Alessandro VENEZIANI
Department of Math&CS
Emory University

Data Driven Computations in the Life Sciences
Autumn School and Workshop,
November 9-13, 2015, IST, Lisbon

A Short Introduction To Estimation Theory and Data Assimilation

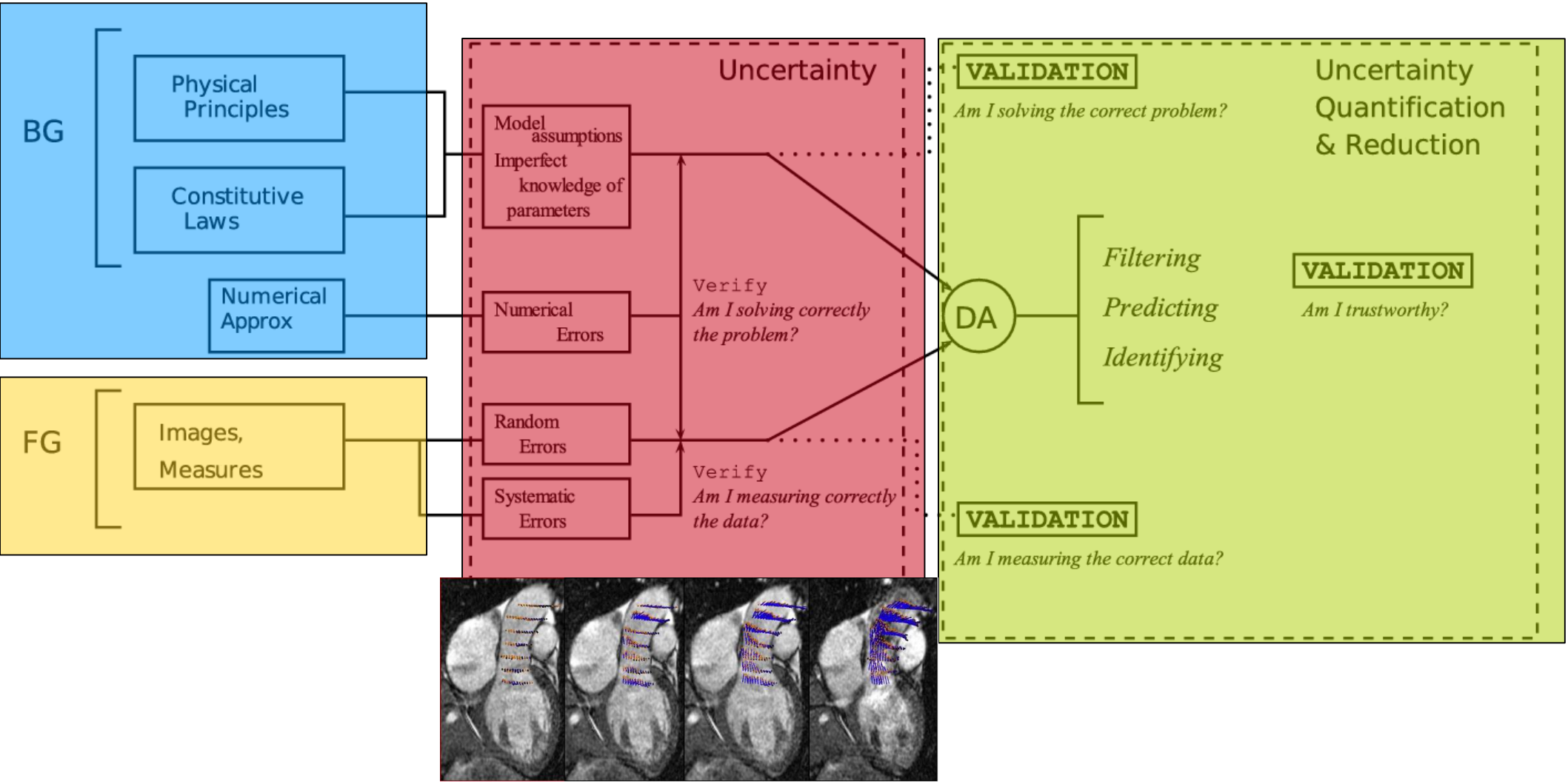
MERGE EXPERIMENTAL DATA AND NUMERICAL SIMULATIONS

ESTIMATION
PREDICTION
IDENTIFICATION

$$\begin{cases} \dfrac{\partial \mathbf{u}}{\partial t} - \mu \triangle \mathbf{u} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} + \nabla p = \mathbf{f} \\ \nabla \cdot \mathbf{u} = 0 \end{cases} \quad \mathbf{x} \in \Omega, t > 0$$

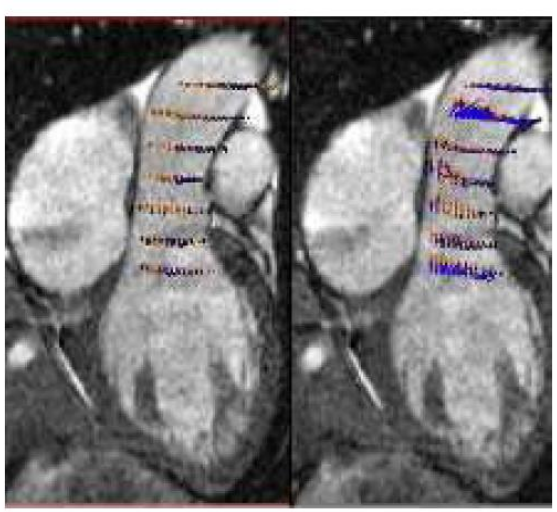
REFERENCE

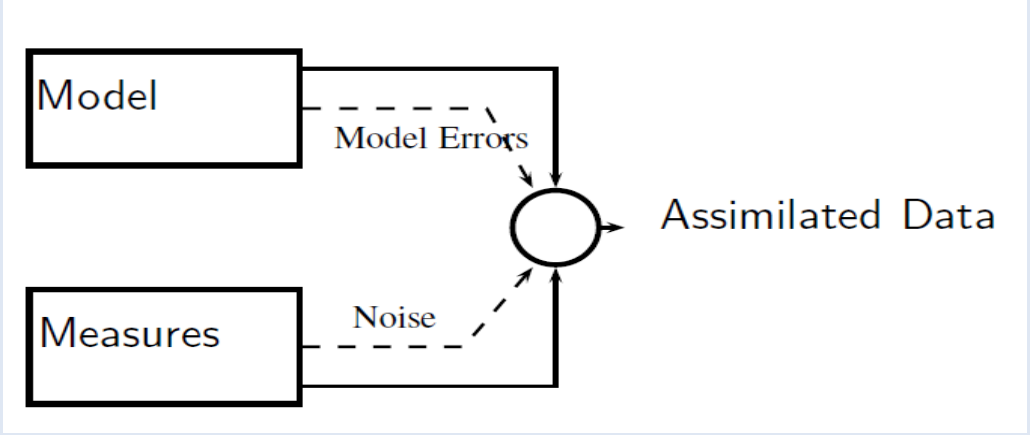
Data Assimilation in Cardiovascular Fluid-Structure Interaction Problems: an introduction, Luca Bertagna, M. D'Elia, M. Perego, A. V.,



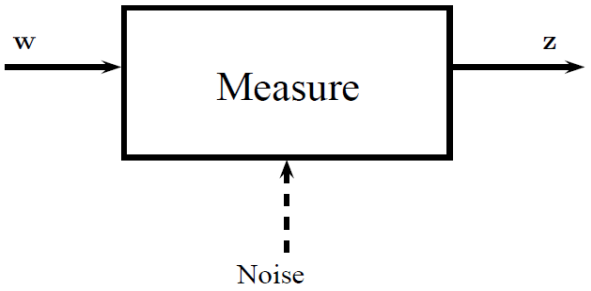
Outline 1

- Probabilistic estimation theory
  - In-Out estimators (stationary)
  - Linear Kalman Filter
  - Extended Kalman Filter
  - Unscented Kalman Filter





Probabilistic Estimation Theory



In-Out Estimation (stationary)

$\mathbf{w}$ = vector to be estimated

$\mathbf{z}$ = vector of observation

$\hat{\mathbf{w}}$ = estimator $\qquad \mathbf{e} \equiv \hat{\mathbf{w}} - \mathbf{w}$

Possible approach: minimize a "cost function" $J(\mathbf{e}) = \mathbf{e}^T \mathrm{E} \mathbf{e}$ where E is a s.p.d matrix

$$\hat{\mathbf{w}} = \arg\min \mathcal{J}(\mathbf{e}) \equiv \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} J(\mathbf{e}) p(\mathbf{w},\mathbf{z}) d\mathbf{z} d\mathbf{w}.$$

After some algebra we get $\boxed{\hat{\mathbf{w}} = \mathcal{E}(\mathbf{w}|\mathbf{z}).}$ Minimum Variance estimator

MV estimator is unbiased: $\mathcal{E}(\hat{\mathbf{w}}_{MV}) = \mathcal{E}(\hat{\mathbf{w}} p_{\mathbf{z}} d\mathbf{z}) = \int_{\mathbb{R}^n} \hat{\mathbf{w}}_{MV} p_{\mathbf{z}} d\mathbf{z} = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathbf{w} p_{\mathbf{w}|\mathbf{z}} d\mathbf{w} p_{\mathbf{z}} d\mathbf{z} =$

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \mathbf{w} p(\mathbf{w},\mathbf{z}) d\mathbf{w} d\mathbf{z} = \mathcal{E}(\mathbf{w}).$$

- $\mathcal{E}(\cdot)$ is the *expected value* operator;
- $p_{\mathbf{w}}(\cdot)$ is the marginal p.d.f. of $\mathbf{w}$;
- $p(\mathbf{w},\mathbf{z})$ is the joint p.d.f. of $\mathbf{w}$ and $\mathbf{z}$.

Other possible (reasonable) estimators

MAP: The estimator is the result of maximization of the *a posteriori* p.d.f. $p_{\mathbf{w}|\mathbf{z}}$

This is the p.d.f. of **w** conditionally to a particular measure

ML: The estimator is the result of maximization of likelihood $p_{\mathbf{z}|\mathbf{w}}$

This is the **w** that maximizes the probability to get a particular measure

Remark:
$\mathbf{w}_{ML}$ can be regarded as the limit of $\mathbf{w}_{MAP}$ when we do not have any statistical *a priori* knowledge on **w** (marginal p.d.f. for **w**)

Remark 2:
A fundamental tool in these considerations is the Bayes Theorem

$$p_{\mathbf{w}|\mathbf{z}} = \frac{p_{\mathbf{wz}}(\mathbf{w},\mathbf{z})}{p_{\mathbf{z}}(\mathbf{z})}$$

Example 1

Assume that w and z are two scalar variables such that

$$z = Hw + \nu$$

with $w \sim \mathcal{G}(0, \lambda_w)$, and the noise $\nu \sim \mathcal{G}(0, r)$ is assumed to be uncorrelated with w. Then it is possible to verify that $z \sim \mathcal{G}(0, H^2\lambda_w + r)$ and $\lambda_{wz} = H\lambda_w$. In fact

$$\begin{aligned}
&\mathcal{E}(z) = H\mathcal{E}(w) + \mathcal{E}(\nu) = 0 \\
&\mathcal{V}(z) = \mathcal{E}\left(z^2\right) = \mathcal{E}\left(H^2w^2 + 2Hw\nu + \nu^2\right) = H^2\lambda_w + 0 + r \\
&\mathcal{V}(z, w) = \mathcal{E}(zw) = \mathcal{E}\left(Hw^2 + \nu w\right) = H\lambda_w.
\end{aligned}$$

In this case

$$\widehat{\mathbf{w}}_{MV} = \widehat{\mathbf{w}}_{MAP} = \frac{H\lambda_w}{H^2\lambda_w + r} z = \frac{H}{H^2 + r/\lambda_w} z$$

$$\widehat{\mathbf{w}}_{ML} = \frac{\lambda_w}{H\lambda_w} z = \frac{z}{H}$$

As expected, $\lim_{\lambda_w \to \infty} \widehat{\mathbf{w}}_{MAP} = \widehat{\mathbf{w}}_{ML}$.

Recall: Gaussian distribution of a n-dimensional vector

$$p_{\mathbf{w}} = \frac{1}{\sqrt{(2\pi)^n |\Lambda|}} \exp\left(-\frac{1}{2}(\mathbf{w} - \mathcal{E}(\mathbf{w}))^T \Lambda^{-1} (\mathbf{w} - \mathcal{E}(\mathbf{w}))\right)$$

Example 2

Example 2.3 Assume now that $\mathbf{w}$ and $\mathbf{z}$ are n-dimensional vectors, $\mathbf{w} \sim \mathcal{G}(\boldsymbol{\mu}, \Lambda)$ and

$$\mathbf{z} = \mathrm{H}\mathbf{w} + \boldsymbol{\nu}$$

where $\boldsymbol{\nu} \sim \mathcal{G}(0, \mathrm{R})$ is noise independent of $\mathbf{w}$. H is called *observation matrix*.

It is possible to prove that $\mathbf{z} \sim \mathcal{G}(\boldsymbol{\mu}_\mathbf{z}, \Lambda_\mathbf{z},)$ with

$$\begin{aligned}\boldsymbol{\mu}_\mathbf{z} &= \mathcal{E}\left(\mathrm{H}\mathbf{w} + \boldsymbol{\nu}\right) = \mathrm{H}\boldsymbol{\mu} \\ \Lambda_\mathbf{z} &= \mathcal{E}\left((\mathbf{z} - \boldsymbol{\mu}_\mathbf{z})^T(\mathbf{z} - \boldsymbol{\mu}_\mathbf{z})\right) = (\ldots) = \mathrm{H}\Lambda\mathrm{H}^T + \mathrm{R}.\end{aligned}$$

For the conditional probabilities, we find that

$$p_{\mathbf{w}|\mathbf{z}} = \frac{\sqrt{|\Lambda_\mathbf{z}|}}{\sqrt{(2\pi)^n |\Lambda||\mathrm{R}|}} \exp(-\frac{1}{2}\mathrm{J})$$

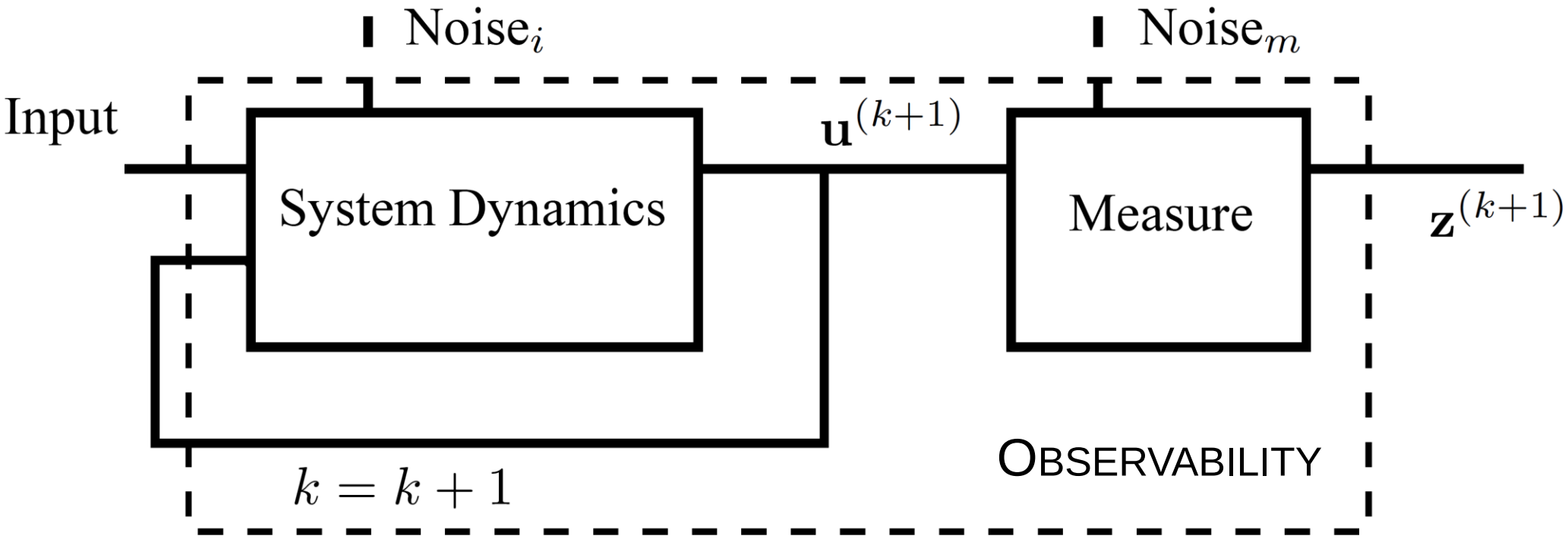
where $\mathrm{J} = (\mathbf{w} - \widehat{\mathbf{w}})^T \Lambda_\mathbf{e}^{-1} (\mathbf{w} - \widehat{\mathbf{w}})$ and $\Lambda_\mathbf{e}^{-1} = \Lambda^{-1} + \mathrm{H}^T\mathrm{R}\mathrm{H}$ and $\widehat{\mathbf{w}} \equiv \mathcal{E}\left(p_{\mathbf{w}|\mathbf{z}}\right) = \Lambda_\mathbf{e}\left(\mathrm{H}^T\mathrm{R}^{-1}\mathbf{z} + \Lambda^{-1}\boldsymbol{\mu}\right)$. This is both the MV and MAP estimator.

Moreover, we find that $p_{\mathbf{z}|\mathbf{w}}$ has average $\mathrm{H}\mathbf{w}$ and $\Lambda_{\mathbf{z}|\mathbf{w}} = \mathrm{R}$. If we maximize the likelihood, we obtain

$$\widehat{\mathbf{w}}_{ML} = \left(\mathrm{H}^T\mathrm{R}^{-1}\mathrm{H}\right)^{-1} \mathrm{H}^T\mathrm{R}^{-1}\mathbf{z}.$$

It is easy to verify that the MV/MAP estimator is unbiased and the ML estimator is obtained by the MAP, when $\Lambda^{-1} \to 0$.

Dynamical (sequential) estimation



$$\mathbf{u}^{(k)} = \mathrm{A}_{k-1}\mathbf{u}^{(k-1)} + \mathbf{b}^{(k-1)}$$

$$\mathbf{z}^{(k)} = \mathrm{H}_k\mathbf{u}^{(k)} + \boldsymbol{\nu}^{(k)}$$

$$\mathbf{b}^{(k)} \sim \mathcal{G}(\mathbf{0}, \mathrm{Q}_{k-1})$$

$$\boldsymbol{\nu}^{(k)} \sim \mathcal{G}(\mathbf{0}, \mathrm{R}_{k-1})$$

Estimate of the state: let us split the operation into steps.

THE KALMAN FILTER

- PREDICTION, deterministic and based on the past

$$\mathbf{u}_p^{(k)} = \mathrm{A}_{k-1}\mathbf{u}_*^{(k-1)}$$

$\mathbf{u}_*^{(k-1)} =$ our best knowledge of state in *(k-1)*

- CORRECTION, i.e. improvement due to the current observation

$$\mathbf{u}_c^{(k)} = \mathrm{L}_k\mathbf{u}_p^{(k)} + \mathrm{K}_k\mathbf{z}^{(k)}$$

How to select L_k and K_k?

Set

$$\mathbf{e}_p^{(k)} = \mathbf{u}_p^{(k)} - \mathbf{u}^{(k)}, \quad \mathbf{e}_c^{(k)} = \mathbf{u}_c^{(k)} - \mathbf{u}^{(k)}$$

If we force the correction to be unbiased,

$$\mathcal{E}\left(\mathbf{e}_c^{(k)}\right) = 0$$

$$\mathrm{L}_k \mathcal{E}\left(\mathbf{e}_p^{(k)}\right) + \mathrm{K}_k \mathcal{E}\left(\boldsymbol{\nu}^{(k)}\right) + \left(\mathrm{L}_k + \mathrm{K}_k \mathrm{H}_k - \mathrm{I}\right) \mathcal{E}\left(\mathbf{u}^{(k)}\right) = 0$$

$$\mathrm{L}_k + \mathrm{K}_k \mathrm{H}_k - \mathrm{I} = 0 \Rightarrow \mathrm{L}_k = \mathrm{I} - \mathrm{K}_k \mathrm{H}_k$$

$$\mathbf{u}_c^{(k)} = \mathbf{u}_p^{(k)} + \mathrm{K}_k\left(\mathbf{z}^{(k)} - \mathrm{H}_k \mathbf{u}_p^{(k)}\right)$$

Gain matrix

Innovation

Selection of the gain matrix: MV approach

Variance matrices: recursive formulas

$$\Lambda_p^{(k)} = \mathrm{A}_{k-1}\Lambda_c^{(k-1)}\mathrm{A}_{k-1}^T + \mathrm{Q}_{k-1}$$

$$\boxed{\Lambda_c^{(k)} = (\mathrm{I} - \mathrm{K}_k\mathrm{H}_k)\Lambda_p^{(k)}(\mathrm{I} - \mathrm{K}_k\mathrm{H}_k)^T + \mathrm{K}_k\mathrm{R}_k\mathrm{K}_k^T}$$

Joseph formula

If we minimize the variance of the estimate error

$$\mathcal{E}\left(\|\mathbf{e}_c^{(k)}\|^2\right)$$

$$\mathrm{H}_k\Lambda_p^{(k)}(\mathrm{I} - \mathrm{K}_k\mathrm{H}_k)^T - \mathrm{R}_k\mathrm{K}_k^T = 0 \Rightarrow \boxed{\mathrm{K}_k = \Lambda_p^{(k)}\mathrm{H}_k^T\left(\mathrm{H}_k\Lambda_p^{(k)}\mathrm{H}_k^T + \mathrm{R}_k\right)^{-1}}$$

Kalman GM

KALMAN FILTER: SUMMARY

1. Prediction

(a) $\mathbf{u}_p^{(k)} = \mathrm{A}_{k-1}\mathbf{u}_c^{(k-1)}$

(b) $\Lambda_p^{(k)} = \mathrm{A}_{k-1}\Lambda_c^{(k-1)}\mathrm{A}_{k-1}^T + \mathrm{Q}_{k-1}$

2. Correction

Kalman gain:

$$\mathrm{K}_k = \Lambda_p^{(k)}\mathrm{H}_k^T\left(\mathrm{H}_k\Lambda_p^{(k)}\mathrm{H}_k^T + \mathrm{R}_k\right)^{-1}.$$

