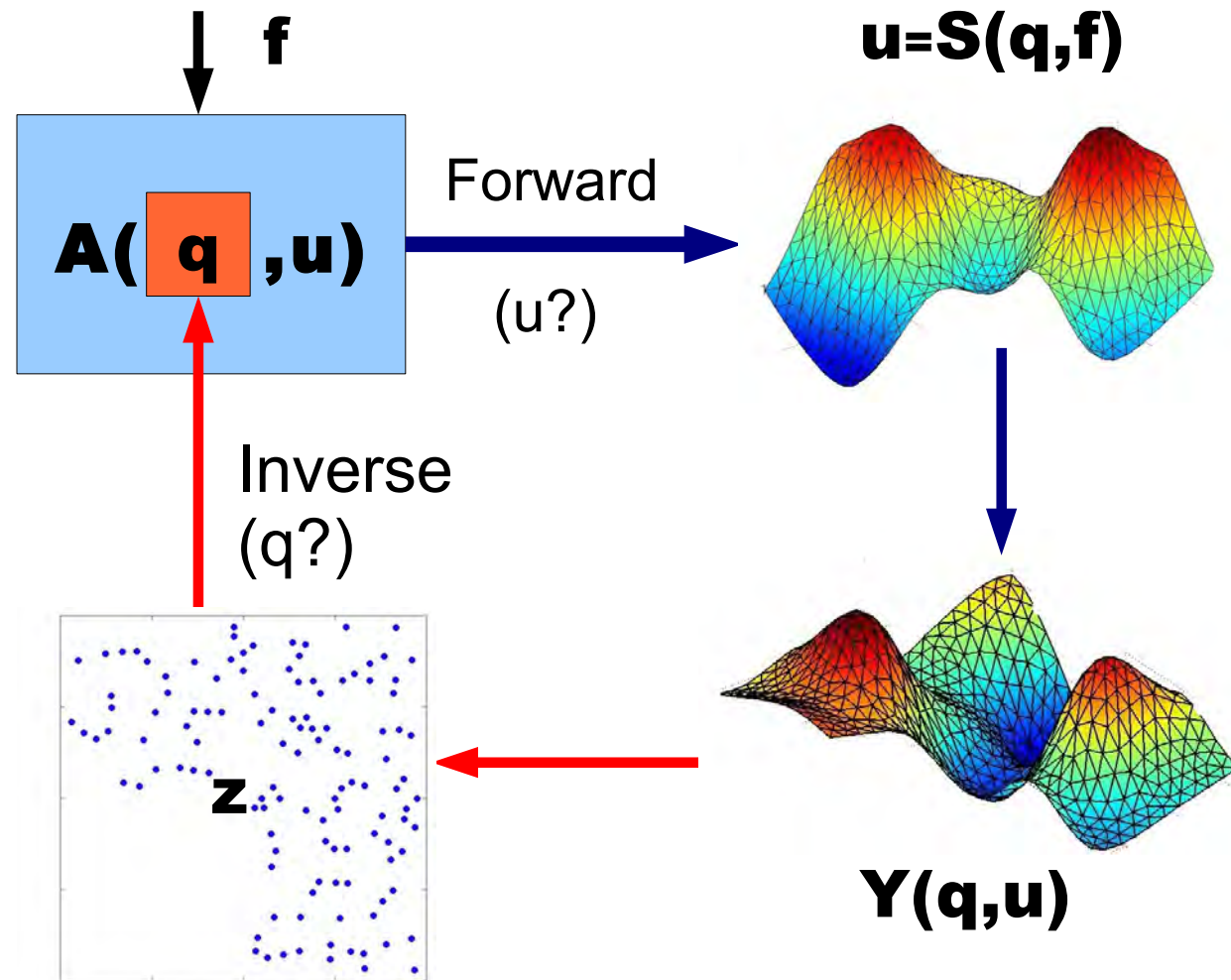
EnKF



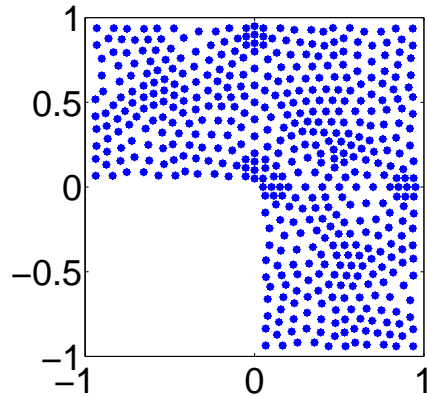
truth \times measurement $+$

posterior prior

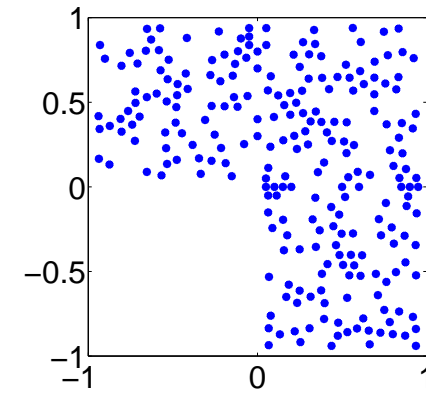
Example 3: diffusion—schematic representation