(a) State estimate

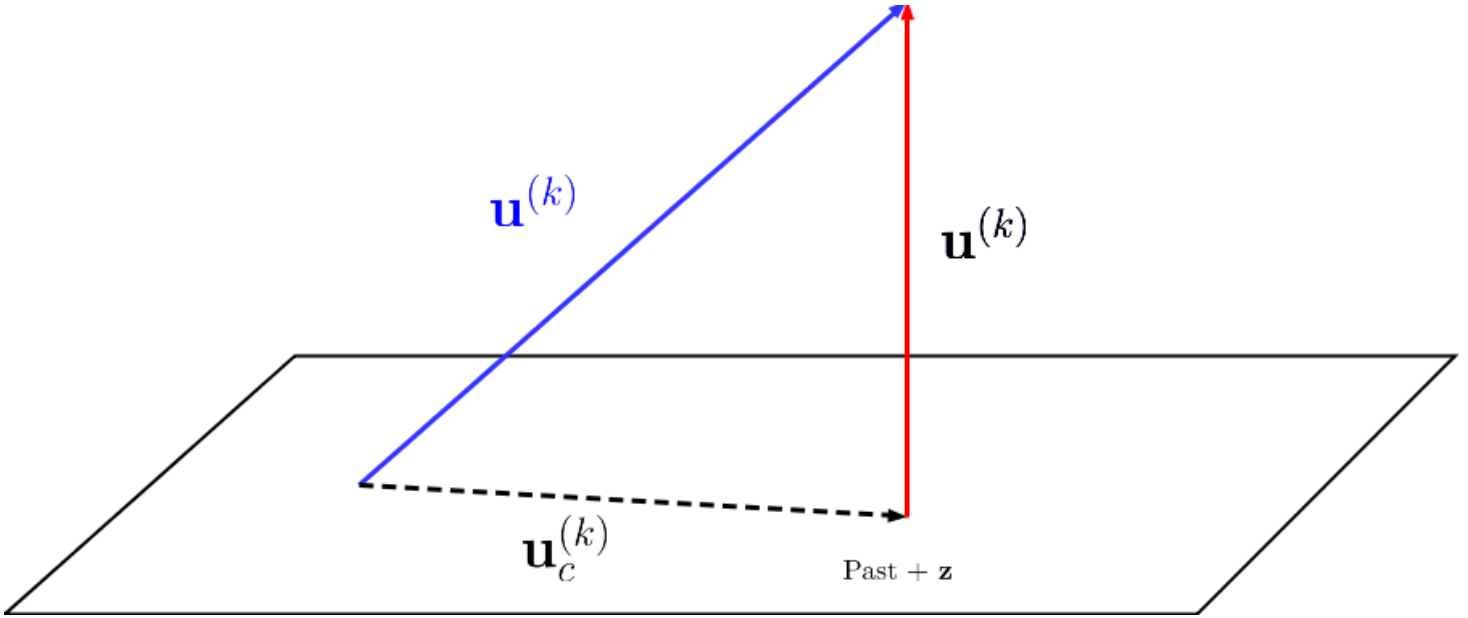
$$\mathbf{u}_c^{(k)} = \mathbf{u}_p^{(k)} + \mathrm{K}_k(\mathbf{z}^{(k)} - \mathrm{H}_k\mathbf{u}_p^{(k)}).$$

(b) Estimator Covariance

$$\Lambda_c^{(k)} = (\mathrm{I} - \mathrm{K}_k\mathrm{H}_k)\,\Lambda_p^{(k)}.$$

PROPERTIES OF THE KF

Innovation



1. $\mathbf{z}^{(k)} - \mathrm{H}_k \mathbf{u}_p^{(k)} = \mathrm{H}_k \mathbf{u}^{(k)} + \boldsymbol{\nu}^{(k)} - \mathrm{H}_k \mathbf{u}_p^{(k)} = \mathrm{H}_k \mathbf{e}_p^{(k)} + \boldsymbol{\nu}^{(k)}.$

2. $\mathcal{E}\left(\mathbf{z}^{(k)} - \mathrm{H}_k \mathbf{u}_p^{(k)}\right) = \mathrm{H}_k \mathcal{E}\left(\mathbf{e}_p^{(k)}\right) + \mathcal{E}\left(\boldsymbol{\nu}^{(k)}\right) = 0.$

3. Innovation is **orthogonal to the past** (with respect to the covariance scalar product)

Moreover: $\mathcal{E}\left(\mathbf{u}_c^{(k)} \mathbf{e}_c^{(k),T}\right) = 0.$

2.Variance reduction of the correction

Let us consider the ``pseudo-observation''

$$\mathbf{z}_{po}^{(k)} = \mathrm{H}_k \mathbf{u}_p^{(k)} + \boldsymbol{\nu}^{(k)}.$$

Then we can write:

$$\Lambda_{po} = \mathrm{H}_k \Lambda_p^k \mathrm{H}_k^T + \mathrm{R}_k, \quad \Lambda_{p,po}^{(k)} = \Lambda_p^{(k)} \mathrm{H}_k^T$$

Then the Kalman matrix and the correction variance read

$$\mathrm{K}_k = \Lambda_{p,po}^{(k)} (\Lambda_p^{(k)})^{-1}, \quad \Lambda_c^{(k)} = \Lambda_p^{(k)} - \underbrace{\Lambda_{p,po}^{(k)} (\Lambda_{po}^{(k)})^{-1} \Lambda_{p,po}^{(k)}}_{\text{s.p.d.}}.$$

3.Asymptotic behavior (for time-independent dynamics)

Recursive formula for the predictor variance:

$$\Lambda_p^{(k+1)} = \mathrm{A}_k \Lambda_c^{(k)} \mathrm{A}_k^T + \mathrm{Q}_k =$$
$$\mathrm{A}_k \left(\mathrm{I} - \mathrm{K}_k \mathrm{H}_k \Lambda_p^{(k)} \mathrm{A}_k^T + \mathrm{Q}_k =\right.$$
$$\mathrm{A}_k \Lambda_p^{(k)} \mathrm{A}_k^T + \mathrm{Q}_k - \mathrm{A}_k \Lambda_p^{(k)} \mathrm{H}_k^T (\mathrm{H}_k \Lambda_p^{(k)} \mathrm{H}_k^T + \mathrm{R}_k)^{-1} \mathrm{H}_k \Lambda_c^{(k)} \mathrm{A}_k^T$$

If the matrices A, Q and H,R are independent of time:

$$\Lambda_p^{(k+1)} = \mathrm{A} \Lambda_p^{(k)} \mathrm{A}^T + \mathrm{Q} - \mathrm{A} \Lambda_p^{(k)} \mathrm{H}^T (\mathrm{H} \Lambda_p^{(k)} \mathrm{H}^T + \mathrm{R})^{-1} \mathrm{H} \Lambda_p^{(k)} \mathrm{A}^T$$

Difference Riccati Equation

The equilibrium solution of this matrix reads:

$$\Lambda_p = \mathrm{A} \Lambda_p \mathrm{A}^T + \mathrm{Q} - \mathrm{A} \Lambda_p \mathrm{H}^T (\mathrm{H} \Lambda_p \mathrm{H}^T + \mathrm{R})^{-1} \mathrm{H} \Lambda_p \mathrm{A}^T$$

Algebraic Riccati Equation

AN EXAMPLE (easy)

Let us consider the scalar case, with

$$u^{(k)} = u^{(k-1)} \quad (A = 1, b = 0)$$
$$z^{(k)} = u^{(k)} + \nu^{(k)} \quad (H = 1, \nu \sim \mathcal{G}(0,1)).$$

Assume also that the initial data $u^{(1)} \sim \mathcal{G}(\mu, 1)$. Set $u_p^{(1)} = \mu$.
Then, the Kalman filter formulas read

$$u_p^{(k)} = u_c^{(k-1)}, \quad \lambda_p^{(k)} = \lambda_c^{(k-1)}$$
$$K_k = \frac{\lambda_p^{(k)}}{\lambda_p^{(k)} + r^{(k)}} = \frac{\lambda_p^{(k)}}{\lambda_p^{(k)} + 1}$$
$$u_c^{(k)} = u_p^{(k)} + \frac{\lambda_p^{(k)}}{\lambda_p^{(k)} + 1}(z^{(k)} - u_p^{(k)}) = \frac{1}{\lambda_p^{(k)} + 1} u_p^{(k)} + \frac{\lambda_p^{(k)}}{\lambda_p^{(k)} + 1} z^{(k)}$$
$$\lambda_c^{(k)} = \frac{\lambda_p^{(k)}}{\lambda_p^{(k)} + 1}$$

So now we have

$$u_p^{(1)} = \mu, \lambda_p^{(1)} = 1, K_1 = \frac{1}{2}, u_c^{(1)} = \frac{\mu + z^{(1)}}{2} = u_c^{(2)}.$$

Notice that the prediction at $k = 2$ is just the sample average of the "past" and the new data. Similarly we obtain at a generic step k

$$u_p^{k+1} = u_c^k = \frac{\mu + \sum_{j=1}^{k} z^{(j)}}{k+1}.$$

Actually, we have the arithmetic average of the available data at t^k: it makes a lot of sense!

By induction one can check that $\lambda_p^{(k)} = \frac{1}{k}$. Consequences:

1. $\lim_{k\to\infty} \lambda_p^{(k)} = 0$, i.e. the prediction is asymptotically exact.
2. 0 is the only solution to the ARE $\lambda = \lambda/(1+\lambda)$.
3. The Kalman filter is asymptotically stable. Notice that the dynamic system is NOT asymptotically stable.

REMARKS:

The interplay between stability of the system and of its Kalman estimator is crucial!

We can barely accept an unstable estimator of a stable system!

Theorem: If the system is stable, the estimator is stable and Λ_p converges to the solution of the ARE

Numerical issues for the KF: the cost is proportional to n^3 where n is the size of the system.

Numerical stabiliy issues in the propagation of rounding errors

THE NON-LINEAR CASE: Extended KF (EKF)

$$\mathbf{u}^{(k)} = \mathrm{A}(\mathbf{u}^{(k-1)}) + \mathbf{b}^{(k)}\mathbf{z}^{(k)} = \mathrm{H}(\mathbf{u}^{(k)}) + \boldsymbol{\nu}^{(k)}$$

Linearization:

$$\mathcal{A}'(\cdot) = \frac{\partial \mathrm{A}(\cdot)}{\partial \mathbf{u}}, \quad \mathcal{H}'(\cdot) = \frac{\partial \mathrm{H}(\cdot)}{\partial \mathbf{u}}.$$

1. Prediction: $\mathbf{u}_p^{((k))} = \mathrm{A}(\mathbf{u}_c^{((k-1))})$, with

$$\Lambda_p^{(k)} = \mathcal{A}'(\mathbf{u}_c^{(k-1)})\Lambda_c^{(k-1)}(\mathcal{A}'(\mathbf{u}_c^{(k-1)}))^T + \mathrm{Q}_{k-1}.$$

2. Kalman gain: $\mathrm{K}_k = \Lambda_p^{(k)}(\mathcal{H}'(\mathbf{u}_p^{((k))}))^T \left(\mathcal{H}'(\mathbf{u}_p^{((k))})\Lambda_p^{(k)}\mathcal{H}'(\mathbf{u}_p^{((k))}) + \mathrm{R}_k\right)^{-1}$.

3. Correction: $\mathbf{u}_c^{(k)} = \mathbf{u}_p^{(k)} - \mathrm{K}_k\left(\mathbf{z}^{(k)} - \mathrm{H}(\mathbf{u}_p^{(k)})\right)$.

4. $\Lambda_c^{(k)} = \left(\mathrm{I} - \mathrm{K}_k\mathcal{H}'(\mathbf{u}_p^{((k))})\right)\Lambda_p^{(k)}$.

There are two main drawbacks in this approach.

1. *Computational Costs.* The computation of the tangent operators can be fairly expensive.

2. *Bias*: The estimates are in general biased, $\mathcal{E}\left(\mathbf{e}_c^{(k)}\right) \neq 0$.

Identification problem: the *state augmentation* approach

$$\begin{aligned}\mathbf{u}^{(k)} &= \mathrm{A}(\vartheta)\mathbf{u}^{(k-1)} + \mathbf{b}^{(k)} \\ \mathbf{z}^{(k)} &= \mathrm{H}(\vartheta)\mathbf{u}^{(k)} + \boldsymbol{\nu}^{(k)}. \\ \vartheta^{(k)} &= \vartheta^{(k-1)} + \varepsilon^{(k)}\end{aligned}$$

$$\mathbf{v}^{(k)} = \begin{bmatrix}\mathbf{u}^{(k)} \\ \vartheta^{(k)}\end{bmatrix} \Rightarrow \mathbf{v}^{(k)} = \begin{bmatrix}\mathrm{A}(\vartheta^{(k-1)})\mathbf{u}^{(k-1)} \\ \vartheta^{(k-1)}\end{bmatrix} + \begin{bmatrix}\mathbf{b}^{(k)} \\ \varepsilon^{(k)}\end{bmatrix}.$$

$$\mathbf{z}^{(k)} = \mathrm{H}(\vartheta^{(k)})\mathbf{u}^{(k)} + \boldsymbol{\nu}^{(k)}.$$

EKF with :

$$\mathcal{A}'(\mathbf{v}) = \begin{bmatrix}\mathrm{A}(\vartheta) & \dfrac{\partial \mathrm{A}}{\partial \vartheta}\mathbf{u} \\ O & \mathrm{I}\end{bmatrix}, \mathcal{H}'(\mathbf{v}) = \begin{bmatrix}\mathrm{H}(\vartheta) & \dfrac{\partial \mathrm{H}}{\partial \vartheta}\mathbf{u}\end{bmatrix}$$

THE NON-LINEAR CASE II: Unscented KF (EKF)

Basic idea: we take some samples and then approximate the non-linear dynamics with the average/variance of the evolving samples.

Example: scalar variable

$$u^{(k)} \text{ with } \mathcal{E}\left(u^{(k)}\right) = \mu \text{ and } \mathcal{V}\left(u^{(k)}\right) = \lambda^2$$

Samples: $s_{1,2} = \mu \pm \lambda$

Approximate dynamics

$$\begin{cases} \mathcal{E}\left(\mathrm{A}(u^{(k)})\right) \approx \dfrac{1}{2}\left(\mathrm{A}(s_1) + \mathrm{A}(s_2)\right) = \overline{\mathrm{A}}. \\ \mathcal{V}\left(\mathrm{A}(u^{(k)})\right) \approx \dfrac{1}{2}\left((\mathrm{A}(s_1) - \overline{A})^2 + (\mathrm{A}(s_2) - \overline{\mathrm{A}})^2\right) \end{cases}$$

For vector state variables, a possible choice of the smaples (s-points):

$$S^{(i)} = \lambda_i \mathbf{e}_i$$

Std deviation

i-th unit vector

1. Sampling Let $Chol()$ denote the Choleski decomposition of a s.p.d. matrix. We take

$$\begin{aligned}
\mathrm{C}_{k-1} &= Chol(\Lambda_p^{(k-1)}) = \left(\Lambda_p^{(k-1)}\right)^{1/2} \\
\mathbf{u}_i^{(k-1)} &= \mathbf{u}_c^{(k-1)} + \mathrm{C}_{(k-1)}^T S^{(i)}, \quad i = 1, 2, \ldots r.
\end{aligned}$$

2. Prediction. Let α_i be the coefficients of the sample average.

$$\begin{aligned}
\mathbf{u}_{p,i}^{(k)} &= \mathrm{A}(\mathbf{u}_{p,i}^{(k-1)}) \quad \text{sample evolution} \\
\mathbf{u}_p^{(k)} &= \sum_i \alpha_i \mathbf{u}_{p,i}^{(k)}, \quad \Lambda_p^{(k)} = \sum_i \alpha_i \left(\mathbf{u}_{p,i}^{(k)} - \mathbf{u}_p^{(k)}\right)\left(\mathbf{u}_{p,i}^{(k)} - \mathbf{u}_p^{(k)}\right)^T
\end{aligned}$$

3. Correction

$$\begin{aligned}
\mathbf{z}_i^{(k)} &= \mathrm{H}(\mathbf{u}_{p,i}^{(k)}) \\
\Lambda_{po}^{(k)} &= \sum_i \alpha_i \left(\mathbf{z}_i^{(k)} - \mathrm{H}(\mathbf{u}_p^{(k)})\right)\left(\mathbf{z}_i^{(k)} - \mathrm{H}(\mathbf{u}_p^{(k)})\right)^T + \mathrm{Q}_k \\
\Lambda_{p,po}^{(k)} &= \sum_i \alpha_i \left(\mathbf{z}_i^{(k)} - \mathrm{H}(\mathbf{u}_p^{(k)})\right)\left(\mathbf{u}_{p,i}^{(k)} - \mathbf{u}_p^{(k)}\right)^T \\
\mathrm{K}_k &= \Lambda_{p,po}^{(k)} \left(\Lambda_{po}^{(k)}\right)^{-1} \\
\mathbf{u}_c^{(k)} &= \mathbf{u}_p^{(k)} + \mathrm{K}_k \left(\mathbf{z}^{(k)} - \mathrm{H}(\mathbf{u}_p^{(k)})\right) \\
\Lambda_c^k &= \Lambda_p^k - \Lambda_{p,po}^{(k)} \left(\Lambda_{po}^{(k)}\right)^{-1} \Lambda_p^{(k)}.
\end{aligned}$$

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Sequential parameter estimation for fluid-structure problems. Application to hemodynamics

Cristóbal Bertoglio — Philippe Moireau — Jean-Frederic Gerbeau

N° 7657

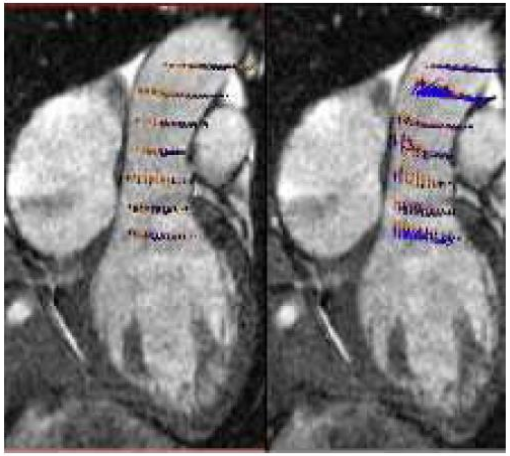
Juin 2011

Observation, Modeling, and Control for Life Sciences

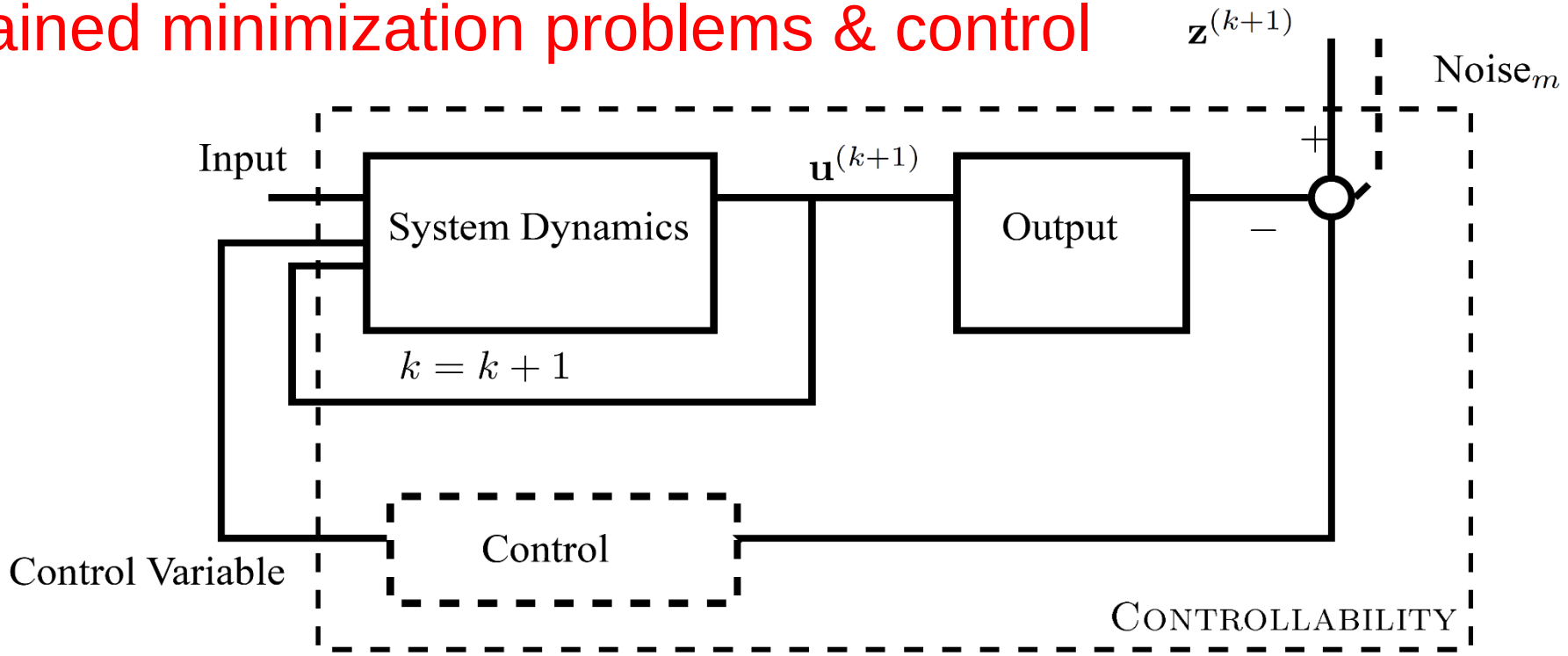
ISRN INRIA/RR--7657--FR+ENG

Outline 2

- Variational methods of DA
  - (In-Out estimates: Least Squares)
  - Constrained minimization
    - One-shot methods (the KKT system)
    - Sensitivies
    - Adjoint computation of the gradient
  - Computational Costs: solution reduction



Constrained minimization problems & control



Ingredients

A functional $\mathcal{J}$ to be minimized. In general, this will be discrepancy between the results obtained by the mathematical (numerical) model and the available data.

A mathematical model, describing the dynamics of interest. In our problems, this will be given by an incompressible fluid, possibly coupled with other problems, such the structure or models for the peripheral circulation. The state variables will include the pressure and the velocity of the fluid.

A set of measurements (output) referring to a function of the state variable.

A control variable (CV), a variable which needs to be tuned to get the minimization done.

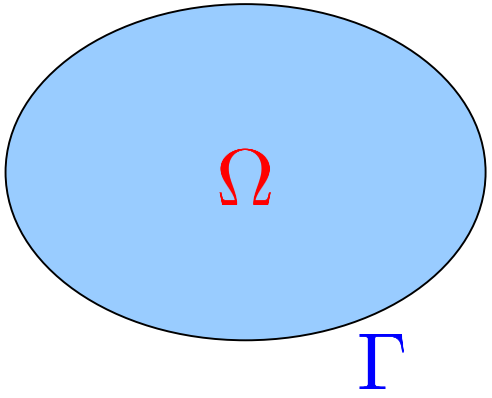
A reference example

$\mathcal{F}(u, \boldsymbol{\alpha}) = 0$

$$-\nabla \cdot (\mu \nabla u) + \mathbf{b} \cdot \nabla u + u^3 = \sum_{i=1}^{k} \alpha_i f_i \qquad \text{in } \Omega$$

$$u = 0 \quad \text{on} \quad \Gamma.$$

$\boldsymbol{\alpha} = \mathrm{CV}$



Functional to be minimized

$$\mathcal{J}(u) = \frac{1}{2} \int_{\Omega} (u(\mathbf{x}) - d(\mathbf{x}))^2 d\mathbf{x},$$

Regularized version

$$\mathcal{J}_R(u) = \mathcal{J}(u) + \frac{\sigma}{2} \|\boldsymbol{\alpha}\|^2$$

Limitation of the control cost

Tychonov regularization

The Lagrange multiplier approach:

$$\mathcal{L}(u, \boldsymbol{\alpha}, \chi) = \mathcal{J}_R - \chi^* \mathcal{F}(u, \boldsymbol{\alpha})$$

Unconstrained minimization

duality

In the example:

$$\mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2) = \mathcal{J}_R(u, \boldsymbol{\alpha}) - \int_\Omega \chi_1 \left(-\nabla \cdot (\mu \nabla u) + \mathbf{b} \cdot \nabla u + u^3 - \sum_{i=1}^{k} \alpha_i f_i \right) - \int_\Gamma \chi_2 u.$$

Stationary points (KKT system)

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \chi} = 0 & \text{State equations} \\ \dfrac{\partial \mathcal{L}}{\partial u} = 0 & \text{Adjoint/Co} - \text{state equations} \\ \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}} = 0 & \text{Optimality conditions} \end{cases}$$

Gateaux derivative

$$\frac{\partial \mathcal{L}}{\partial \chi_1} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\mathcal{L}(u, \boldsymbol{\alpha}, \chi_1 + \varepsilon \delta_{\chi_1}, \chi_2) - \mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2) \right)$$

δ_{χ_1} = admissible variation

$$\Rightarrow \quad \int_\Omega \chi_1 \left(-\nabla \cdot (\mu \nabla u) + \mathbf{b} \cdot \nabla u + u^3 - \sum_{i=1}^{k} \alpha_i f_i \right) = 0.$$

$$\frac{\partial \mathcal{L}}{\partial \chi_2} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2 + \varepsilon \delta_{\chi_2}) - \mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2) \right) = \int_\Gamma \delta_{\chi_2} u = 0$$

$$\Rightarrow \quad u = 0 \text{ on } \Gamma$$

$$\frac{\partial \mathcal{L}}{\partial u} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\mathcal{L}(u + \varepsilon \delta_u, \boldsymbol{\alpha}, \chi_1, \chi_2) - \mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2) \right) =$$

$$\int_\Omega (u - d)\, \delta_u - \int_\Omega \chi_1 \left(-\nabla \cdot (\mu \nabla \delta_u) + \mathbf{b} \cdot \nabla \delta_u + 3u^2 \delta_u \right) - \int_\Gamma \chi_2 \delta_u = 0.$$

$$\begin{cases} -\nabla \cdot (\mu \nabla \chi_1) - \nabla \cdot (\mathbf{b}\chi_1) + 3u^2\chi_1 = u - d & \text{in } \Omega \\ \chi_1 = 0 & \text{on } \Gamma \\ \chi_2 = -\mu \nabla \chi_1 \cdot \mathbf{n} - \mathbf{b} \cdot \mathbf{n}\chi_1 & \text{on } \Gamma. \end{cases}$$

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(\mathcal{L}(u, \alpha_1, \alpha_2, \ldots, \alpha_i + \delta_{\alpha_i}, \ldots, \alpha_k, \chi_1, \chi_2) - \mathcal{L}(u, \boldsymbol{\alpha}, \chi_1, \chi_2) \right) =$$

$$\left(\int_\Omega \chi_1 f_i + \sigma \alpha_i \right) \delta_{\alpha_i}$$

State
$$\begin{cases} -\nabla \cdot (\mu \nabla u) + \mathbf{b} \cdot \nabla u + u^3 = \sum_{i=1}^{k} \alpha_i f_i & \text{in } \Omega \\ u = 0 \quad \text{on } \Gamma \end{cases}$$

Adjoint
$$\begin{cases} -\nabla \cdot (\mu \nabla \chi_1) - \nabla \cdot (\mathbf{b}\chi_1) + 3u^2\chi_1 = u - d \text{ in } \Omega \\ \chi_1 = 0 \text{ on } \Gamma \\ \chi_2 = -\mu \nabla \chi_1 \cdot \mathbf{n} - \mathbf{b} \cdot \mathbf{n}\chi_1 \text{ on } \Gamma. \end{cases}$$

Optimality $\alpha_i = \alpha_i - \dfrac{\rho}{\sigma} \displaystyle\int_\Omega \chi_1 f_i \qquad i = 1, \ldots, k.$

Summary
(abstract framwork)

Problem Find the control variable α and the state u that minimizes

$$\mathcal{J}_R(u,\alpha) = \mathcal{J}(\alpha,u) + \sigma\|\alpha\|^2,$$

under the constraint $\mathcal{F}(u,\alpha) = 0$.

Lagrangian functional and its stationary points (One-shot) Set

$$\mathcal{L}(u,\alpha,\chi) = \mathcal{J}_R(u,\alpha) - \underbrace{\chi^*\mathcal{F}(u,\alpha)}_{\text{duality product}}$$

The stationary points of this (unconstrained) functional are s.t.

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\chi} &= 0 \quad \text{state} \\ \frac{\partial\mathcal{L}}{\partial u} &= 0 \quad \text{adjoint} \\ \frac{\partial\mathcal{L}}{\partial\alpha} &= 0 \quad \text{optimality.} \end{aligned} \tag{3.9}$$

More precisely, we have

$$\frac{\partial\mathcal{L}}{\partial\chi} = \lim_{\varepsilon\to 0}\left(\mathcal{L}(u,\alpha,\chi+\delta_\chi) - \mathcal{L}(u,\alpha,\chi)\right) = \delta^*_\chi\mathcal{F}(u,\alpha) = 0 \Rightarrow \boxed{\mathcal{F}(u,\alpha) = 0}$$

$$\frac{\partial\mathcal{L}}{\partial u} = \lim_{\varepsilon\to 0}\left(\mathcal{L}(u+\delta_u,\alpha,\chi) - \mathcal{L}(u,\alpha,\chi)\right) = \left(\frac{\partial\mathcal{J}_R}{\partial u}\right)|_{u,\alpha}\delta_u - \chi^*\left(\frac{\partial\mathcal{F}}{\partial u}\right)|_{u,\alpha}\delta_u = 0 \Rightarrow$$
$$\boxed{\left(\frac{\partial\mathcal{F}}{\partial u}\right)^*|_{u,\alpha}\chi = \left(\frac{\partial\mathcal{J}_R}{\partial u}\right)^*|_{u,\alpha}}$$

$$\frac{\partial\mathcal{L}}{\partial\alpha} = \lim_{\varepsilon\to 0}\left(\mathcal{L}(u,\alpha+\delta_\alpha,\chi) - \mathcal{L}(u,\alpha,\chi)\right) = \left(\frac{\partial\mathcal{J}_R}{\partial\alpha}\right)|_{u,\alpha}\delta_\alpha - \chi^*\left(\frac{\partial\mathcal{F}}{\partial\alpha}\right)|_{u,\alpha}\delta_\alpha = 0 \Rightarrow$$
$$\boxed{\left(\frac{\partial\mathcal{F}}{\partial\alpha}\right)^*|_{u,\alpha}\chi = \left(\frac{\partial\mathcal{J}_R}{\partial\alpha}\right)^*|_{u,\alpha}}$$

A DIFFERENT (non-Lagrange) APPROACH

1. Given a guess for α, called $\alpha^{(k)}$, we compute the corresponding state $\mathcal{F}(u^{(k)}, \alpha^{(k)}) = 0$.

2. With $u^{(k)}, \alpha^{(k)}$, we compute the gradient of $\mathcal{J}$ or more precisely $D\mathcal{J}/D\alpha$ (total derivative). $(\mathcal{J}_{\mathcal{R}})$

3. With $D\mathcal{J}/D\alpha$ we find an increment $\delta\alpha$ to update

$$\alpha^{(k+1)} = \alpha^{(k)} + \delta\alpha.$$

4. Test the convergence.

Critical step

Gradient computation through sensitivities

$$\frac{D\mathcal{J}_R}{D\alpha_i}\Big|_{\alpha_i^{(k)}} = \frac{\partial \mathcal{J}_R}{\partial u}\Big|_{u^{(k)}} \frac{\partial u}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}} + \frac{\partial \mathcal{J}_R}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}}$$

SENSITIVITIES:

$$\frac{\partial u}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}}, \quad \forall i = 1, 2, \ldots k$$

Notice that:

$$\mathcal{F}(u^{(k)}, \alpha^{(k)}) = 0 \Rightarrow \frac{D\mathcal{F}}{D\alpha_i}\Big|_{\alpha_i^{(k)}} = \frac{\partial \mathcal{F}}{\partial u}\Big|_{u^{(k)}} \frac{\partial u}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}} + \frac{\partial \mathcal{F}}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}} = 0$$

SENSITIVITIES Equations

$$\frac{\partial \mathcal{F}}{\partial u}\Big|_{u^{(k)}} \frac{\partial u}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}} = -\frac{\partial \mathcal{F}}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}}$$

In the Example:

$$\frac{D\mathcal{J}_R}{D\alpha_i}\Big|_{\alpha_i^{(k)}} = \int_\Omega (u-d)\frac{\partial u}{\partial \alpha_i} + \sigma\alpha_i = \int_\Omega (u-d)\phi_i + \sigma\alpha_i$$

$$\frac{\partial \mathcal{F}}{\partial u}\Big|_{u^{(k)}}(\bullet) = -\nabla\cdot(\mu\nabla\bullet) + \mathbf{b}\cdot\nabla\bullet + 3(u^{(k)})^2\bullet, \qquad \frac{\partial \mathcal{F}}{\partial \alpha_i}\Big|_{\alpha_i^{(k)}} = f_i$$

Set $\dfrac{\partial u}{\partial \alpha_i} = \phi_i$

We get the *sensitivities equations*

$$\begin{cases} -\nabla\cdot(\mu\nabla\phi_i) + \mathbf{b}\cdot\nabla\phi_i + 3(u^{(k)})^2\phi_i = f_i & \text{in } \Omega \\ \\ \phi_i = 0 & \text{on } \Gamma. \end{cases}$$

$$\frac{D\mathcal{J}}{D\alpha_i}\Big|_{\alpha_i^{(k)}} = \int_\Omega (u-d)\phi_i + \sigma\alpha_i.$$

Major drawback:
1 equation per CV

Gradient computation through adjoint problems

Let us consider the adjoint problem, where ρ now is the unknown

$$\left(\frac{\partial \mathcal{F}}{\partial u}\right)^* |_{u,\alpha}\rho = \left(\frac{\partial \mathcal{J}_R}{\partial u}\right)^* |_{u,\alpha}$$

Then, we can write

$$\frac{D\mathcal{J}_R}{D\alpha} = \frac{\partial \mathcal{J}_R}{\partial u}\frac{\partial u}{\partial \alpha} + \frac{\partial \mathcal{J}_R}{\partial \alpha} =$$

$$\boxed{\rho^* \frac{\partial \mathcal{F}}{\partial u}\frac{\partial u}{\partial \alpha}} + \frac{\partial \mathcal{J}_R}{\partial \alpha} = \boxed{\rho^* \frac{\partial \mathcal{F}}{\partial \alpha}} + \frac{\partial \mathcal{J}_R}{\partial \alpha}.$$

In the case of our example, the adjoint problem reads:

$$\begin{cases} -\nabla \cdot (\mu \nabla \rho) - \nabla \cdot (\mathbf{b}\rho) + 3u^2\rho = u - d \text{ in } \Omega \\ \rho = 0 \text{ on } \Gamma \end{cases}$$

Then, we can write the gradients as:

$$\frac{D\mathcal{J}_R}{D\alpha_k} = \sigma\alpha_k + \int_\Omega (u-d)\frac{\partial u}{\partial \alpha_k} =$$

$$\sigma\alpha_k + \int_\Omega \left(-\nabla\cdot(\mu\nabla\rho) - \nabla\cdot(\mathbf{b}\rho) + 3u^2\rho\right)\frac{\partial u}{\partial \alpha_k} = \text{by parts}$$

$$\sigma\alpha_k + \int_\Omega \boxed{\left(-\nabla\cdot(\mu\nabla\phi_k) + \mathbf{b}\cdot\nabla\phi_k + 3u^2\phi_k\right)}\rho =$$

$$\sigma\alpha_k + \int_\Omega f_k\rho.$$

Sensitivities equation

Pro: by solving 1! differential equation we retrieve all the sensitivities

Con: unsteady adjoint problems are backward in time – at the continuous level we need to know the solution of the state problem over the entire time interval and then come back! High storage/computational costs

A (classical) dilemma

First Discretize then Optimize (DO) vs.

First Optimize then Discretize (OD).

Which is better? The answer is basically problem-dependent.

With DO:

- we avoid inconsistencies induced by the numerical differentiation of the KKT system;
- we can even use automatic differentiation software;
- we can split an unsteady problem into a sequence of pseudo-steady optimization problems.

With OD:

- we do not deal with the differentiation of numerical artificial terms (like advection stabilization);
- We can use different grids on the different problems (usually they do need different grids);
- It can be easier to manage moving boundary problems: we do not need the derivative of the grid (which is part of the state of the system in DO with moving boundary problems) with respect to the optimization parameters.

Model reduction in a nutshell

In order to reduce computational costs, we may replace the state model with a ``low-dimensional'' or reduced order model.

We can represent the solution with a Galerkin expansion

$$u_h = \sum_{i=1}^{N} U_i \varphi_i .$$

We want to resort to a reduced model in the form

$$u_r = \sum_{i=1}^{M} R_i \rho_i .$$

To do: (1) find an appropriate basis set ρ_i

A possible approach is to pick up off-line snapshots of the state solution (no reduction) for different values of the CV (or different values in time)

(2) find M so to include only the relevant information

A possible approach is the Proper Orthogonal Decomposition (POD)

A general statement on Solution Reduction methods

$$u = \sum_{i=1}^{N} U_i \varphi_i$$

N ← Number of degrees of freedom

φ_i ← Basis functions

Two ``extreme'' choices

1 JEEP (General purpose)

Basis functions do not include any info of the problem, they can be (virtually) used for any problem

→ we need a large number of DOF N "

EXAMPLES: Finite Elements, Spectral Method

2 SMART (Educated)

Basis functions knows something of the problem, they are problem specific, not general but

→ we may (significantly) reduce N #

EXAMPLES: Sturm-Liouville Eigenfunctions, Reduced Basis Method, Proper Orthogonal Decomposition


forfun²
forfun²
BFGoodrich
301
35C

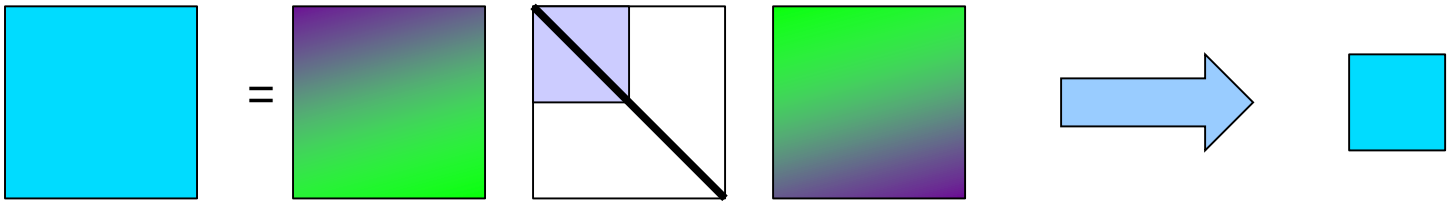
Two words on POD

Assume we have a basis that we want to reduce, from M_1 to M

1) average and covariance of the snapshots basis

$$\overline{\rho} = \sum_{j=1}^{M_1} \rho_j \qquad \mathbb{R}^{M_1 \times M_1} \ni \mathrm{C} = \frac{1}{M_1} \left[\rho_1 - \overline{\rho}, \ldots \rho_{M_1} - \overline{\rho}\right]^T \left[\rho_1 - \overline{\rho}, \ldots \rho_{M_1} - \overline{\rho}\right]$$

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{M_1} \qquad \{\mathbf{x}_i\} = \text{eigenvectors of } \mathrm{C}$$



2) Thresholding: find *M* s.t. For a given threshold γ:

$$\frac{\sum_{i=1}^{M} \lambda_i}{\sum_{i=1}^{M_1} \lambda_i} \geq \gamma$$

3) New basis

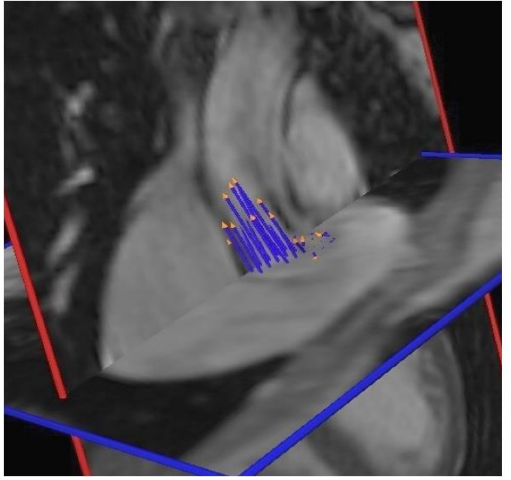
$$\mathbf{y}_i = \sum_{j=1}^{M_1} (\mathbf{x}_i)_j (\rho_j - \overline{\rho})$$

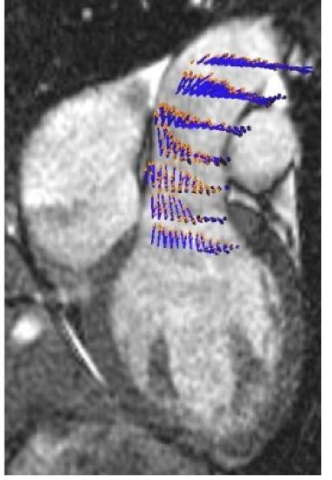
i=1,2,...M

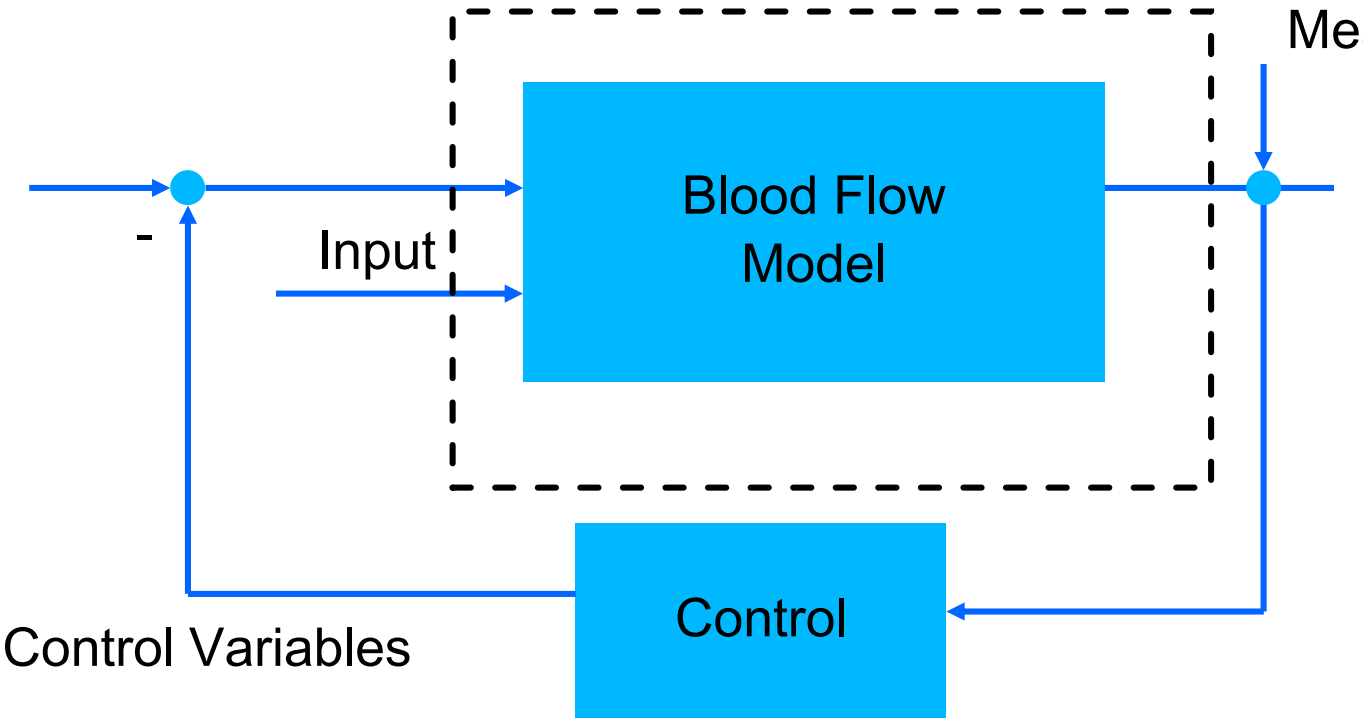
Normalized

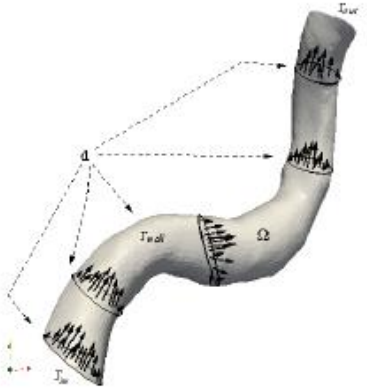
$$\mathbf{y}_i^* = \frac{1}{\|\mathbf{y}_i\|} \mathbf{y}_i$$