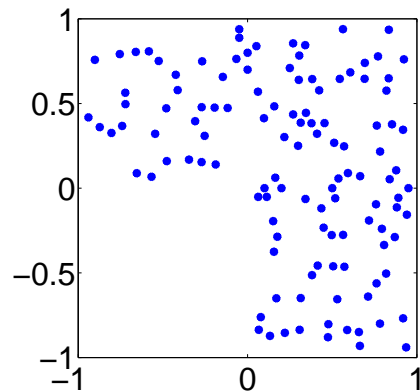
Measurement patches



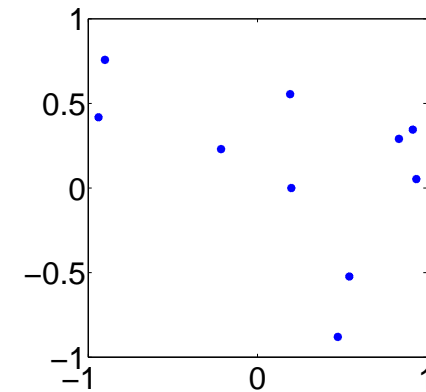
447 measurement patches



239 measurement patches

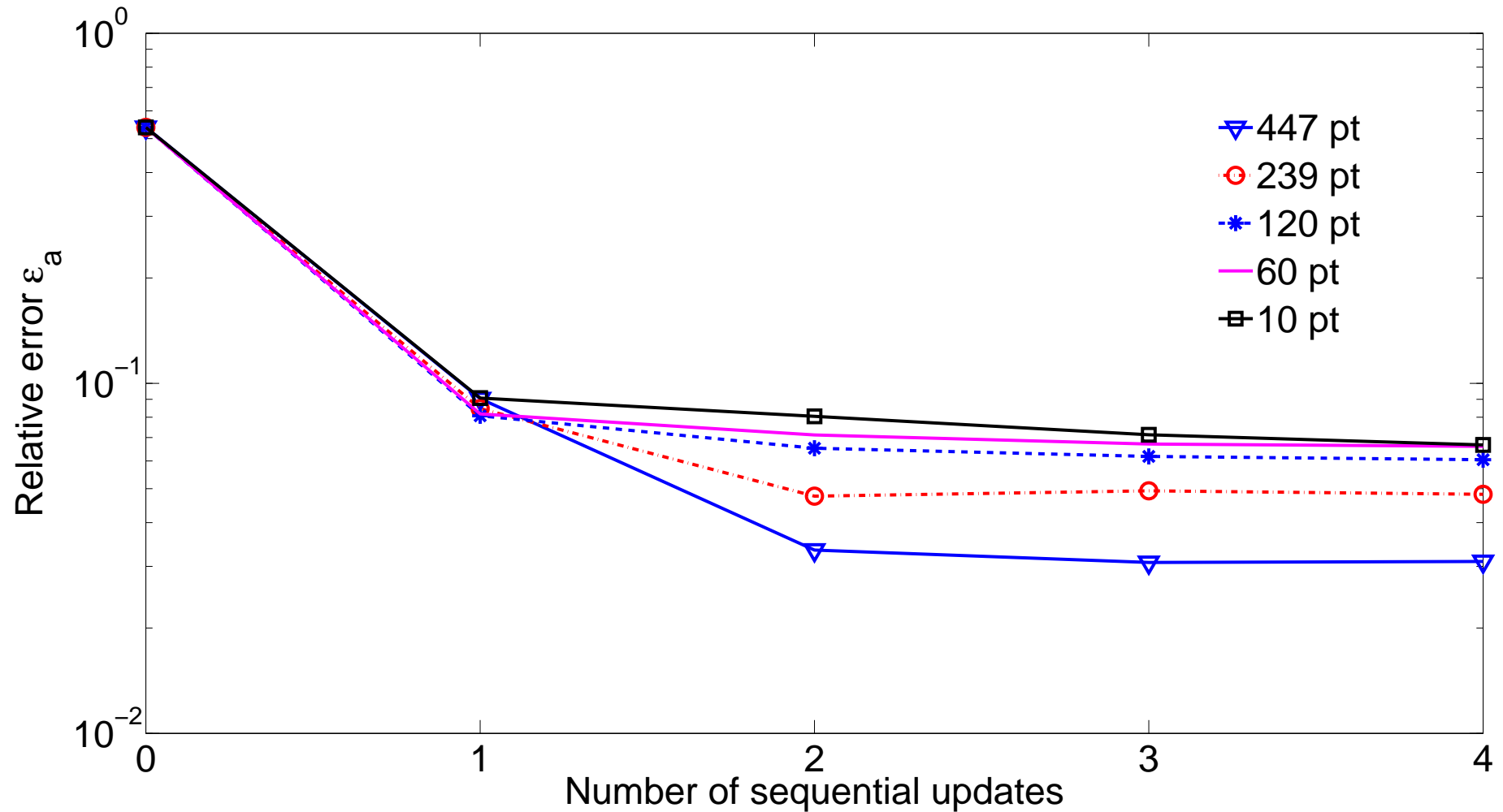


120 measurement patches

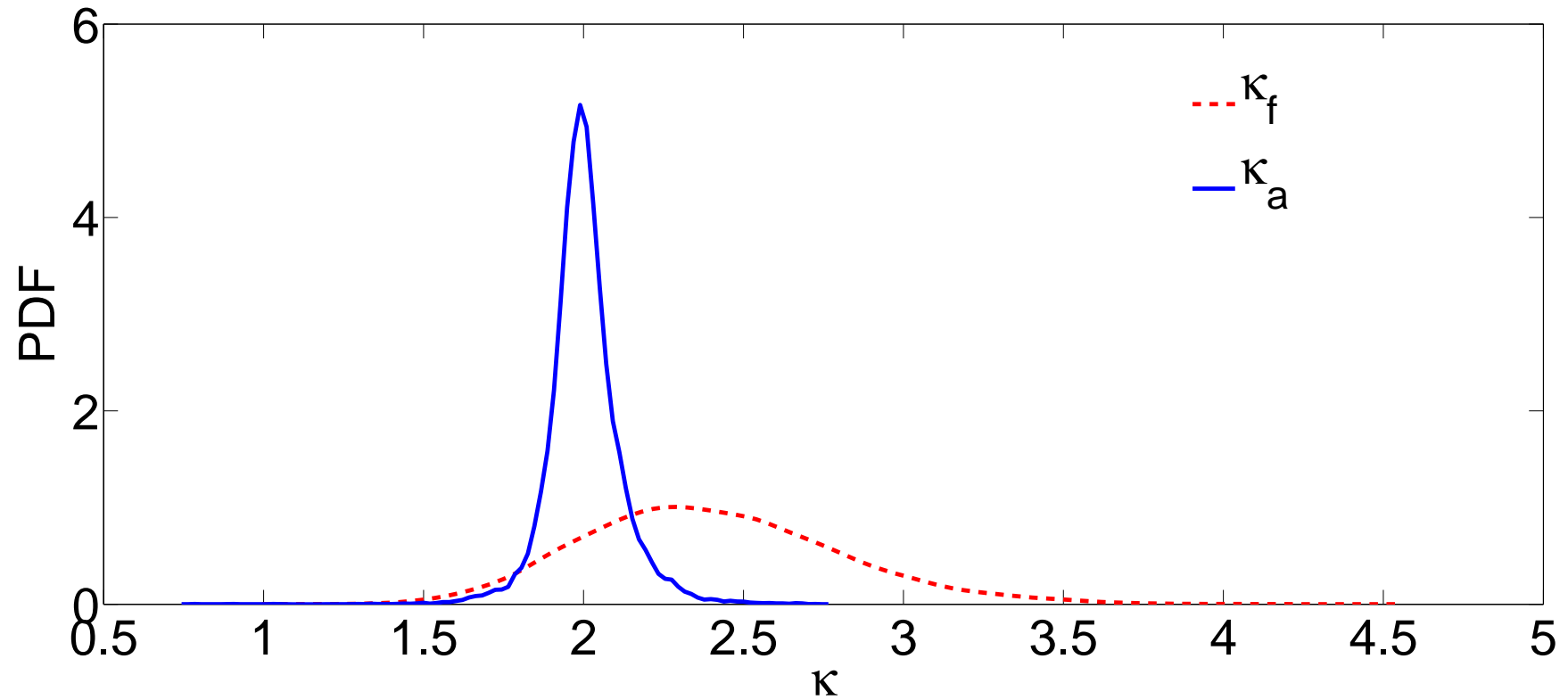


10 measurement patches