Variational assimilation of velocity measures in the domain of interest







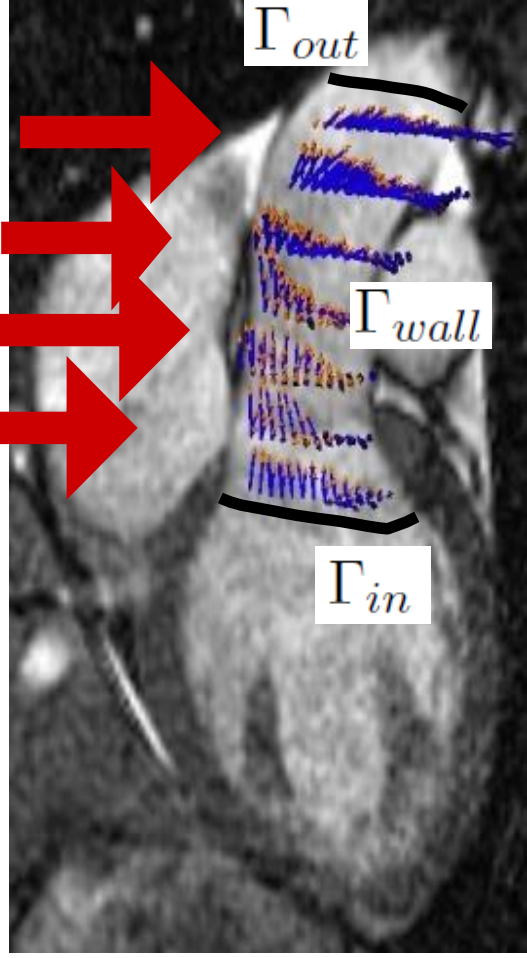


The deterministic mathematical problem

Constraint: the NS equations

Data **d**

$$
\text{(NS)} \quad \begin{cases}
-\nu\,\nabla\cdot(\nabla\mathbf{u}+\nabla\mathbf{u}^{\mathrm{T}})+(\mathbf{u}\cdot\nabla)\mathbf{u}+\nabla p=\mathbf{s} & \text{in } \Omega \\
\nabla\cdot\mathbf{u}=0 & \text{in } \Omega \\
\mathbf{u}=\mathbf{0} & \text{on } \Gamma_{wall}; \\
-\nu\,(\nabla\mathbf{u}+\nabla\mathbf{u}^{\mathrm{T}})\cdot\mathbf{n}+p\cdot\mathbf{n}=\mathbf{h} & \text{on } \Gamma_{in}; \\
-\nu\,(\nabla\mathbf{u}+\nabla\mathbf{u}^{\mathrm{T}})\cdot\mathbf{n}+p\cdot\mathbf{n}=\mathbf{g} & \text{on } \Gamma_{out}.
\end{cases}
$$



Control Variable (arbitrary choice): inlet flux **h**

Minimization problem:

find **h*** s.t.
dist (**u(h*),d**) ≤ *dist* (**u(h),d**)
under the constraint of (NS)

Approach: Discretize then Optimize

The discrete problem (Oseen)

Ingredients:

$\mathbf{V} = \begin{bmatrix} \mathbf{U} \\ \mathbf{P} \end{bmatrix}$ discretized $\mathbf{u}$ and p

$\mathrm{R}_{in} = [\mathrm{R}_{in,u} \;\; \mathrm{O}]$ inflow restriction matrix

$\mathrm{D} = [\mathrm{Q} \;\; \mathrm{O}]$ selection matrix of the measurements sites

M_{in} inflow mass matrix

$\mathrm{S} = \begin{bmatrix} \mathrm{C} & \mathrm{B}^{\mathrm{T}} \\ \mathrm{B} & \mathrm{O} \end{bmatrix}$ Oseen matrix

Regularizing term

$$\min_{\mathbf{H}} \mathcal{J}(\mathbf{V}, \mathbf{H}) = \frac{1}{2}\|\mathrm{D}\mathbf{V} - \mathbf{d}\|_2^2 + \frac{a}{2}\|\mathrm{L}\mathbf{H}\|_2^2$$
$$\text{s.t. } \mathrm{S}\mathbf{V} = \mathrm{R}_{in}^{\mathrm{T}}\mathrm{M}_{in}\mathbf{H} + \mathbf{F}.$$

KKT:

$$\mathrm{D}^{\mathrm{T}}(\mathrm{D}\mathbf{V} - \mathbf{d}) + \mathrm{S}^{\mathrm{T}}\mathbf{\Lambda} = \mathbf{0}$$
$$a\mathrm{L}^{\mathrm{T}}\mathrm{L}\mathbf{H} - \mathrm{M}_{in}^{\mathrm{T}}\mathrm{R}_{in}\mathbf{\Lambda} = \mathbf{0};$$
$$\mathrm{S}\mathbf{V} - \mathrm{R}_{in}^{\mathrm{T}}\mathrm{M}_{in}\mathbf{H} - \mathbf{F} = \mathbf{0};$$

Reduced Hessian

$$(\mathrm{Z}^{\mathrm{T}}\mathrm{Z} + a\mathrm{L}^{\mathrm{T}}\mathrm{L})\mathbf{H} = \mathrm{Z}^{\mathrm{T}}(\mathbf{d} - \mathrm{D}\mathrm{S}^{-1}\mathbf{F})$$

where $\mathrm{Z} = \mathrm{D}\mathrm{S}^{-1}\mathrm{R}_{in}^{\mathrm{T}}\mathrm{M}_{in}$

(sensitivity matrix)

Non-singularity of the discrete problem

Sufficient condition for the non-singularity is that
the reduced Hessian $Z^T Z + aL^T L$ is p.d.

1. This is always the case for $a>0$
2. For $a=0$, a sufficient condition is

$$Null(D) \cap Range(S^{-1} R_{in}^T M_{in}) = \{0\}$$

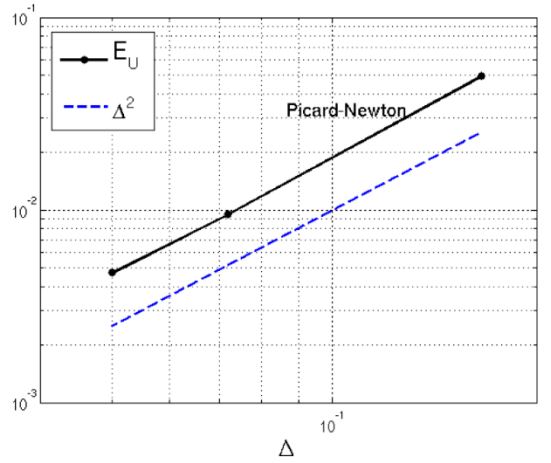
As expected, this is forcing a **condition on the measurements sites** (represented by the **matrix Q in D**), forcing that *a significant part of them is on the inflow* .

M. D'Elia, M. Perego, A. V., J Sci Comp

The Nonlinear Case

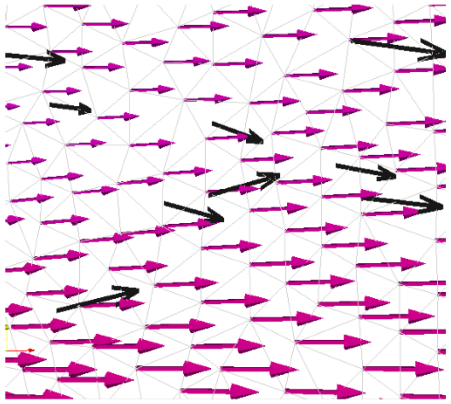
noise-free

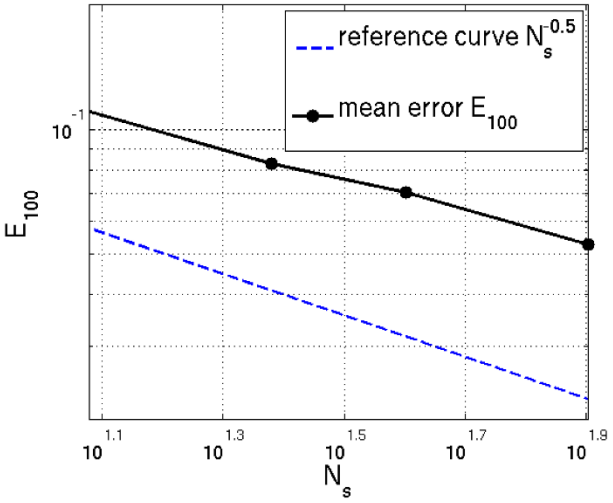
consistency: finite element convergence rate



noisy

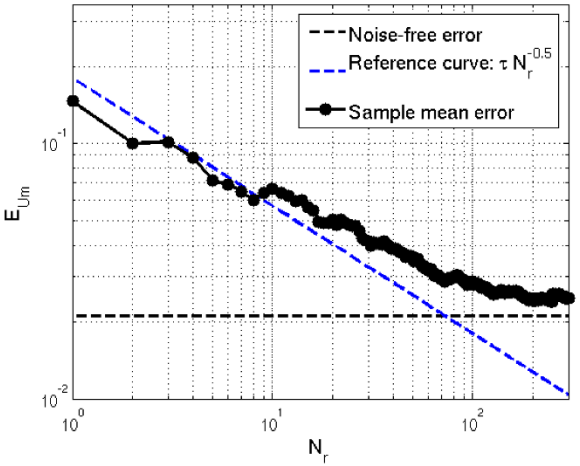
effective and robust noise filtering method:





- $E_{\mathbf{U}}$ **decreases** as more data are available: $E_{\mathbf{U}}(N_s) = \mathcal{O}(N_s^{-0.5})$

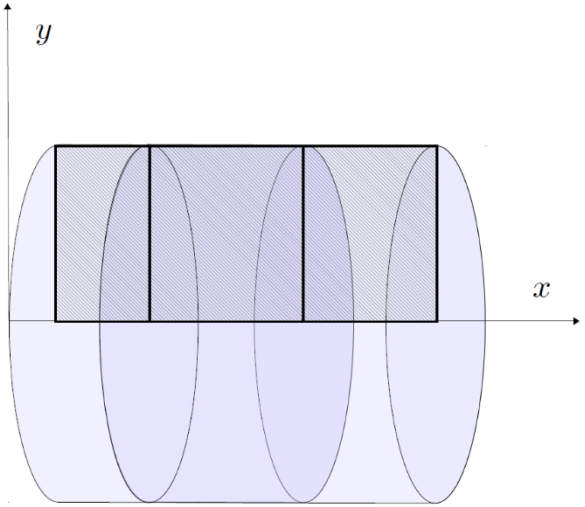
- the sample mean of the velocity over N_r noise realizations **converges** to the FE noise-free solution with rate $\mathcal{O}(N_r^{-0.5})$

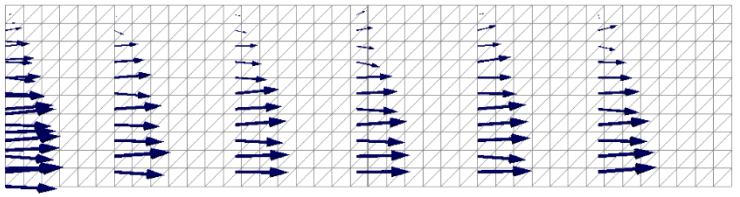


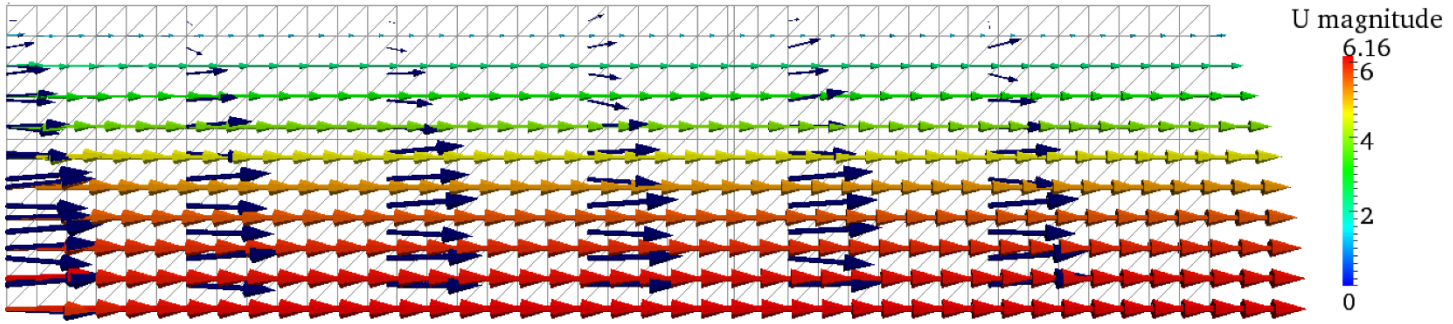
Axi-symmetric 3D (Poiseuille)

domain: rectangular domain representing a slice of a cylinder

data generation: analytical solution with additional noise, data on the inflow boundary do not satisfy sufficient conditions ⟶ piece-wise linear interpolation



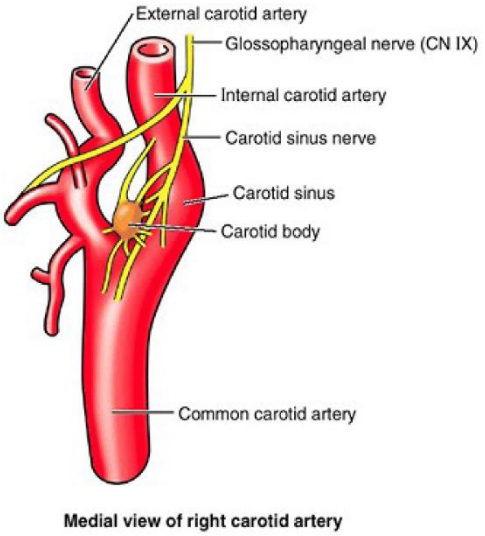


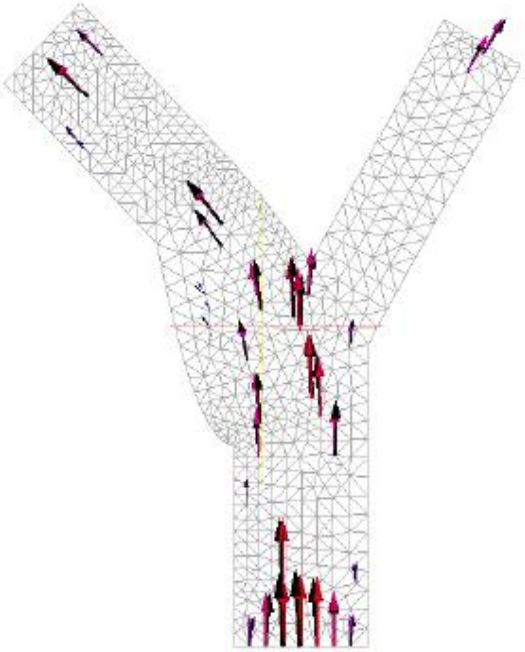


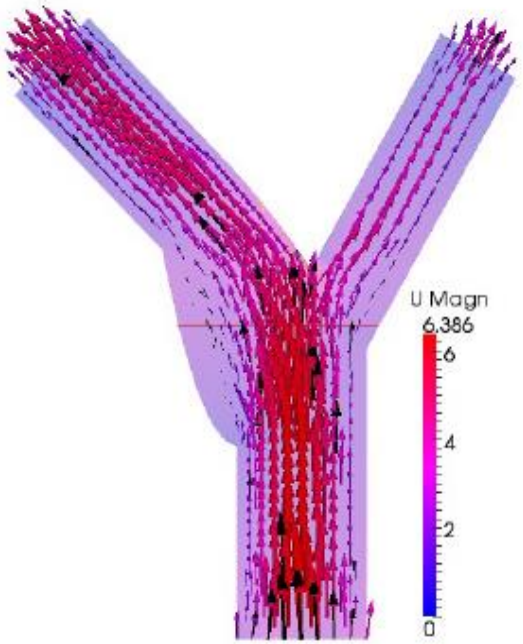
Non-trivial 2D

domain: 2D approximation of a carotid with a bulb

data generation: finite element approximation on a very fine grid with additional noise, data on the inflow boundary do not satisfy sufficient conditions → interpolation







wall shear stress (WSS) comparison

SNR	$E^*_{\mathrm{WSS,DA}}$	$E^*_{\mathrm{WSS,FW}}$
100	0.2536	0.2667
20	0.2591	0.3030
10	0.2738	0.3861
5	0.3149	0.6114

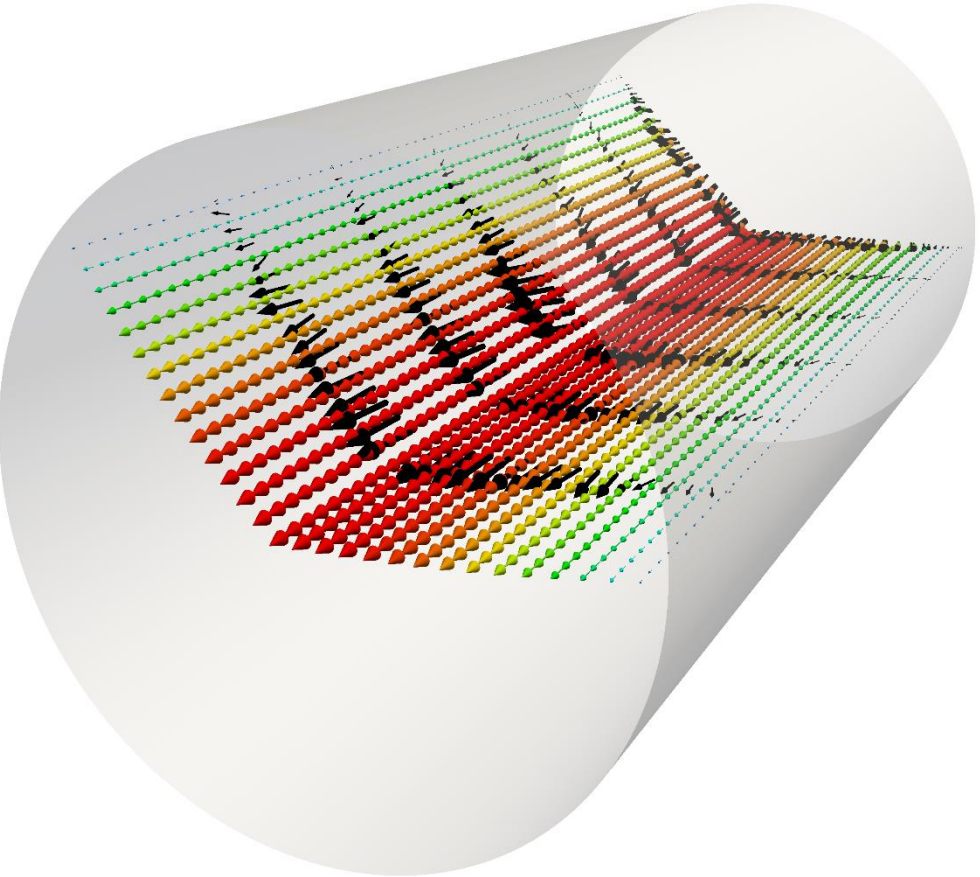
Axi-symmetric 3D (Womersley) – Unsteady!!!

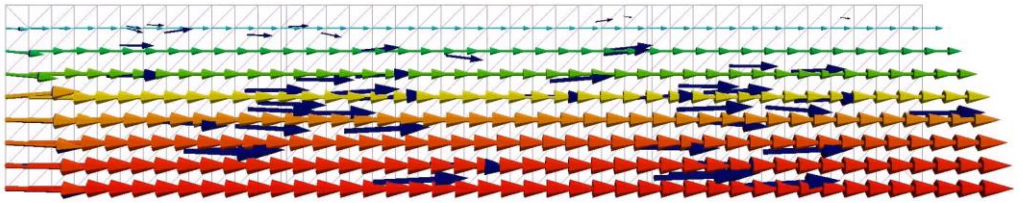
Challenges of the unsteady case:

1) The initial conditions are not known, they should be part of the identification problem. More precisely, periodicity should be enforced.

2) Computational costs) Model Reduction Techniques

Here, we follow a DO approach, so we optimize at each time step.





Bayesian formulation

Goal: estimate the reliability of results → quantification of the uncertainty

We predict stochastic features of the variables of interest

the prediction of the uncertainty is based on the **knowledge** of

- the measurement process
- deterministic models available

Main assumptions:

- all discretized variables are treated as **random**

  information resides in the **probability distributions**

- the entities involved are probability density functions (PDFs)
- the method delivers a **distribution**

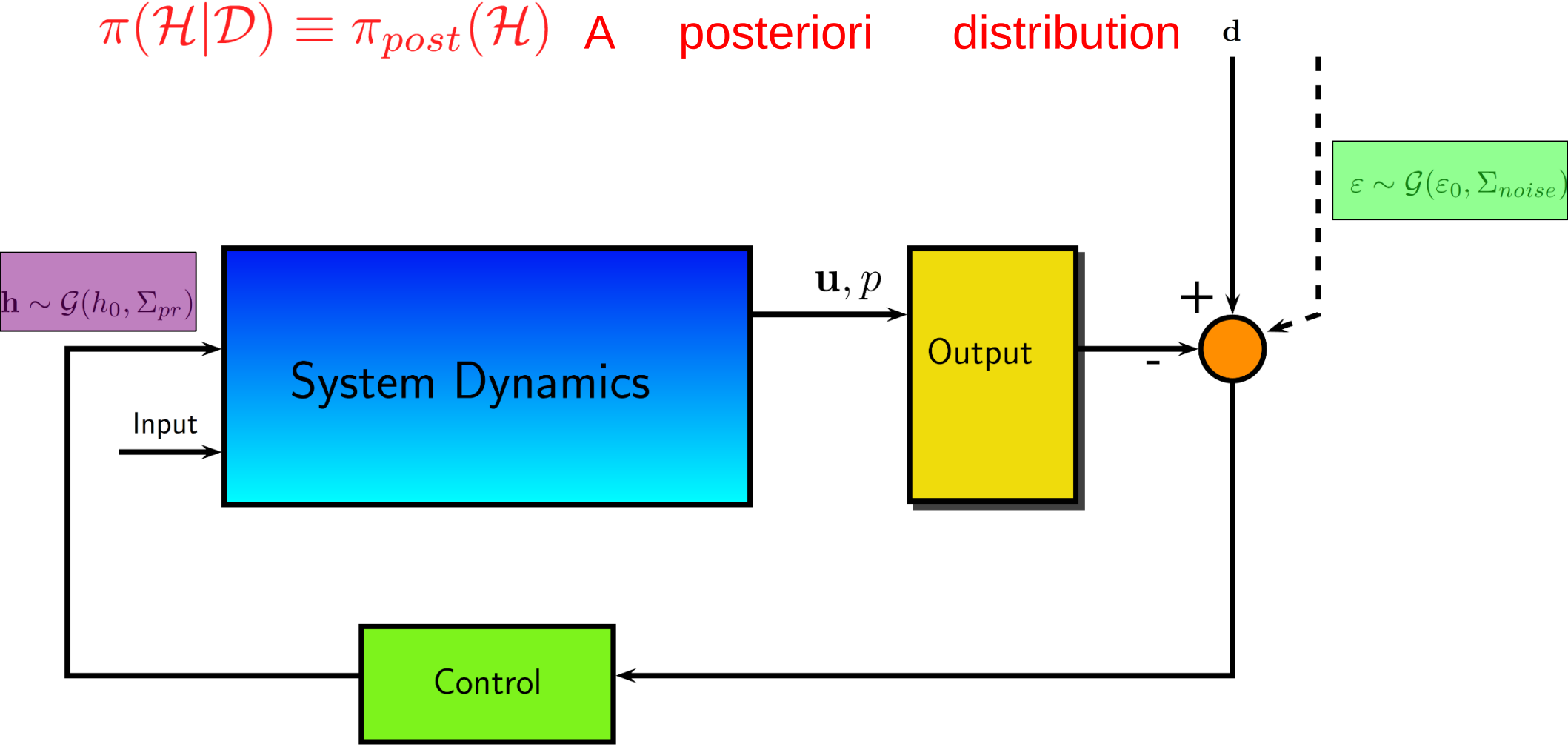
(while deterministic methods produce a **single** estimate)

Notation

- $\mathcal{H}$ is the RV counterpart for $\mathbf{h}$ with pdf (prior) $\pi_{pr}(\mathcal{H})$
- $\mathcal{D}$ is the RV counterpart of the measurements $\mathbf{d}$
- $\mathcal{E}$ is the RV associated with the noise with pdf $\pi_{noise}(\mathcal{E})$

$\pi(\mathcal{D}|\mathcal{H})$ Likelihood

$\pi(\mathcal{H}|\mathcal{D}) \equiv \pi_{post}(\mathcal{H})$ A posteriori distribution



We assume an additive noise model $\mathcal{Z}\mathcal{H} + \mathcal{E} = \mathcal{D}$

with $\mathcal{H}$ and $\mathcal{E}$ mutually independent

Here, $\mathcal{Z}$ is the deterministic model (sensitivity matrix = Neumann-to-Dirichlet map)

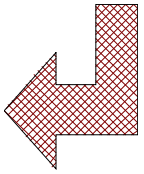
$$\Rightarrow \pi(\mathcal{D}|\mathcal{H}) = \pi_{noise}(\mathcal{D} - \mathcal{Z}\mathcal{H})$$

Thanks to the Bayes Theorem

$$\pi_{post}(\mathcal{H}) = \frac{\pi(\mathcal{D}|\mathcal{H})\pi_{pr}(\mathcal{H})}{\pi(\mathcal{D})} \sim \pi_{noise}(\mathcal{D} - \mathcal{Z}\mathcal{H})\pi_{prior}(\mathcal{H})$$

Assume a Gaussian distribution for prior and noise

$$\pi_{pr} \sim \mathcal{G}(h_0, \Sigma_{pr}) \qquad \pi_{noise} \sim \mathcal{G}(\varepsilon_0, \Sigma_{noise})$$



Posterior is $\pi_{post} \sim \mathcal{G}(h_{post}, \Sigma_{post})$

$$h_{post} = (\Sigma_{pr}^{-1} + Z^T\Sigma_{noise}^{-1}Z)^{-1}(Z^T\Sigma_{noise}^{-1}(d - \varepsilon_0) + \Sigma_{pr}^{-1}h_0)$$

$$\Sigma_{post} = (\Sigma_{pr}^{-1} + Z^T\Sigma_{noise}^{-1}Z)^{-1}$$

Two approaches

Maximum A Posteriori $h = arg \max_h \pi_{post}$

Maximum Likelihood $h = arg \max_h \pi(\mathcal{D}|\mathcal{H})$

$$h_{det} = (a\mathrm{L}^T\mathrm{L} + \mathrm{Z}^T\mathrm{Z})^{-1}\mathrm{Z}^T(\mathbf{d} - \mathrm{DS}^{-1}\mathbf{F})$$

$$h_{MAP} = (\Sigma_{pr}^{-1} + \mathrm{Z}^T\Sigma_{noise}^{-1}\mathrm{Z})^{-1}\mathrm{Z}^T\Sigma_{noise}^{-1}(\mathbf{d} - \mathrm{DS}^{-1}\mathbf{F})$$

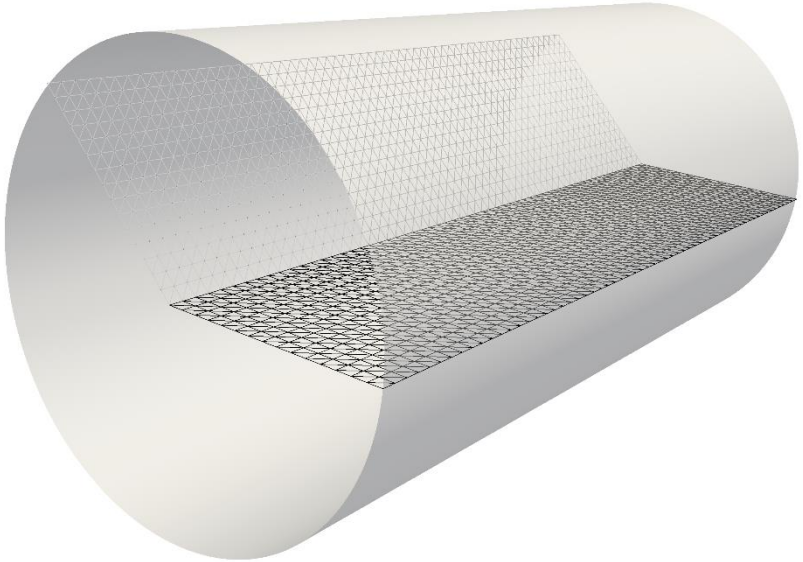
$$h_{ML} = (\mathrm{Z}^T\Sigma_{noise}^{-1}\mathrm{Z})^{-1}\mathrm{Z}^T\Sigma_{noise}^{-1}(\mathbf{d} - \mathrm{DS}^{-1}\mathbf{F})$$

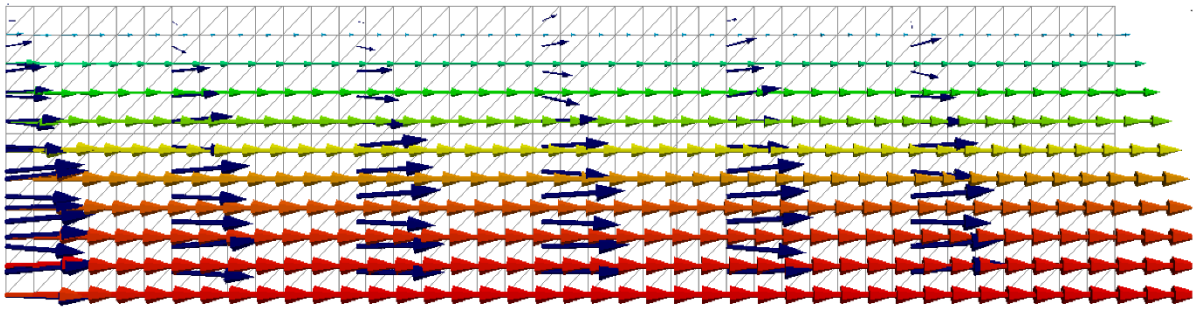
The a priori information acts as a Tychonov regularizer in MAP vs det

ML vs MAP: ML represents the limit case when no a priori knowledge is available on $\mathbf{h}$

As for h_{det} with a = 0, ML estimator requires conditions on Z

numerical results





test case E

- axisymmetric formulation
- h_{det} versus h_{MAP}
- $a = 10^{-7}$
- interpolation active

SNR	$\overline{E}_{\mathbf{U},det}$	$\overline{E}_{\mathbf{U},MAP}$	% gain
20	0.0396	0.0308	22%
10	0.1423	0.0978	31%

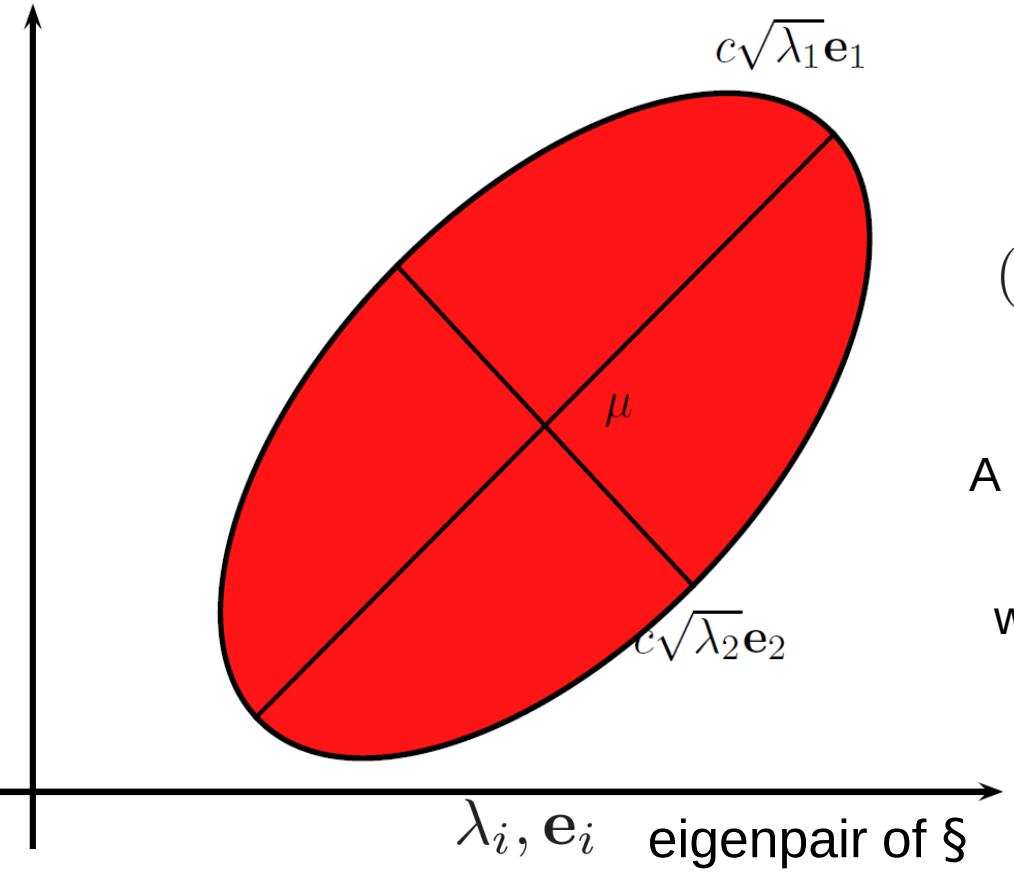
Statistical Spread Estimators

Multivariate normal distribution

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} exp(-(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu))$$

$$\forall x_i \in (-\infty, \infty), i = 1, \ldots, n$$

$$(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu) = c^2$$



$\lambda_i, \mathbf{e}_i$ eigenpair of §

FACTS

$(\mathbf{x}-\mu)\Sigma^{-1}(\mathbf{x}-\mu)$ has a χ^2_n distribution

A point belongs to the ellipsoid with $c^2 = \chi^2_n(\alpha)$

with probability $1-\alpha$

Spread estimators - velocity

Goal: quantify how likely velocity and flow related variables are inside an interval of (critical, significant) values

$$\mathbf{V} = [\mathbf{U}, P], \mathbf{U} = \mathrm{E}\mathbf{V}$$

$$\mathbf{U} = \mathrm{E}(\mathrm{S}^{-1}\mathrm{R}_{in}^{T}\mathrm{M}_{in}\mathbf{H} + \mathrm{S}^{-1}\mathbf{F})$$

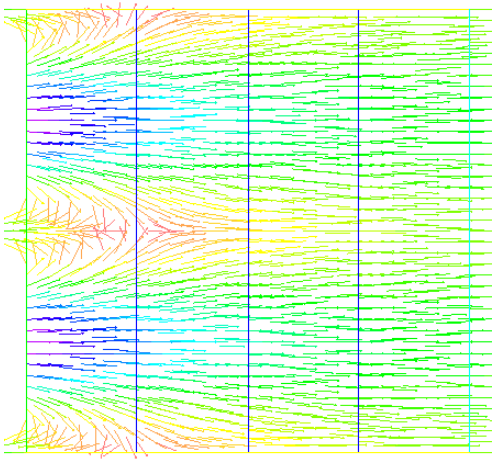
$$\mathrm{T} \equiv \mathrm{E}\mathrm{S}^{-1}\mathrm{R}_{in}^{T}\mathrm{M}_{in}$$

$$\Rightarrow \mathbf{U} \sim \mathcal{G}(\mathrm{T}h_{post} + \mathrm{T}\mathbf{F}, \mathrm{T}\Sigma_{post}\mathrm{T}^{T})$$

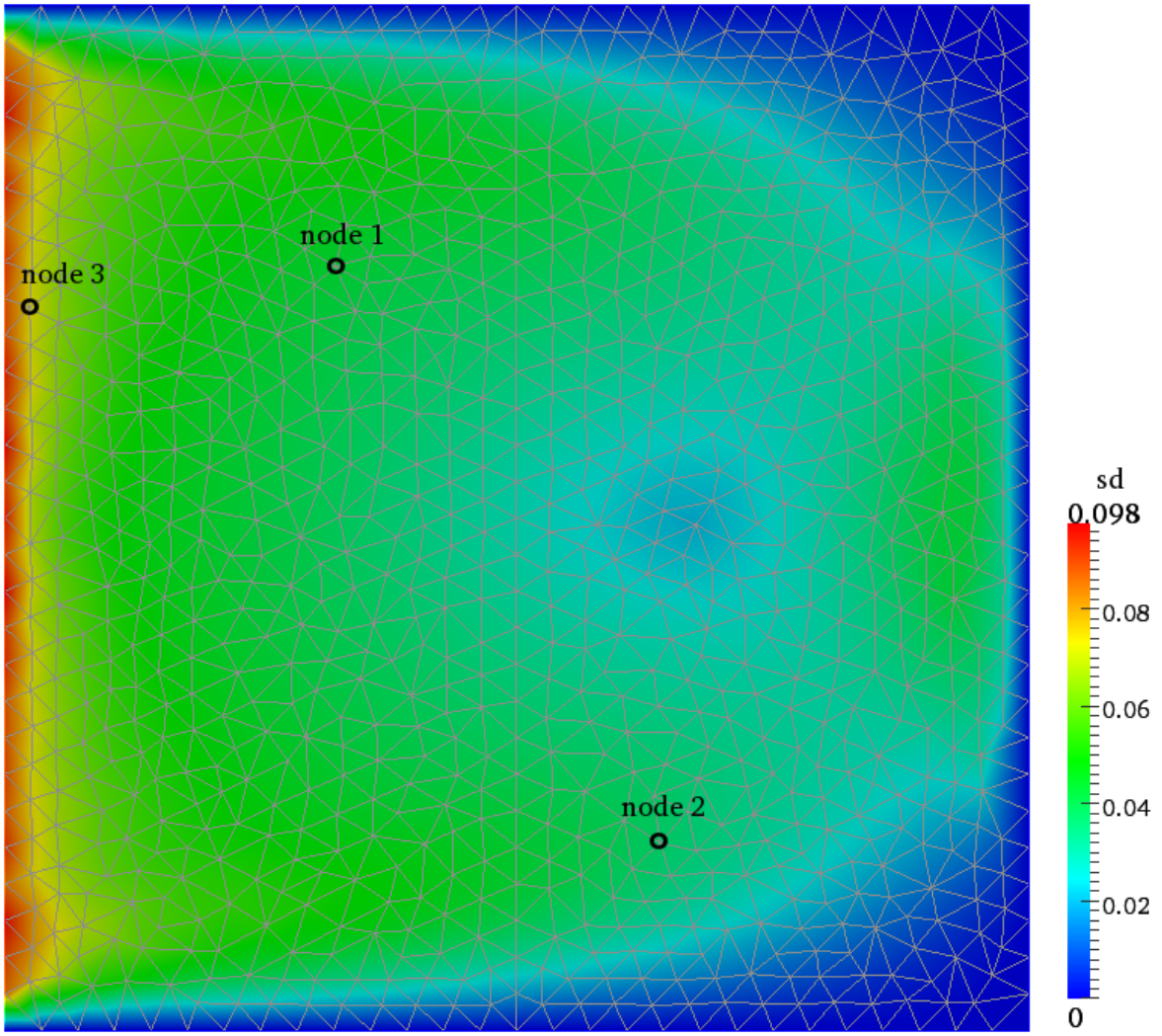
We build credibility regions for the velocity