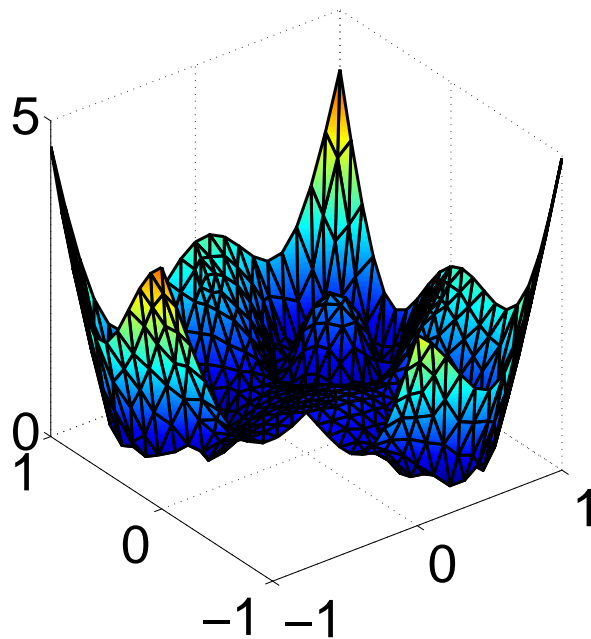
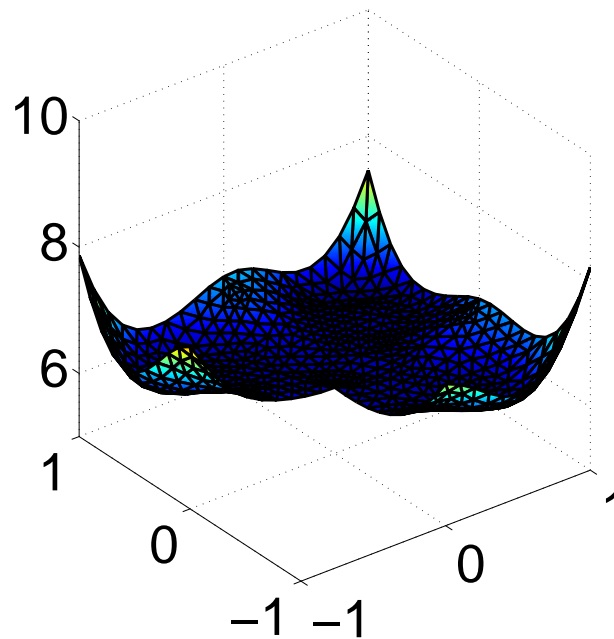
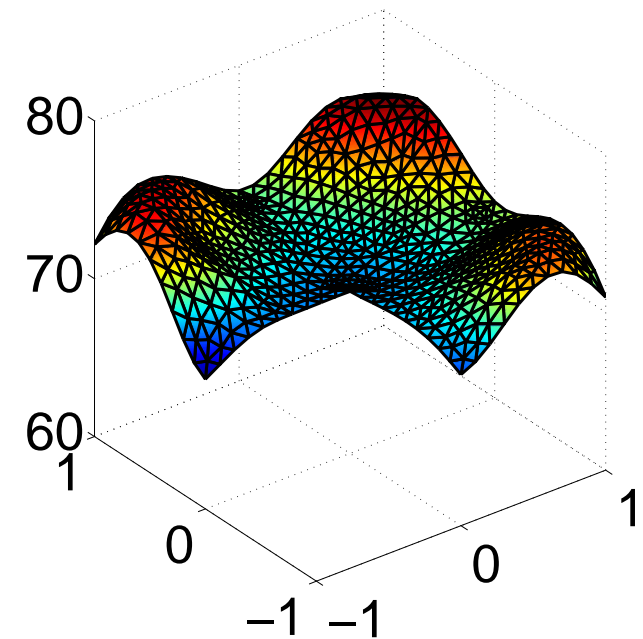
Convergence plot of updates