spread estimators – numerical results

test case: same analytic solution, square domain, SNR = 20



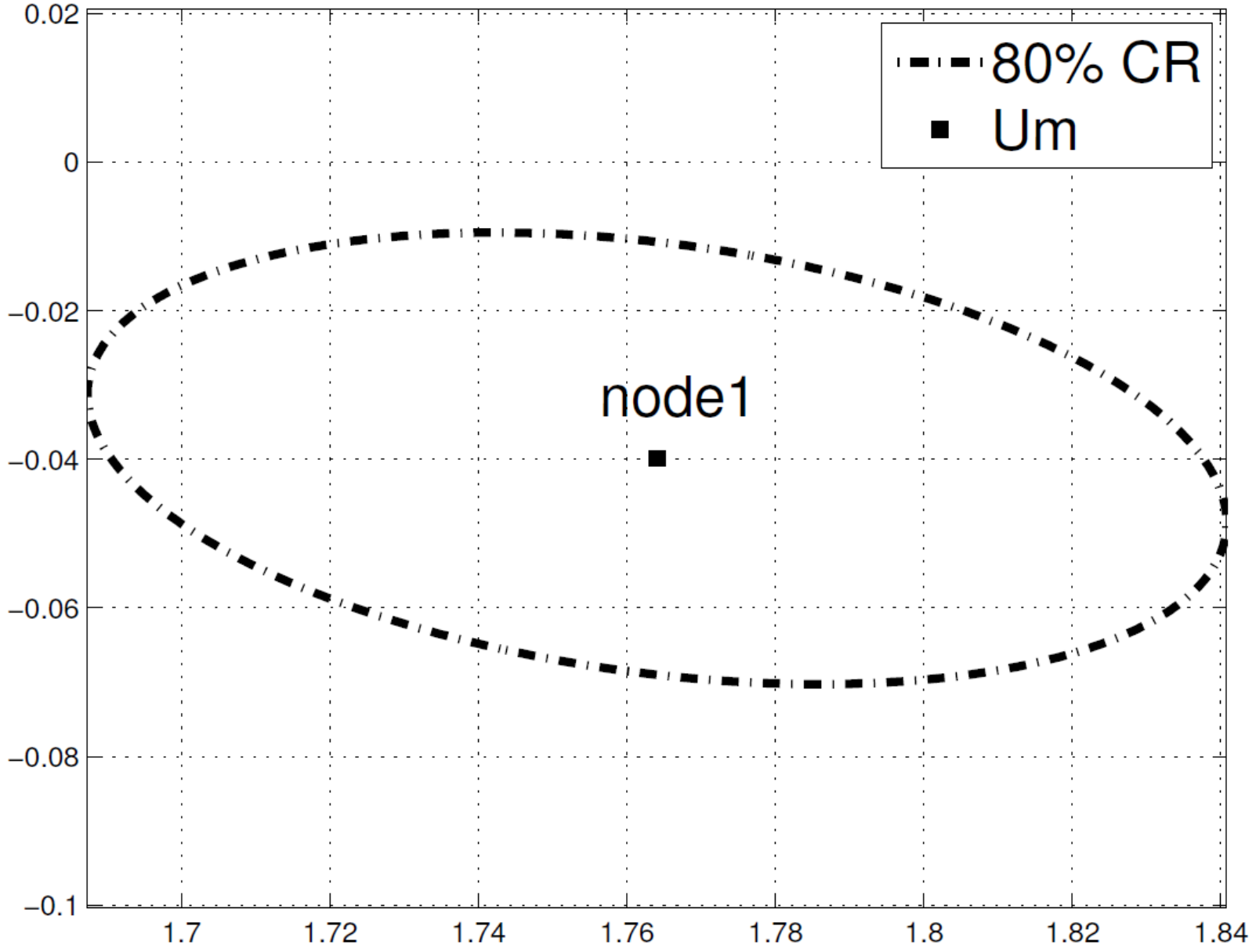
Map of $\sqrt{\lambda_{max}}$



input noise: std = 0.1467

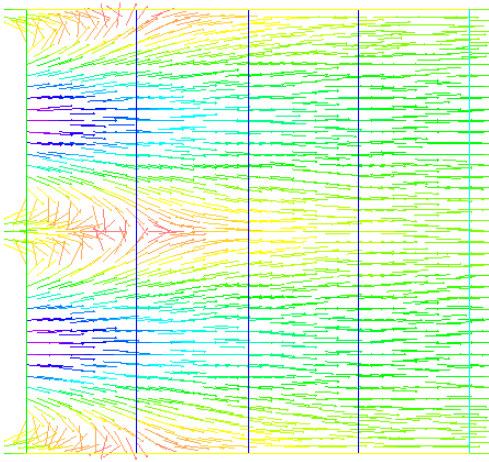
output noise: max std = 0.098

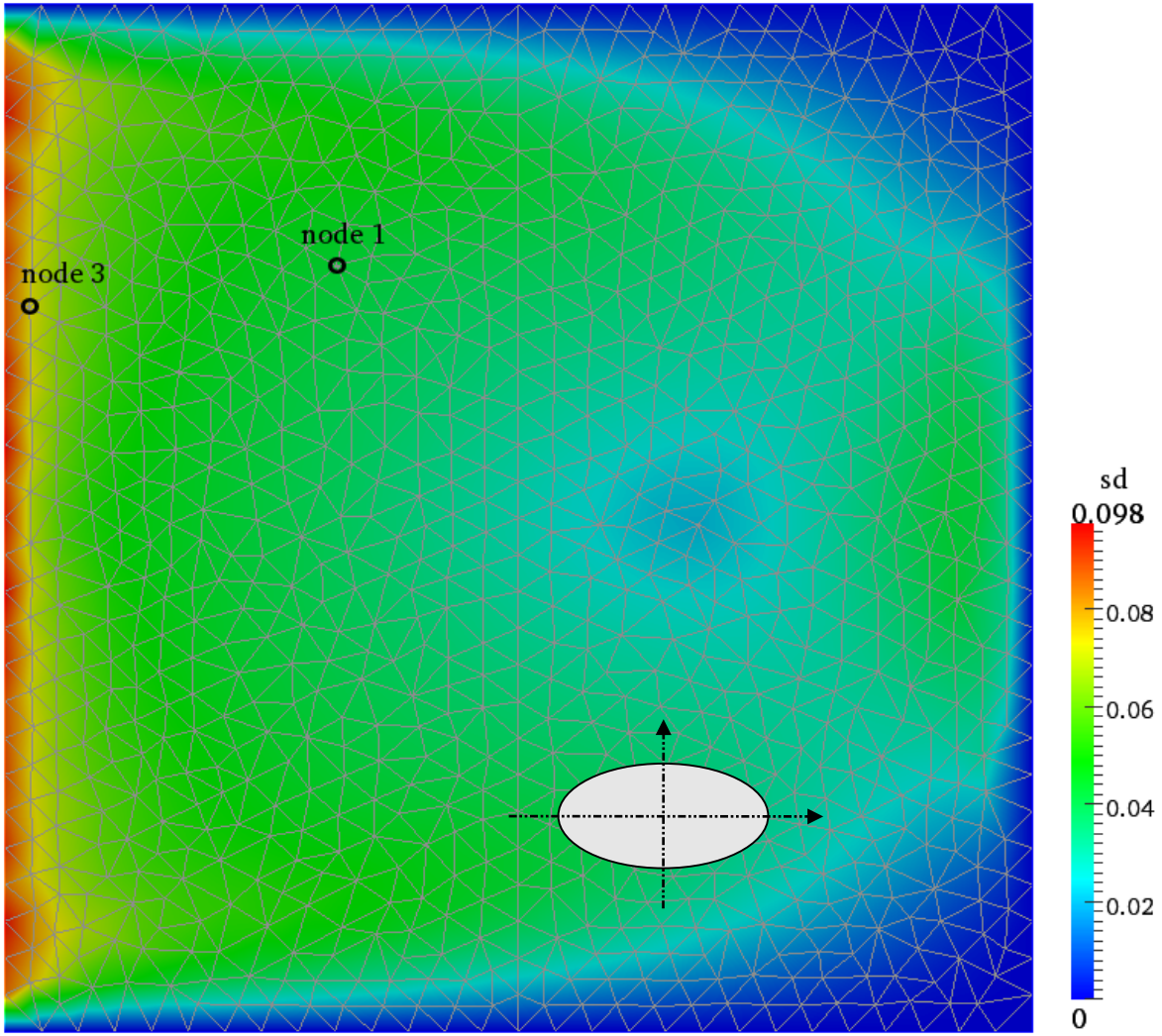
spread estimators – numerical results



spread estimators – numerical results

test case: same analytic solution, square domain, SNR = 20

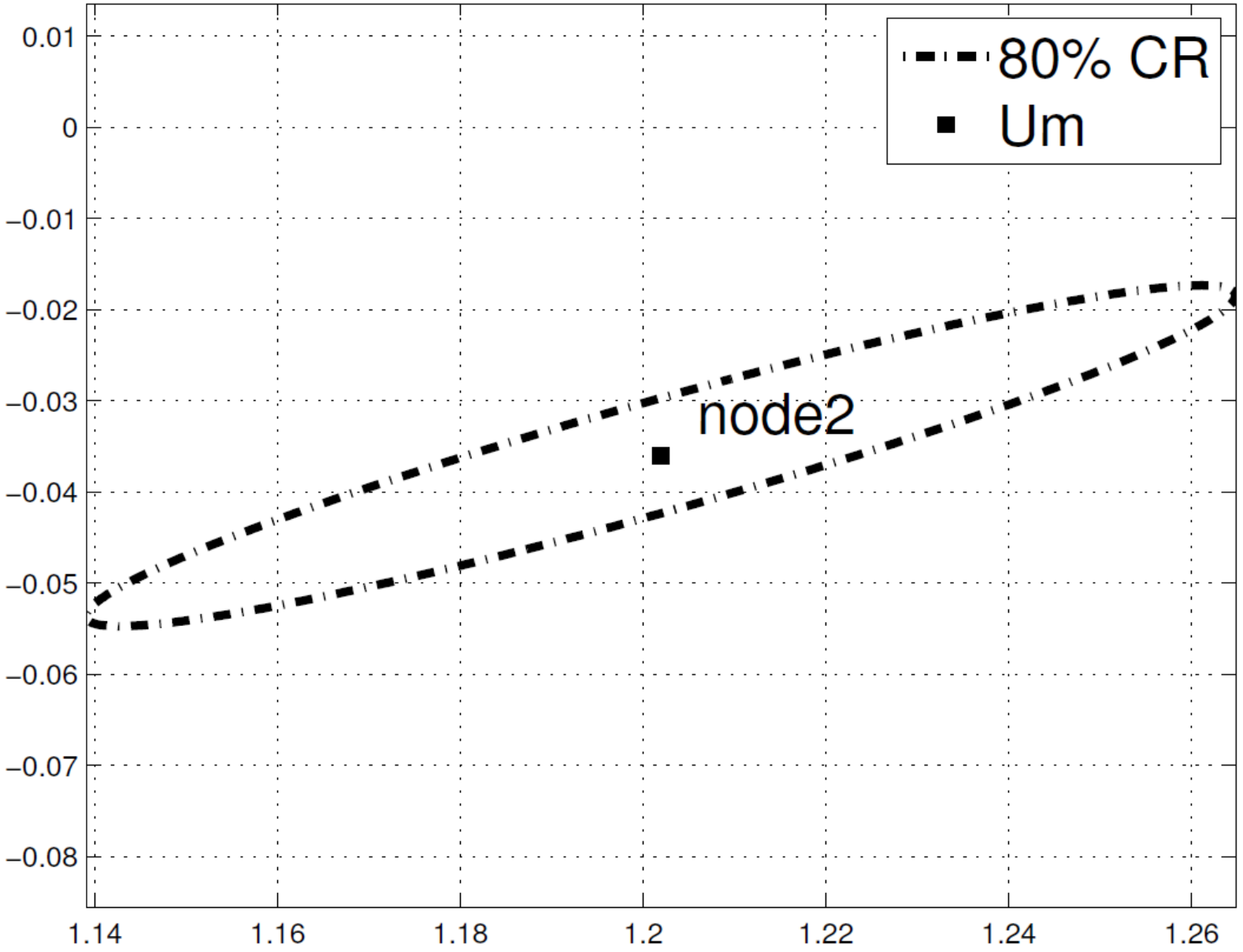




input noise: std = 0.1467

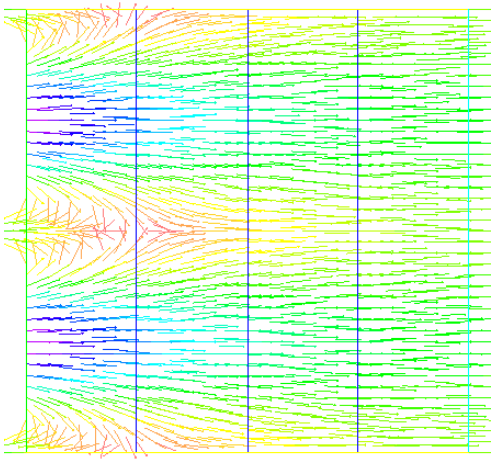
output noise: max std = 0.098

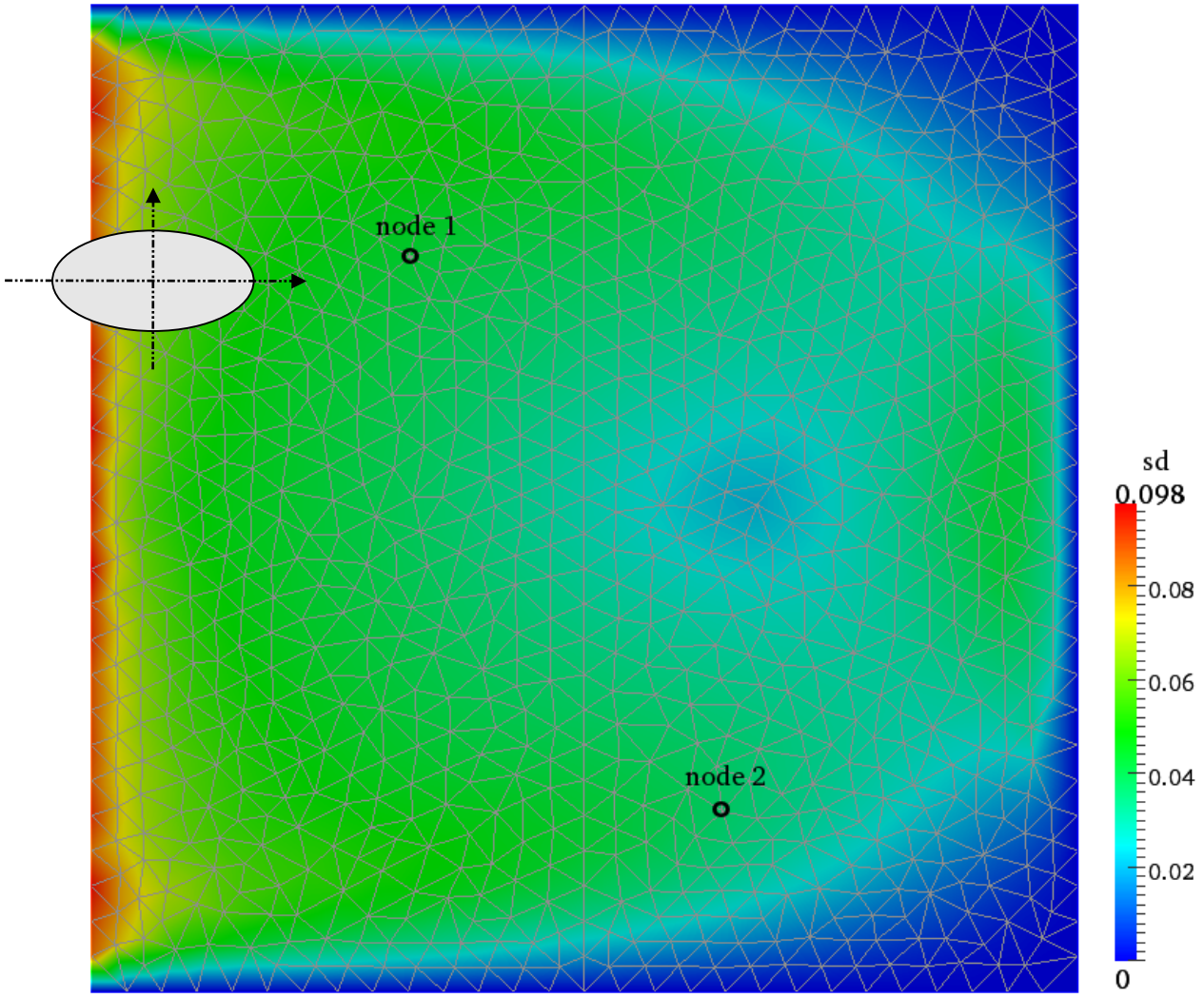
spread estimators – numerical results



spread estimators – numerical results

test case: same analytic solution, square domain, SNR = 20

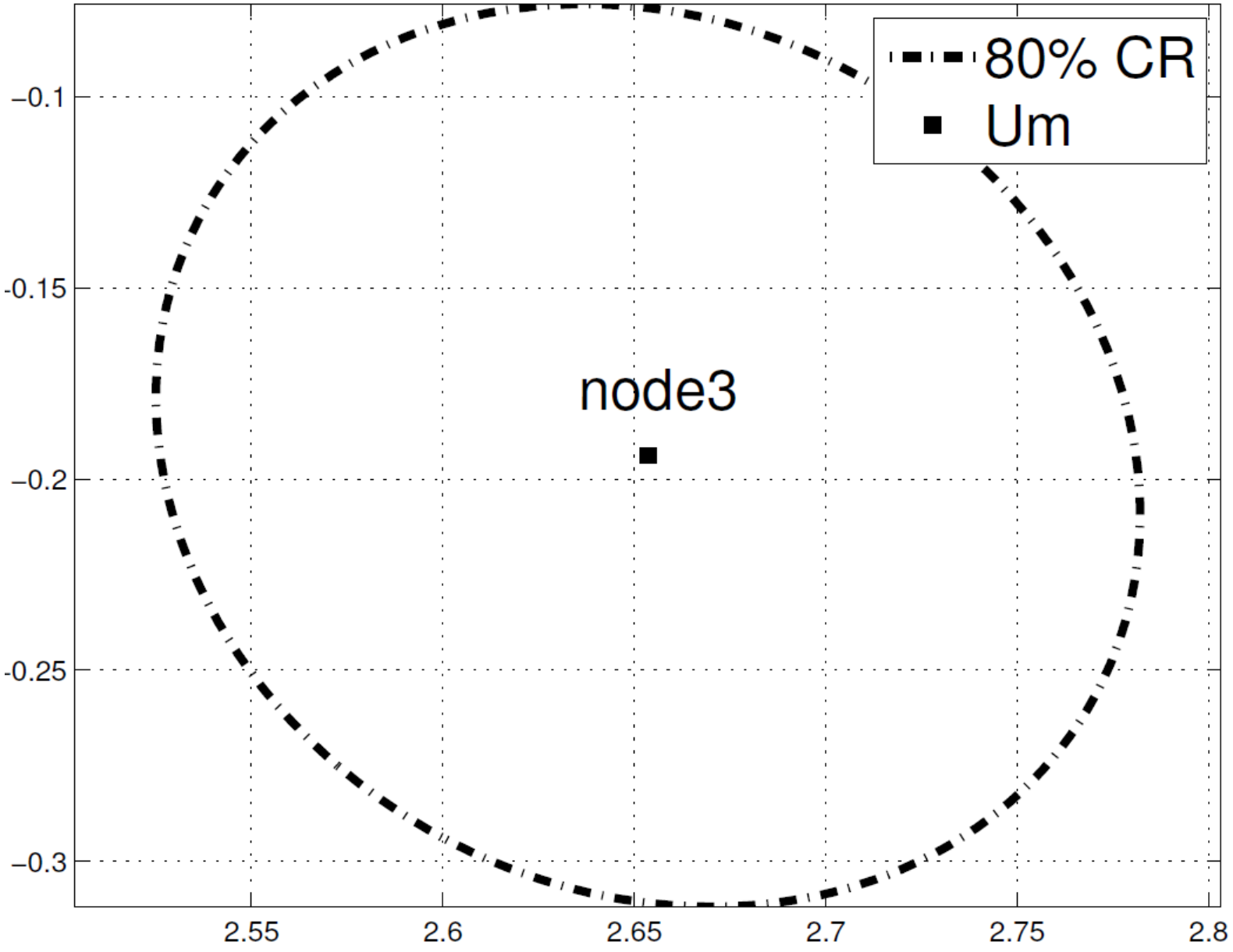




input noise: std = 0.1467

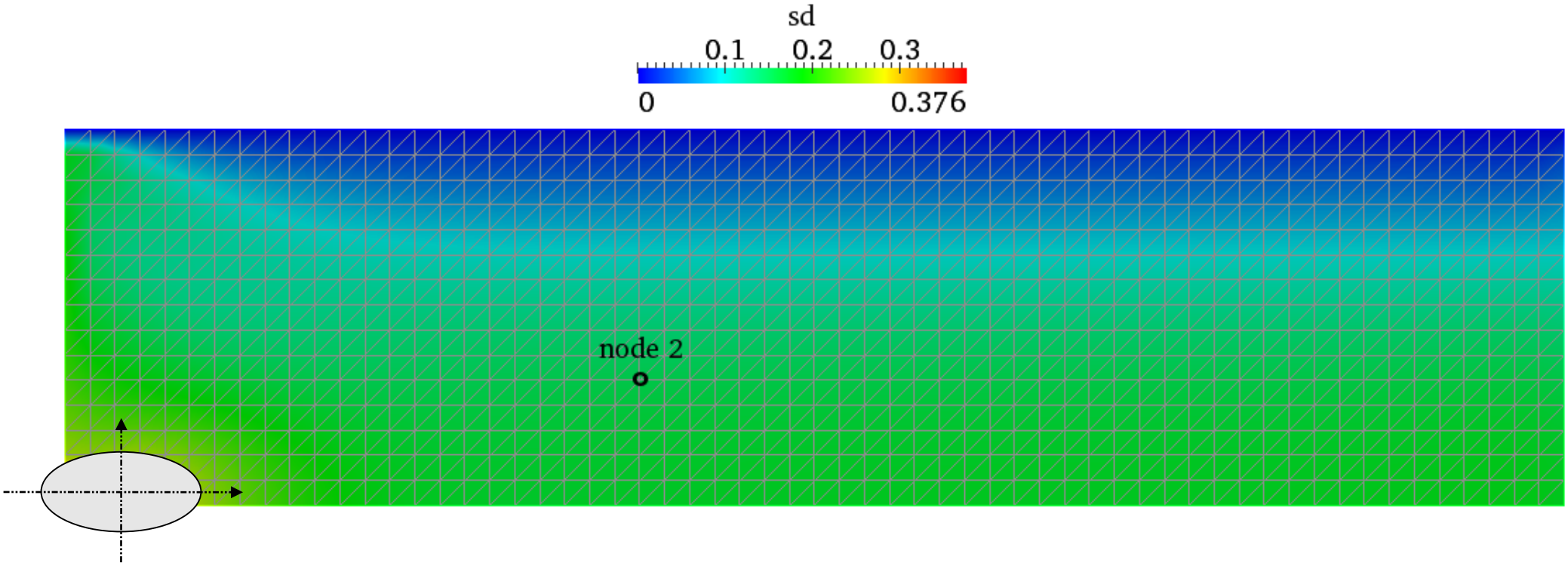
output noise: max std = 0.098

spread estimators – numerical results



spread estimators – numerical results

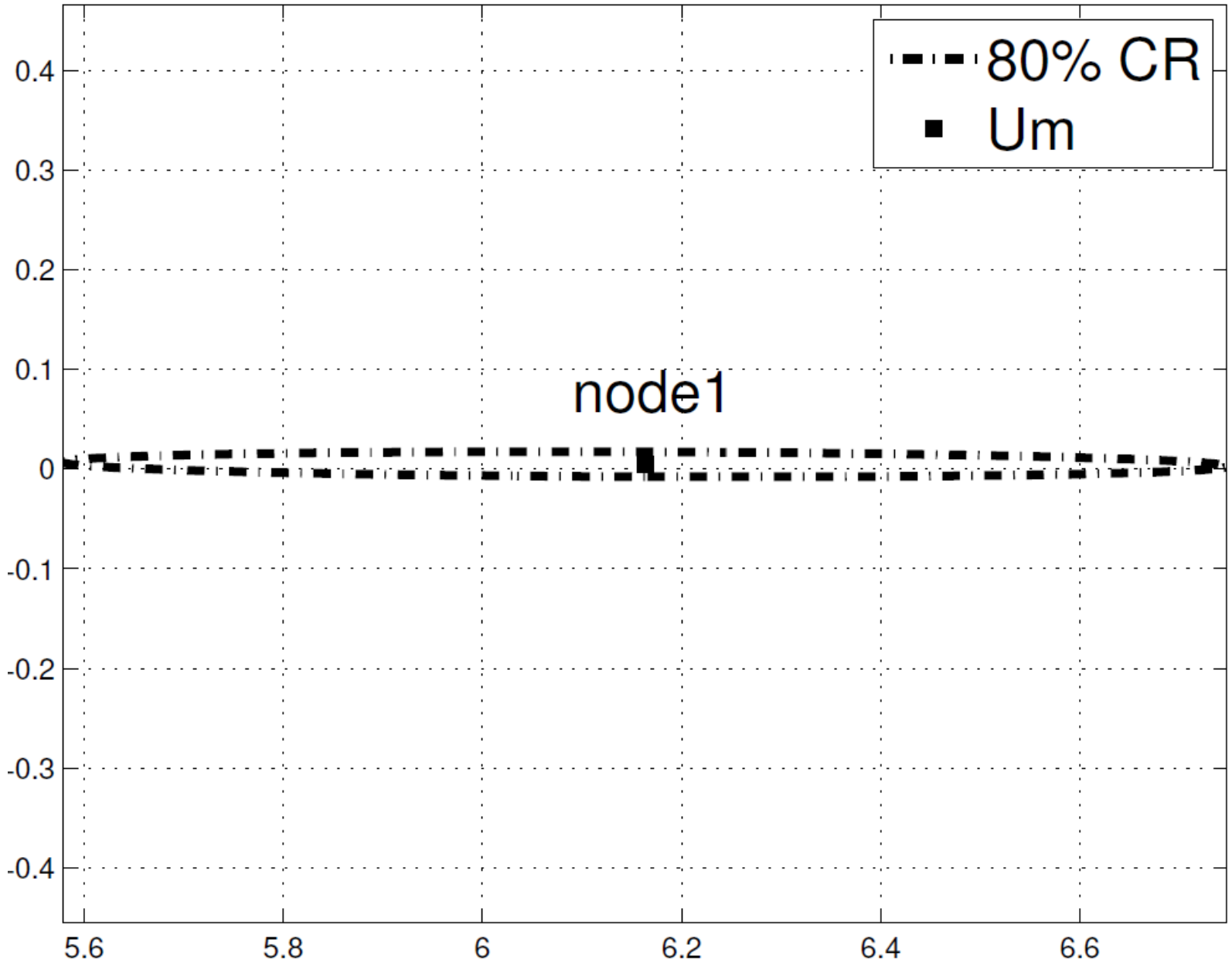
test case: axisymmetric case, cylindrical square domain, SNR = 20



input noise: std = 0.325, all over the domain

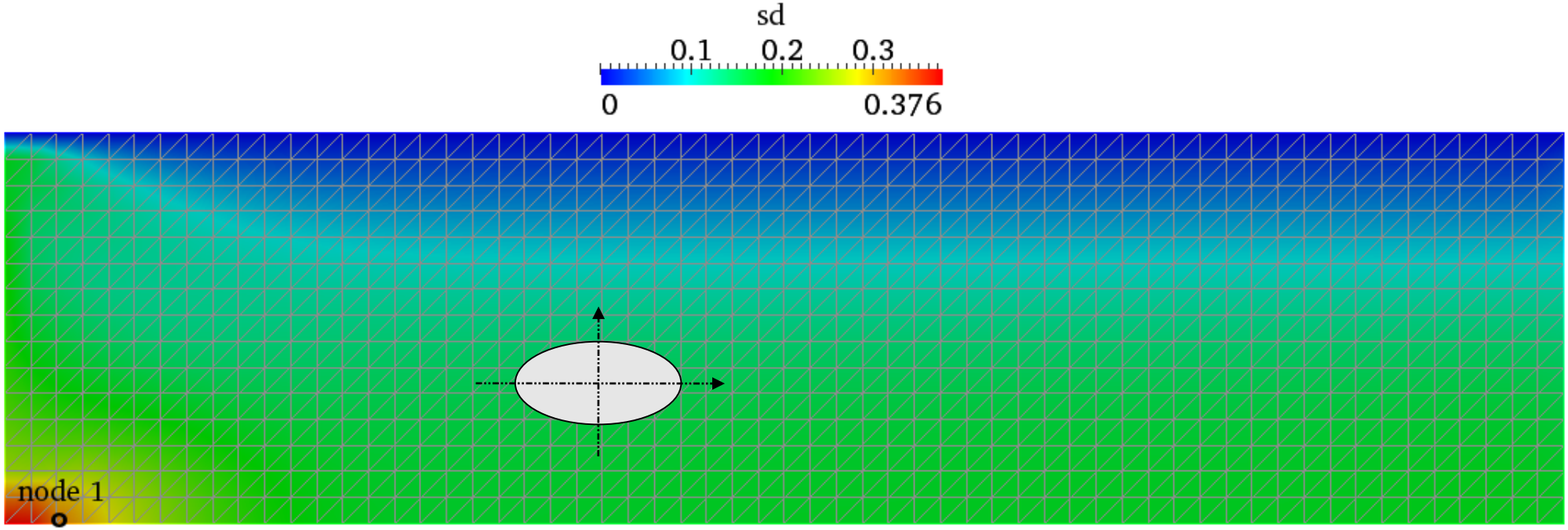
output noise: max std = 0.376, in a restricted area

spread estimators – numerical results



spread estimators – numerical results

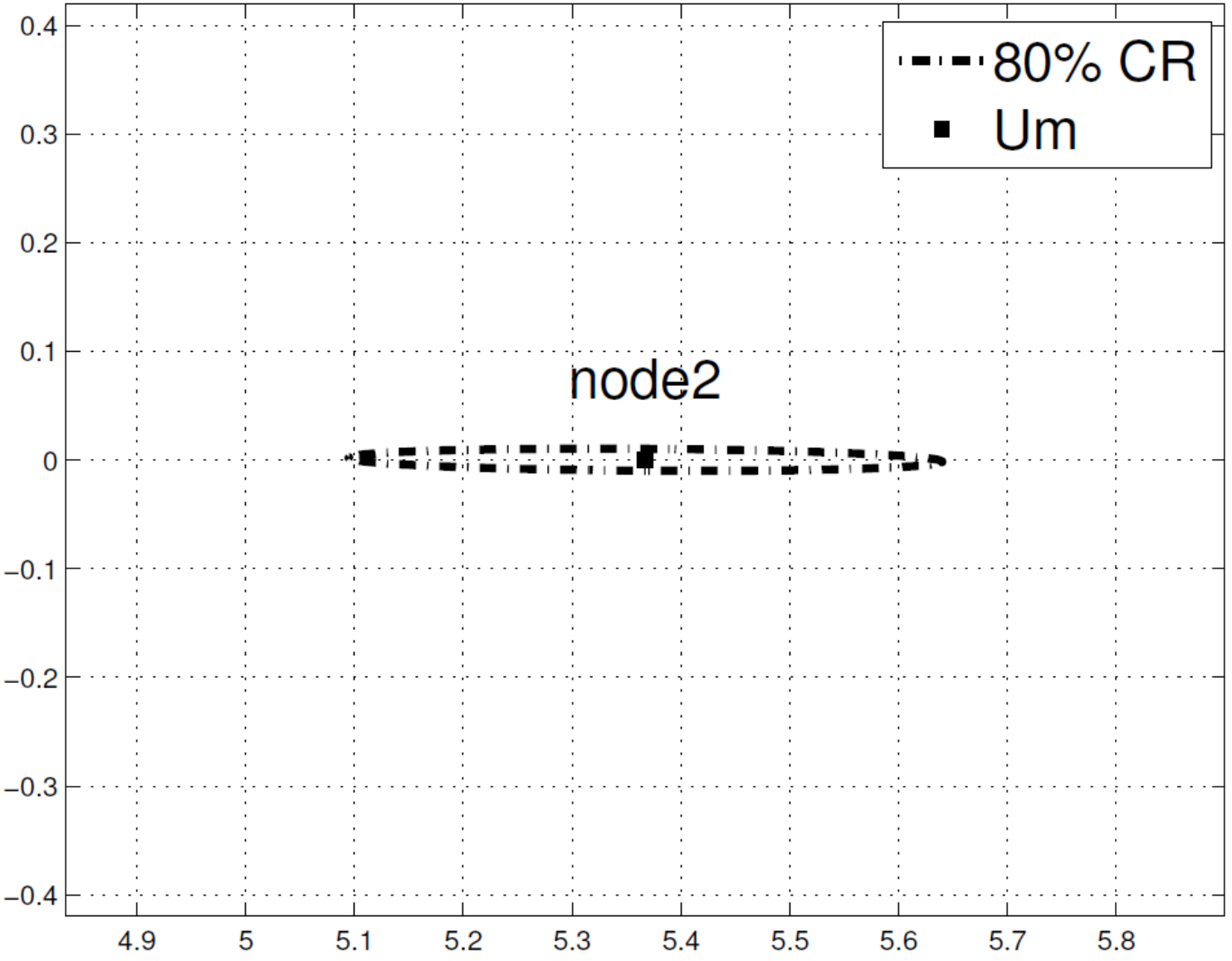
test case: axisymmetric case, cylindrical square domain, SNR = 20



input noise: std = 0.325, all over the domain

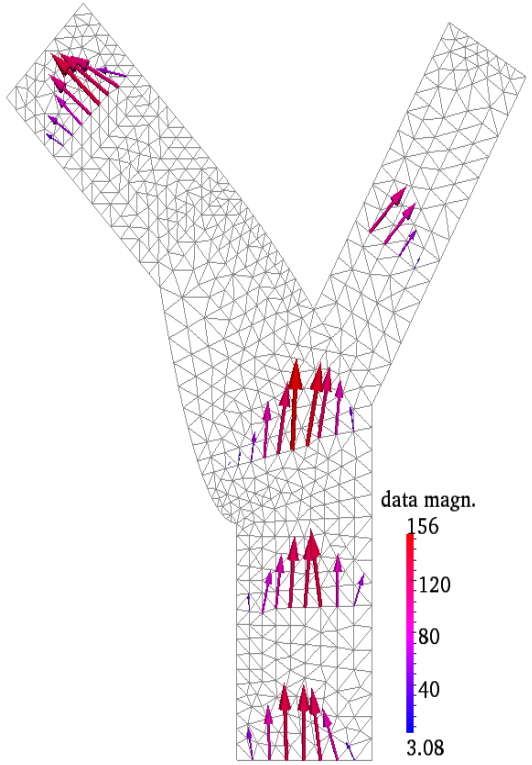
output noise: max std = 0.376, in a restricted area

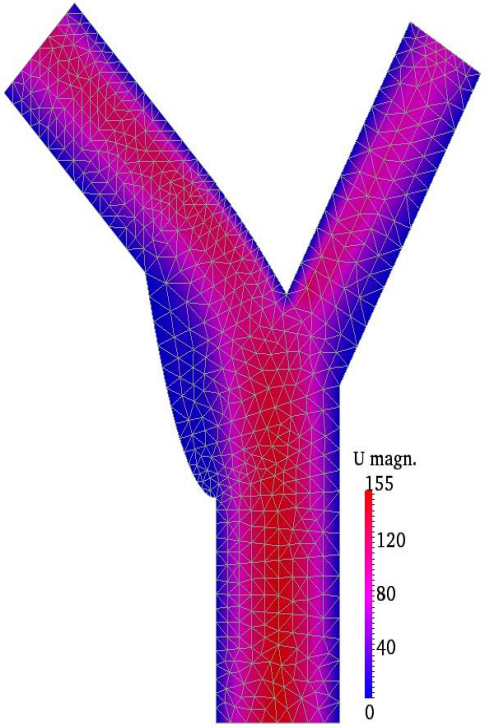
spread estimators – numerical results

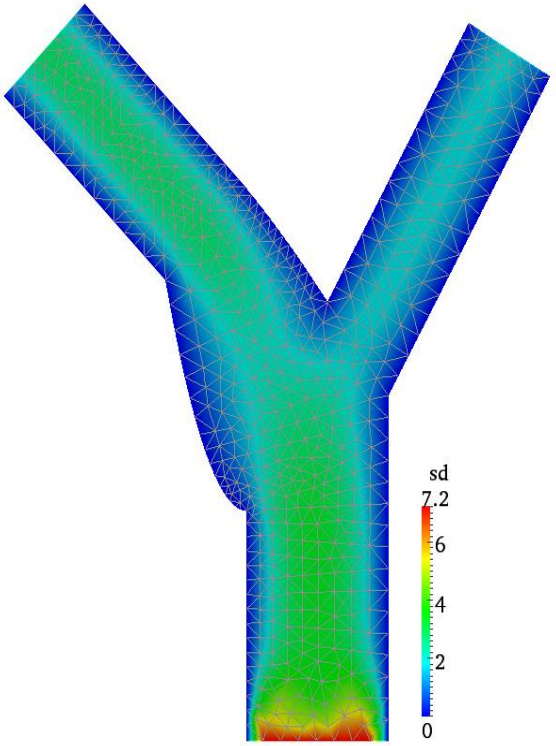


towards real geometries - carotid

Gain ML vs Deterministic estimator: 31%

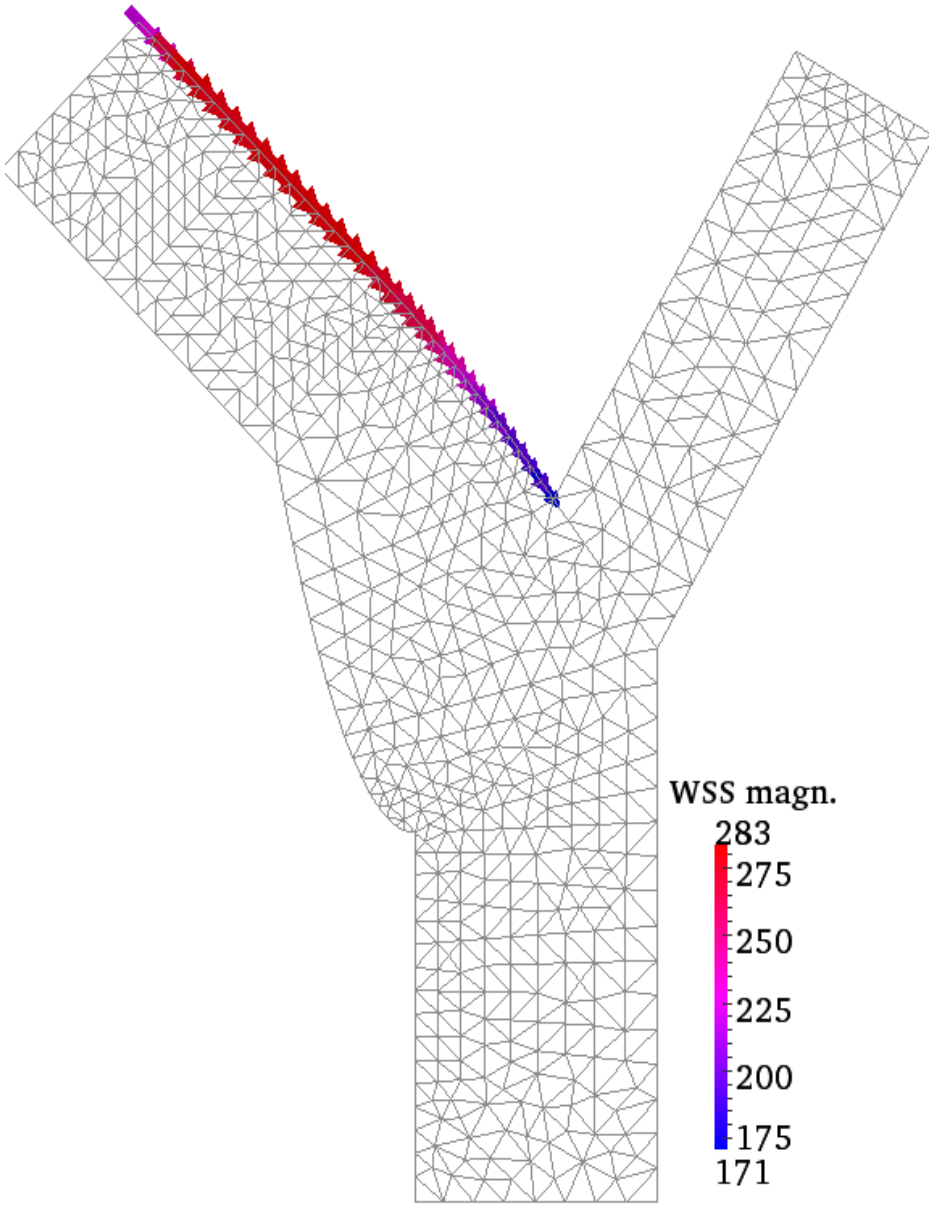


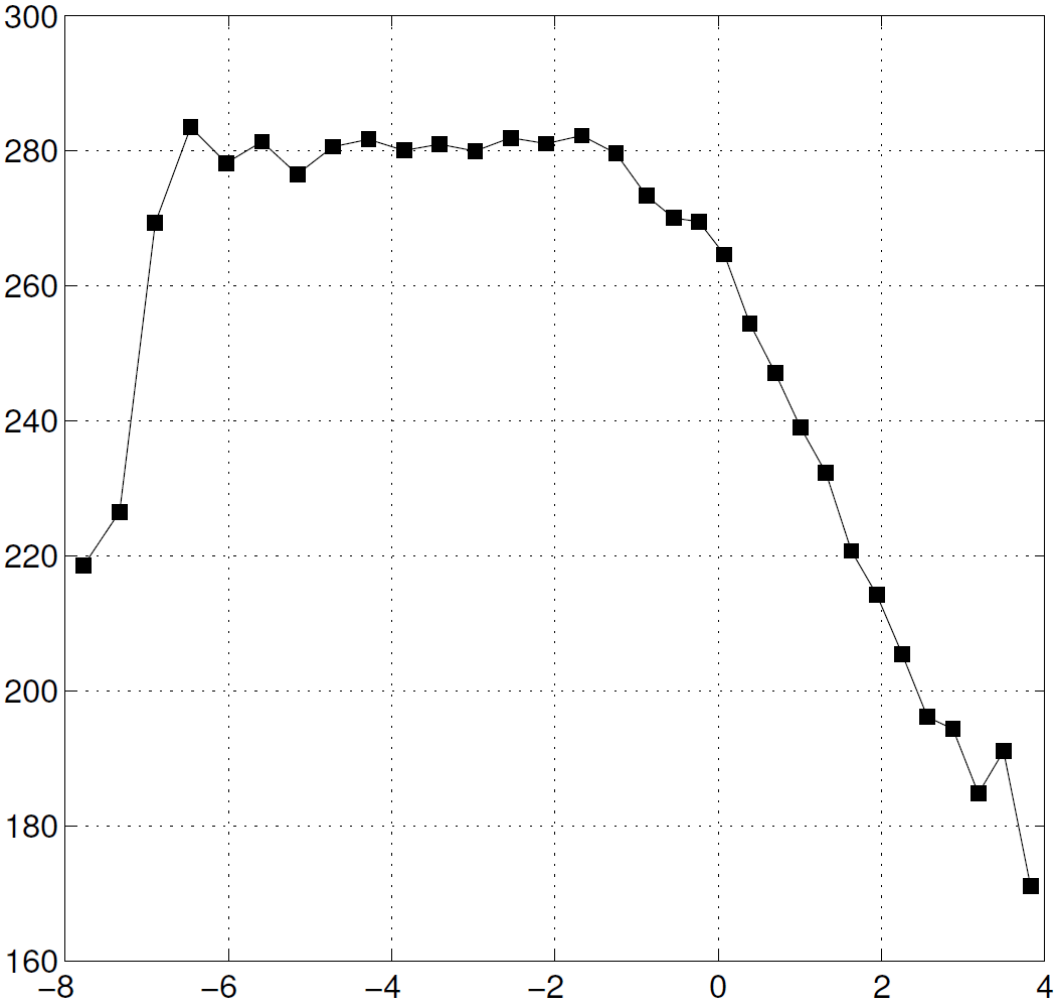




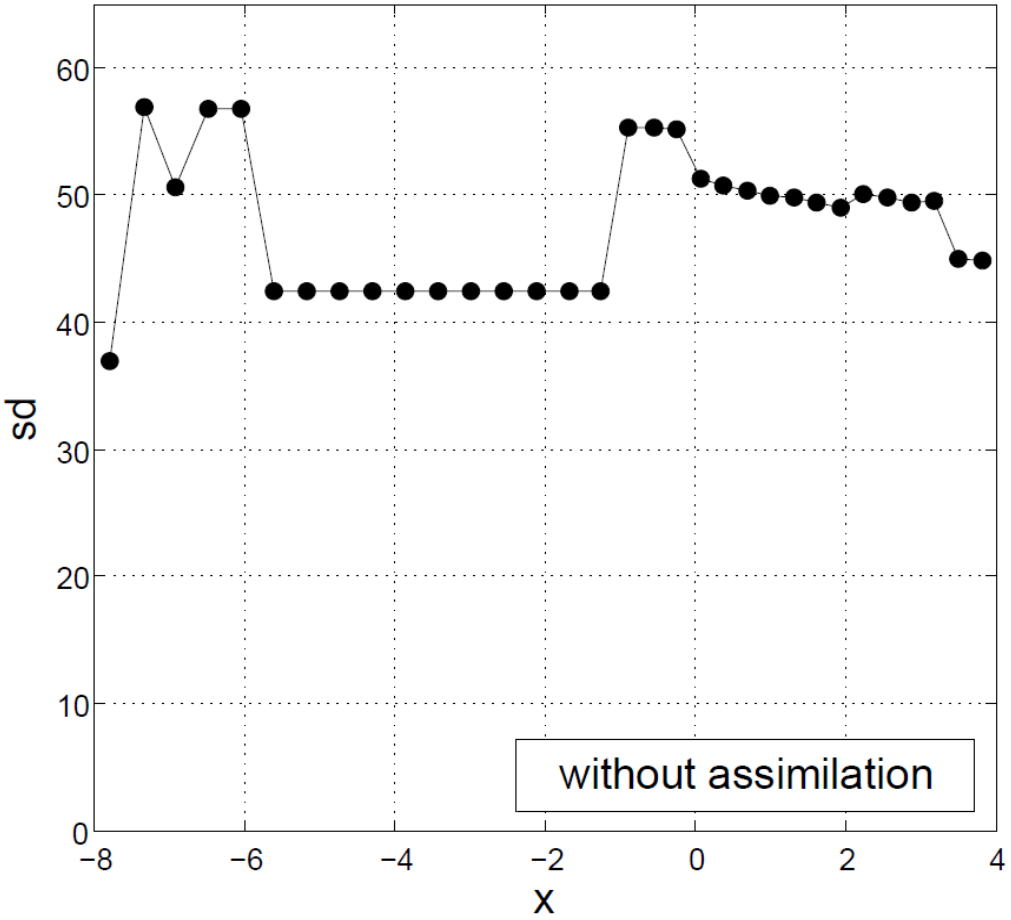
Map of $\sqrt{\lambda_{max}}$

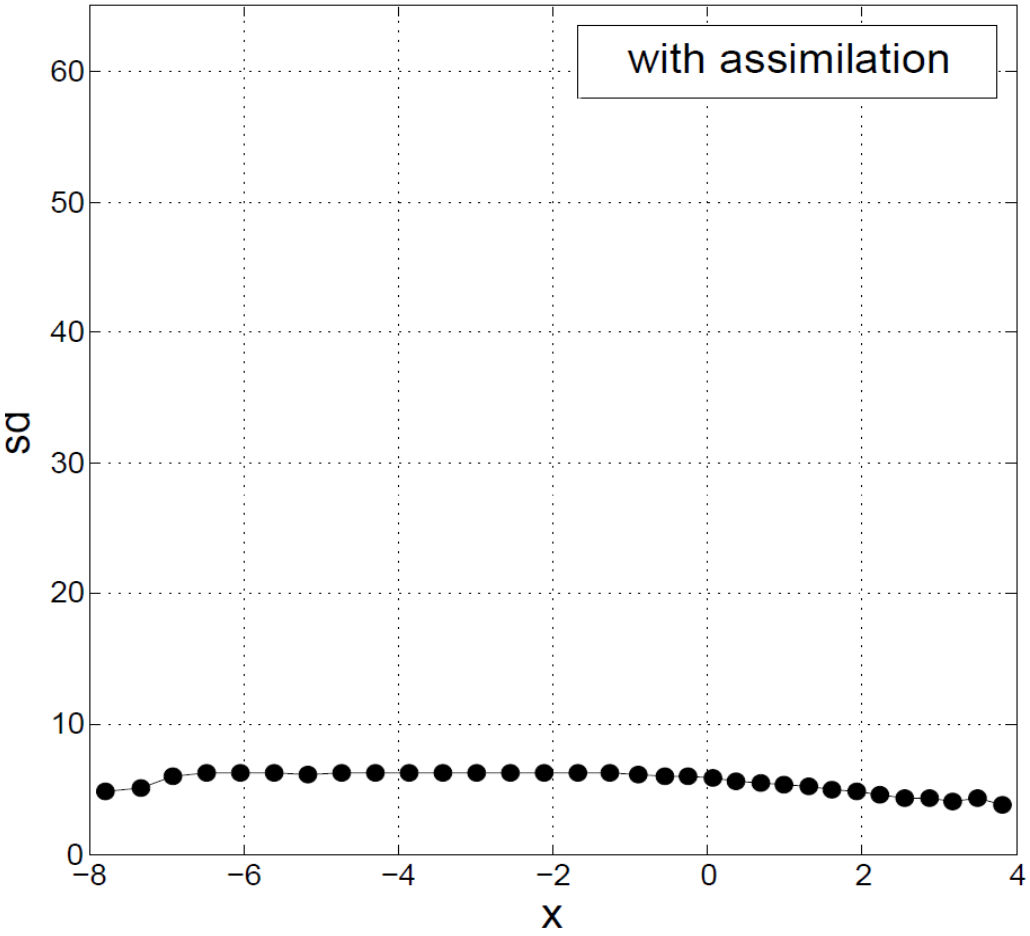
spread estimators – the wall shear stress (WSS)





spread estimators – the wall shear stress (WSS)





Conclusions:

- Progressive use of numerical simulations in clinical practice requires certified reliability

- Combination/Integration of Data and Numerical Methods can be done ***much stronger***
  - New measurements devices
  - New mathematical and numerical methods

- Proper numerical methods for the integration process ***need to be investigated.***

- ***Bayesian analysis is a viable approach for improving reliability of (variational) Data Assimilation***

Journal of Computational Physics 229 (2010) 107–118

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Weighted least-squares finite elements based on particle imaging velocimetry data

J.J. Heys [a,*], T.A. Manteuffel [b], S.F. McCormick [b], M. Milano [c], J. Westerdale [c], M. Belohlavek [d]

[a] Chemical and Biological Engineering, Montana State University, Bozeman, MT 59717-3920, United States
[b] Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309, United States
[c] Mechanical and Aerospace Engineering, Arizona State University, Tempe, AZ 85287, United States
[d] Translational Ultrasound Research Laboratory, Division of Cardiovascular Diseases, Mayo Clinic, Scottsdale, AZ 85259, United States

ARTICLE INFO

Article history:
Received 2 December 2008
Received in revised form 23 July 2009
Accepted 16 September 2009
Available online 20 September 2009

Keywords:
Least-squares
Finite element
Data assimilation
Particle imaging velocimetry

ABSTRACT

The solution of the Navier–Stokes equations requires that data about the solution is available along the boundary. In some situations, such as particle imaging velocimetry, there is additional data available along a single plane within the domain, and there is a desire to also incorporate this data into the approximate solution of the Navier–Stokes equation. The question that we seek to answer in this paper is whether two-dimensional velocity data containing noise can be incorporated into a full three-dimensional solution of the Navier–Stokes equations in an appropriate and meaningful way. For addressing this problem, we examine the potential of least-squares finite element methods (LSFEM) because of their flexibility in the enforcement of various boundary conditions. Further, by weighting the boundary conditions in a manner that properly reflects the accuracy with which the boundary values are known, we develop the weighted LSFEM. The potential of weighted LSFEM is explored for three different test problems: the first uses randomly generated Gaussian noise to create artificial 'experimental' data in a controlled manner, and the second and third use particle imaging velocimetry data. In all test problems, weighted LSFEM produces accurate results even for cases where there is significant noise in the experimental data.