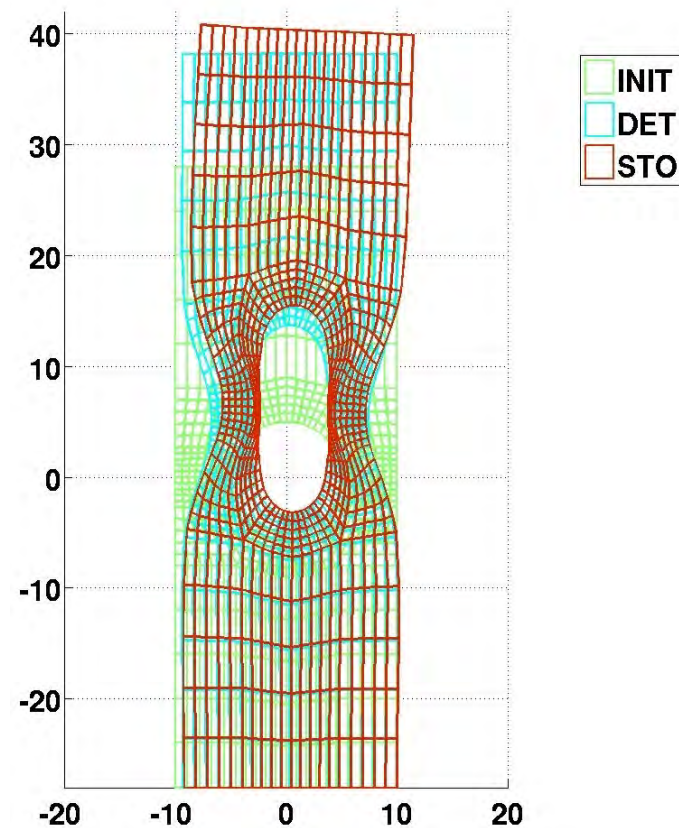
Forecast and Assimilated pdfs



Spatial Error Distribution

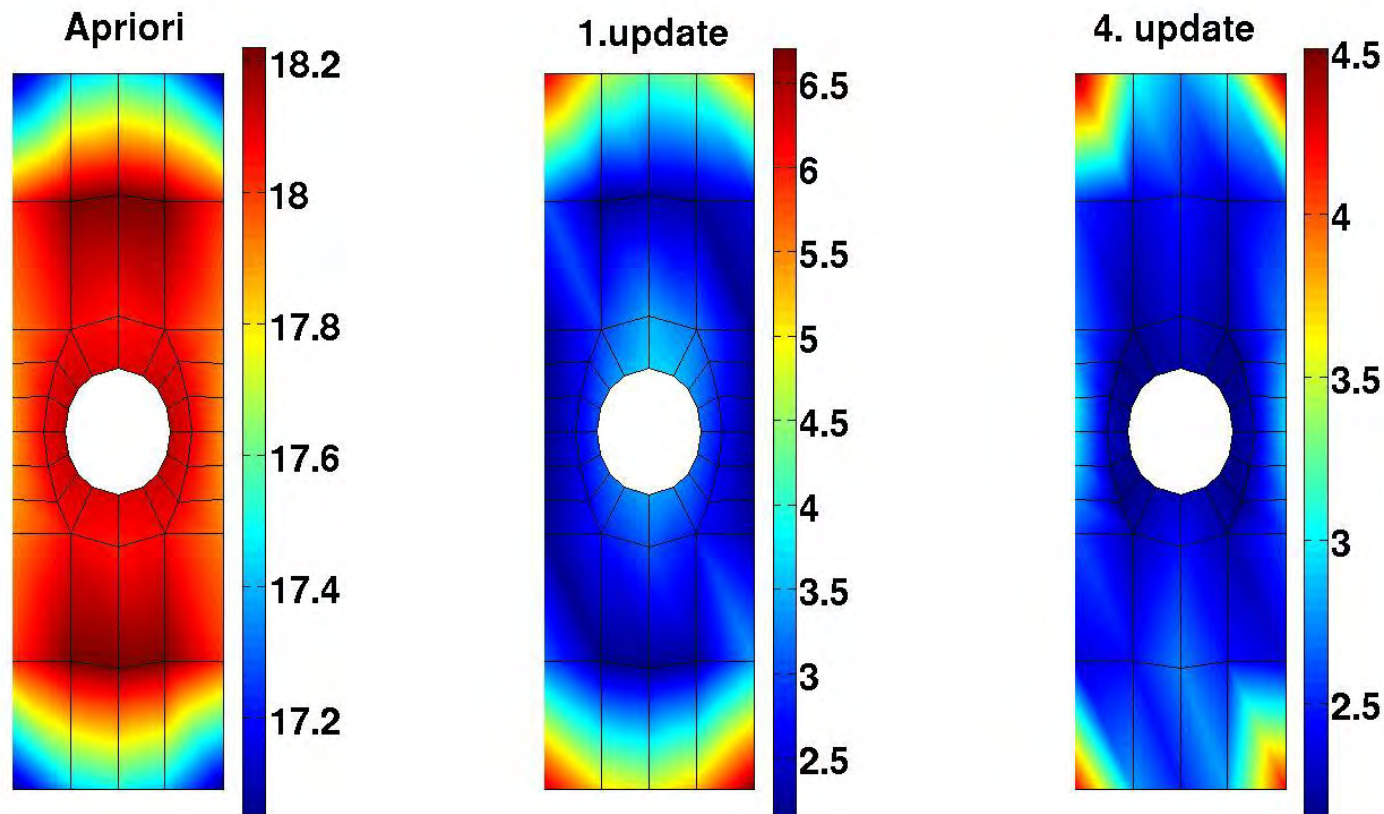
a) $\bar{\varepsilon}_a$ [%]b) ε_a [%]c) I [%]

Example 4: plate with hole



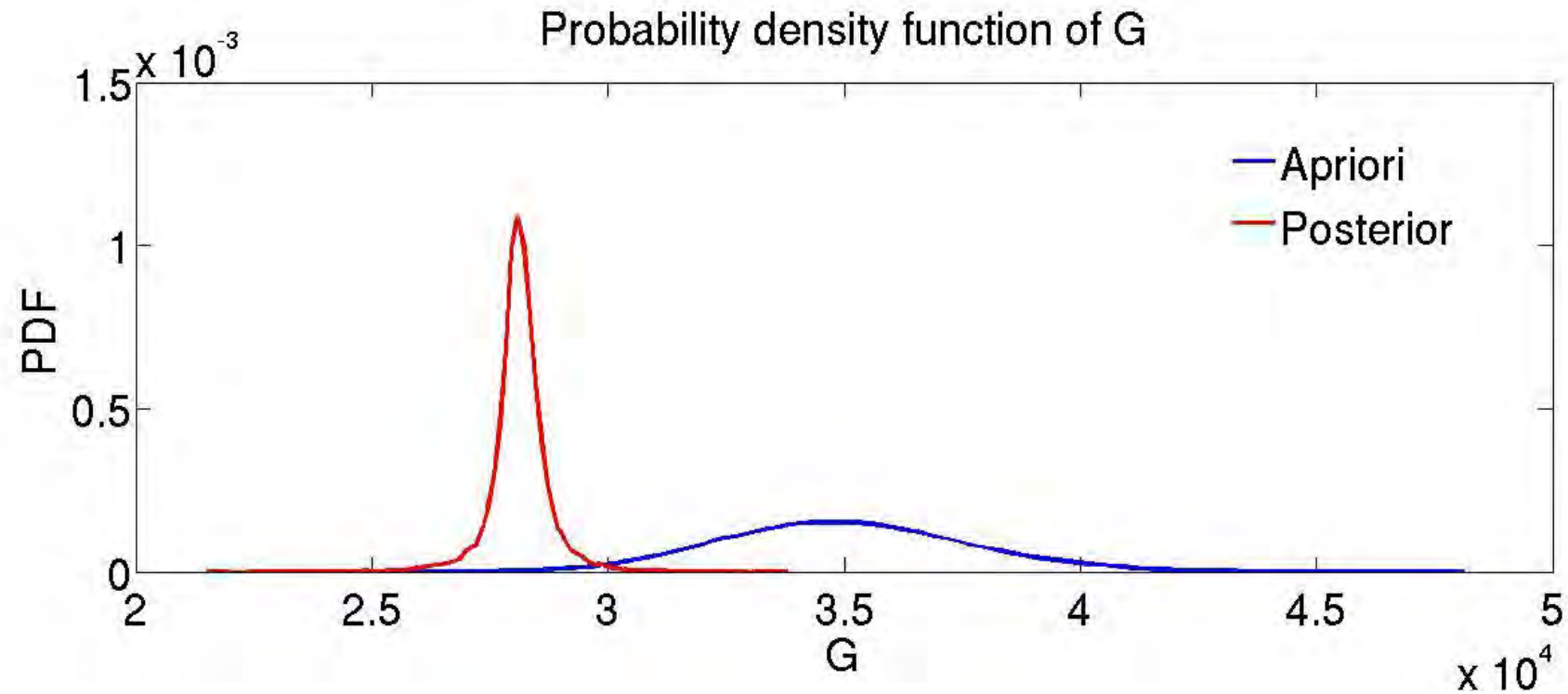
Forward problem: the comparison of the mean values of the total displacement for deterministic, initial and stochastic configuration

Relative variance of shear modulus estimate



Relative RMSE of variance [%] after 4th update in 10% equally distributed measurement points

Probability density shear modulus



Comparison of prior and posterior distribution

Conclusion

- Parametric problems lead to tensor representation.
- **Inverse** problems via Bayes's theorem.
- **Bayesian update** is a **projection**.
- For efficiency try and use **sparse** representation throughout; ansatz in **low-rank** tensor products, **saves** storage as well as computation.
- Bayesian update **compatible** with **low-rank** representation.