1. Introduction

The solution of the Navier–Stokes equations requires the specification of a domain and data along the boundaries of that domain (i.e., boundary conditions). The development of new experimental techniques, including particle imaging velocimetry (PIV), has created a situation in which additional data may be available along a lower-dimensional region of the domain. For example, data may be available along a two-dimensional plane within a three-dimensional domain. Further, it may be desirable to incorporate this experimental data into the solution of the Navier–Stokes equations. For example, echocardiologists can use FDA-approved microbubbles and PIV to determine two components of the blood velocity along a single plane within the left ventricle of the heart [11,23,24]. Despite the abundance of data, ultrasound and PIV alone are insufficient for calculating the flow properties of interest to clinicians, such as the pressure gradient and total energy loss, which require knowledge of the full three-dimensional velocity field. This is an example of an application where it may be useful to assimilate two-dimensional velocity data into a three-dimensional solution of the Navier–Stokes equations, which is distinctly different from using experimental data (e.g., PIV) to validate a computational fluid dynamics result (e.g., [10,25]).

* Corresponding author. Tel.: +1 406 994 7902.
E-mail address: jeff.heys@gmail.com (J.J. Heys).

0021-9991/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2009.09.016

Bayesian Inference for Data Assimilation using Least-Squares Finite Element Methods

Richard P. Dwight

Assistant Professor, Aerodynamics Group, Faculty of Aerospace, TU Delft, P.O. Box 5058, 2600GB Delft, The Netherlands.

E-mail: r.p.dwight@tudelft.nl

Abstract. It has recently been observed that Least-Squares Finite Element methods (LSFEMs) can be used to assimilate experimental data into approximations of PDEs in a natural way, as shown by Heyes *et al.* in the case of incompressible Navier-Stokes flow [1]. The approach was shown to be effective without regularization terms, and can handle substantial noise in the experimental data without filtering. Of great practical importance is that – unlike other data assimilation techniques – it is not significantly more expensive than a single physical simulation. However the method as presented so far in the literature is not set in the context of an inverse problem framework, so that for example the meaning of the final result is unclear. In this paper it is shown that the method can be interpreted as finding a *maximum a posteriori* (MAP) *estimator* in a Bayesian approach to data assimilation, with normally distributed observational noise, and a Bayesian prior based on an appropriate norm of the governing equations. In this setting the method may be seen to have several desirable properties: most importantly discretization and modelling error in the simulation code does not affect the solution in limit of complete experimental information, so these errors do not have to be modelled statistically. Also the Bayesian interpretation better justifies the choice of the method, and some useful generalizations become apparent. The technique is applied to incompressible Navier-Stokes flow in a pipe with added velocity data, where its effectiveness, robustness to noise, and application to inverse problems is demonstrated.

1. Introduction

Data assimilation has long been of importance in atmospheric science and geophysics, where computational models are largely empirical, or model parameters are inaccessible to measurement. In engineering on the other hand, we are motivated by the complementary errors and limitations of simulations and experiments. The particular example considered in this article is illustrative: using the *particle image velocimetry* (PIV) measurement technique the velocity of a fluid can be measured; pairs of photos are taken (with a small time increment) of particles transported by the flow, the particles having been illuminated with laser light sheets [2]. Compared to other measurement techniques an extremely large amount of data is returned, but only in a limited 2d-planar "window" of a flow field. This window can not be too close to walls, otherwise laser reflections pollute the images, it is further limited by optical access to the flow, camera resolution, and many other practical considerations. Important quantities such as the pressure are inaccessible. Measurement errors are approximately Gaussian in distribution. On the other hand a Navier-Stokes numerical simulation can encompass the entire 3d flow field in all state variables, giving a complete picture of the flow, but includes discretization error,

Example 1 - Estimation of Cardiac conductivities

Cardiac conductivities: numerical values in Bidomain simulations matter!!!

Geselowitz (1971) – conductivity estimation in a (spherical) volume by impedance measures

Experimental works

Clerc (1976), Roberts et al. (1979), Roberts et al. (1982)

No common agreement

TABLE 1. The measured bidomain conductivities (mS cm^{-1}). [43]

Symbol	Clerc (1976)[4]	Roberts *et al.* (1979)[40]	Roberts and Scher (1982)[41]
σ_{il}	1.70	2.80	3.40
σ_{it}	0.19	0.26	0.60
σ_{el}	6.20	2.20	1.20
σ_{et}	2.40	1.30	0.80
σ_{il}/σ_{it}	8.95	10.77	5.67
σ_{el}/σ_{et}	2.58	1.69	1.50
σ_{il}/σ_{el}	0.27	1.27	2.83
σ_{it}/σ_{et}	0.08	0.200	0.75

System Dynamics: Bidomain system

$$
\begin{cases}
\dfrac{\partial u}{\partial t} - \nabla \cdot \boldsymbol{\sigma}_i \nabla u - \nabla \cdot \boldsymbol{\sigma}_i \nabla u_e + f(u,w) = I_{si} & \text{in } Q \\
-\nabla \cdot \boldsymbol{\sigma}_i \nabla u - \nabla \cdot (\boldsymbol{\sigma}_i + \boldsymbol{\sigma}_e) \nabla u_e = I_{si} - I_{se} & \text{in } Q \\
\dfrac{\partial w}{\partial t} + g(u,w) = 0 & \text{in } Q \\
(\boldsymbol{\sigma}_i \nabla u + \boldsymbol{\sigma}_i \nabla u_e) \cdot \mathbf{n} = 0, \quad \boldsymbol{\sigma}_e \nabla u_e \cdot \mathbf{n} = 0 & \text{on } \partial Q \\
u(x,0) = u_0, \quad w(x,0) = w_0 & \text{in } \Omega
\end{cases}
$$

Rogers-McCulloch

The Bidomain solutions have been proved to be strongly sensitive to the values of conductivities

with

$Q \equiv \Omega \times [0,T]$

u = transmembrane potential

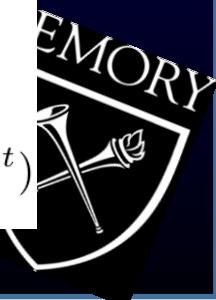
u_e = extracellular potential

``Mondomain model'' is a significant simplification under some (not realistic) assumptions

$$
I_{si,se} = \frac{1}{C_m} \bar{I}_{si,se}, \quad f(u,w) = \frac{1}{C_m} I_{ion}(u,w),
$$

$$
\boldsymbol{\sigma}_i = \frac{1}{\beta C_m} \bar{\boldsymbol{\sigma}}_i = \frac{1}{\beta C_m} (\sigma_{il} \mathbf{a_l} \mathbf{a_l}^t + \sigma_{it} \mathbf{a_t} \mathbf{a_t}^t + \sigma_{in} \mathbf{a_n} \mathbf{a_n}^t),
$$

$$
\boldsymbol{\sigma}_e = \frac{1}{\beta C_m} \bar{\boldsymbol{\sigma}}_e = \frac{1}{\beta C_m} (\sigma_{el} \mathbf{a_l} \mathbf{a_l}^t + \sigma_{et} \mathbf{a_t} \mathbf{a_t}^t + \sigma_{en} \mathbf{a_n} \mathbf{a_n}^t)
$$



Statement of the Inverse Conductivity Problem (Variational approach)

Consider the mismatch functional

$$\mathcal{J} = \frac{1}{2}\int_0^T \int_{\Omega_{obs}} |u(\boldsymbol{\sigma}) - u_m|^2 + \mathcal{R}(\boldsymbol{\sigma})$$

with the admissible set

$$\mathcal{C}_{adm} = \left\{(\sigma_{il}, \sigma_{it}, \sigma_{el}, \sigma_{et}) \in (L^\infty(\Omega))^4, \boldsymbol{\sigma} \in [0 < m, M]^4\right\}$$

$\mathcal{R}$ = Tikhonov-like regularization, Ω_{obs} = Observation subregion

Problem:

Find admissible $\boldsymbol{\sigma}$ that minmizes $\mathcal{J}$
under the constraint of the Bidomain equations.

- Extension to 3D immediate
- Possible regularization (for an available and trustworthy guess $\widehat{\boldsymbol{\sigma}}$)

$$\mathcal{R} = \frac{\alpha}{2}\int_\Omega |\boldsymbol{\sigma} - \widehat{\boldsymbol{\sigma}}|_F^2$$

Statement of the Inverse Conductivity Problem

VARIATIONAL APPROACH

Let us introduce the following functional

$$J = \frac{1}{2}\int_0^T \int_{\Omega_{obs}} \|u(\boldsymbol{\sigma}) - u_m\|^2 + \mathcal{R}(\boldsymbol{\sigma})$$

Ω_{obs} = observation subregion

$\mathcal{R}$ = Tikhonov regularization

Admissible set (in 2D)

$$\mathcal{C}_{ad}^{2D} = \{\boldsymbol{\sigma} = (\sigma_{il}, \sigma_{it}, \sigma_{el}, \sigma_{et}) \in (L^\infty(\Omega))^4 : \boldsymbol{\sigma}(x) \in [m, M]^4, \forall x \in \Omega\}$$

PROBLEM

Find admissible σ that minimizes J under the constraint of the Bidomain equations/conditions

Possible regularization: given a reasonable guess $\hat{\sigma}$ and for $\|\cdot\|_F$ being the Frobenius norm

$$\mathcal{R} = \frac{\alpha}{2}\int_\Omega \|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}\|_F^2$$

Other regularization terms (and 3D problem) may be considered as well.

Well-posedness analysis

Lemma(A priori estimates)(C. Nagaiah, K. Kunisch, and G. Plank)
If Ω has a Lipschitz boundary and the conductivities tensors are uniformly elliptic, there exist positive constants $\tilde{C}$ and $\tilde{c}$, such that

$$|u|_{C(H)} + |u|_{L^2(V)} + |u|_{L^4(Q)} + |u_t|_{L^{4/3}(V^*)} + |v|_{L^2(U)} + |w|_{C(H)} + |w_t|_{L^2(H)}$$

$$\leq \tilde{C}(|u_0| + |v_0| + \tilde{c}|\Omega| + |I_{si}|_{L^2(V^*)} + |I_{si} - I_{se}|_{L^2(U^*)}).$$

where $\tilde{c}$ depends on the ionic model, $\tilde{C}$ depends on m, M and the ionic model, but is independent of $(u_0, w_0), (I_{si}, I_{se})$ and $\boldsymbol{\sigma}_{i,e}$.

Theorem(Existence of minimizer)(H. Yang and A. Veneziani)
Under the assumption condition as before, for $\alpha \geq 0$, there exists at least one minimizer to the optimization problem.

OPEN ISSUE: Uniqueness (of course?)
More precisely:
is there a class of parameters (e.g. constant, piecewise constant) s.t. the minimizing set is unique?

The Minimal ionic model

$$
\begin{aligned}
f(u,v,w,s) &= 85.7\beta C_m(J_{fi}+J_{so}+J_{si})\\
J_{fi} &= -vH(\tilde{u}-\theta_v)(\tilde{u}-\theta_v)(u_u-\tilde{u})/\tau_{fi}\\
J_{so} &= (\tilde{u}-u_o)(1-H(\tilde{u}-\theta_w))/\tau_o + H(\tilde{u}-\theta_w)/\tau_{so}\\
J_{si} &= -H(\tilde{u}-\theta_w)ws/\tau_{si}\\
\partial_t v &= (1-H(\tilde{u}-\theta_v))(v_\infty - v)/\tau_v^- - H(\tilde{u}-\theta_v)v/\tau_v^+\\
\partial_t w &= (1-H(\tilde{u}-\theta_w))(w_\infty - w)/\tau_w^- - H(\tilde{u}-\theta_w)w/\tau_w^+\\
\partial_t s &= ((1+\tanh(k_s(\tilde{u}-u_s)))/2 - s)/\tau_s
\end{aligned}
$$

where $\tilde{u} = (u+84)/85.7$ and $H(x)$ is the standard Heaviside function.

$$
\begin{aligned}
\tau_v^- &= (1-H(\tilde{u}-\theta_v^-))\tau_{v1}^- + H(\tilde{u}-\theta_v^-)\tau_{v2}^-\\
\tau_w^- &= \tau_{w1}^- + (\tau_{w2}^- - \tau_{w1}^-)(1+\tanh(k_w^-(\tilde{u}-u_w^-)))/2\\
\tau_{so} &= \tau_{so1} + (\tau_{so2}-\tau_{so1})(1+\tanh(k_{so}(\tilde{u}-u_{so})))/2\\
\tau_s &= (1-H(\tilde{u}-\theta_w))\tau_{s1} + H(\tilde{u}-\theta_w)\tau_{s2}\\
\tau_o &= (1-H(\tilde{u}-\theta_o))\tau_{o1} + H(\tilde{u}-\theta_o)\tau_{o2}\\
v_\infty &= 1-H(\tilde{u}-\theta_v^-)\\
w_\infty &= (1-H(\tilde{u}-\theta_o)(1-\tilde{u}/\tau_{w\infty}) + H(\tilde{u}-\theta_o)w_\infty^*
\end{aligned}
$$

Bueno-Orovio, Cherry, Fenton

Numerical Solution

Forward Problem

Bidomain equations with Rogers-McCulloch model

- Quadratic Finite Elements
- BDF2 Time discretization (FD)

→ In 2D : FreeFem++ (direct solver)
→ In 3D : LifeV (C++ Solver) using an extension of the Monodomain model as preconditioner (Gerardo-Giorda et al. JCP 2009)

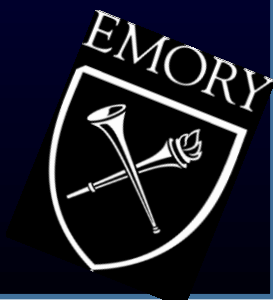
Inverse Problem

- Optimize then Discretize approach
- Dual equations (backward in time) solved with FEM (space) and FD (time)

Steps:
→ solve the system dynamics (over the entire time interval)
→ solve the dual equations (backward in time)
→ compute the functional at the current guess
→ compute the integral of the functional DJ with the dual equations
→ line search (BFGS)

Stopping criterion

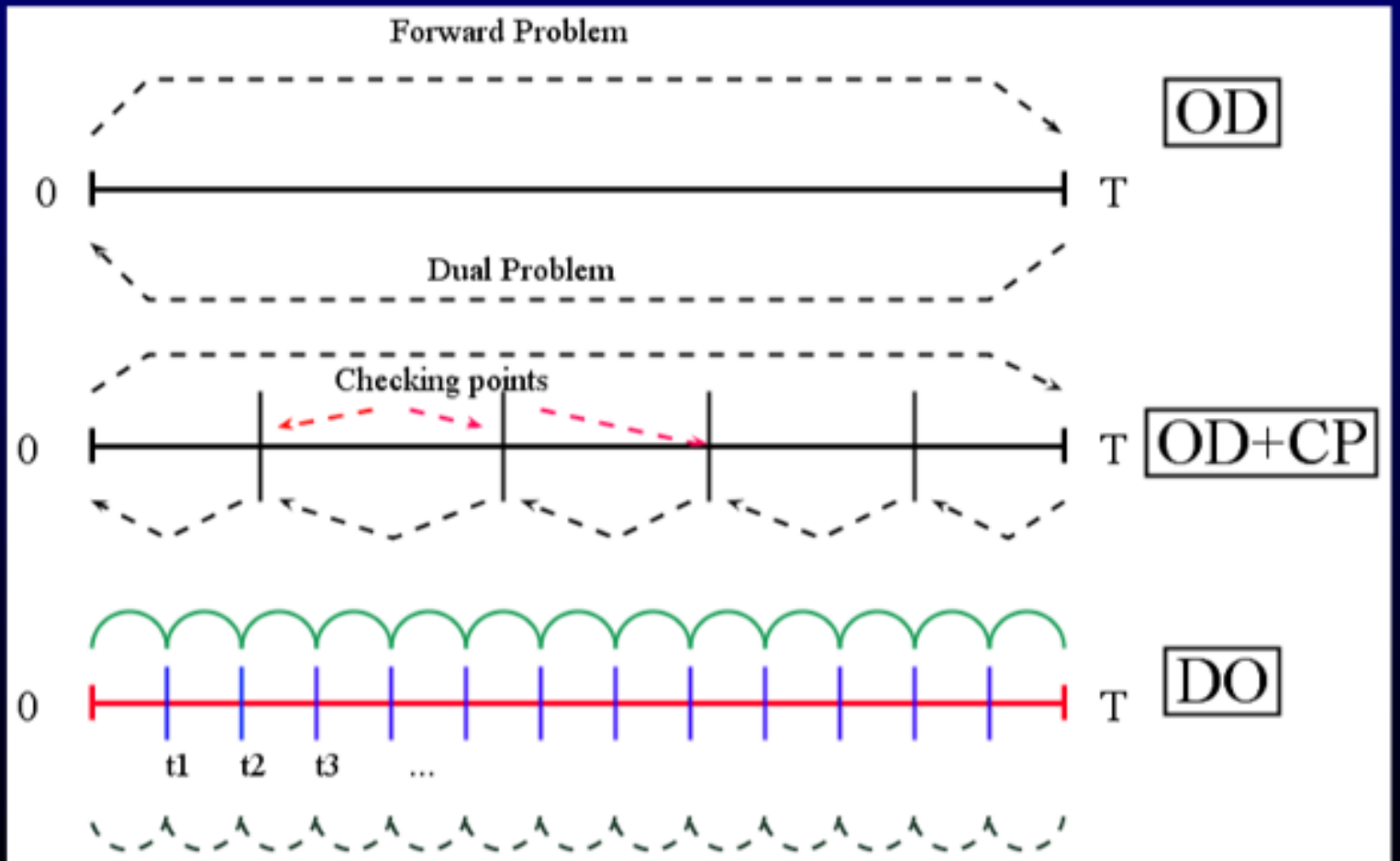
$$\|\mathcal{D}J\| \leq 10^{-6}$$



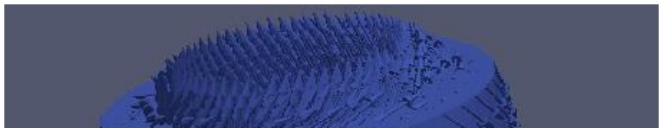
OD vs DO dilemma

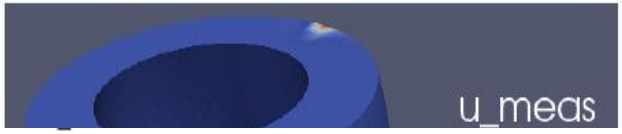
Optimize the Discretize:
Global optimization in time
Dual problem ``backward'' in time:
high computational/storage costs
Checking points:
→ storage only in a limited number of points+
→ local-in-time reconstruction of the forward solution

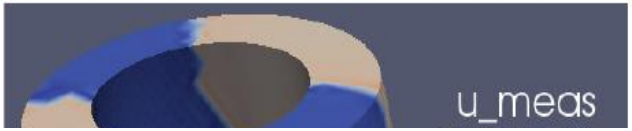
Discretize then Optimize:
Time sliced by the discretization
Optimization at each time step
(suboptimal)
Fast



Consistency check in 3D: Truncated Ellipsoid





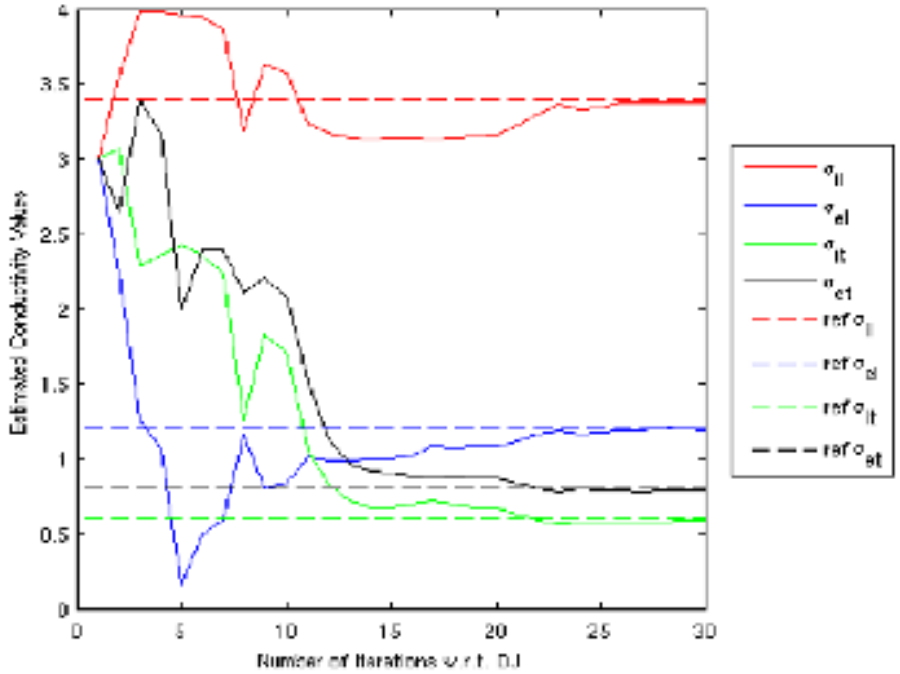


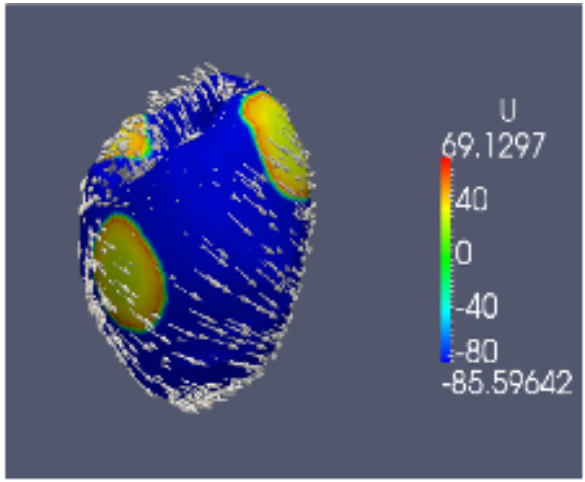
$\boldsymbol{\sigma}_{exact}$ $\boldsymbol{\sigma}_{initial}$	[3.5,3,0.3,1.8] [5, 5, 3, 3]	[3.5,3,0.3,1.8] [5, 5, 3, 3]	[3.5,3,0.3,1.8] [2, 2, 1, 1]	[3.5,3,0.3,1.8] [2, 2, 1, 1]
# sites	100	200	100	200
σ_{il}	3.50637	3.48173	3.51006	3.48027
σ_{el}	3.18176	3.0039	3.17776	3.00454
σ_{it}	0.312358	0.2707	0.311667	0.270614
σ_{et}	1.55211	1.86707	1.55151	1.86914
# fwd|bwd	27|25	27|25	18|17	19|19

$\boldsymbol{\sigma}_{exact}$ $\boldsymbol{\sigma}_{initial}$	[2.8,2.2,0.26,1.3] [5, 5, 3, 3]	[2.8,2.2,0.26,1.3] [2, 2, 1, 1]	[3.5,3,0.3,1.8] [5, 5, 3, 3]	[3.5,3,0.3,1.8] [2, 2, 1, 1]
σ_{il}	2.62885	2.74088	3.4704	3.47661
σ_{el}	2.37449	2.29435	3.12013	3.11378
σ_{it}	0.27818	0.274675	0.306733	0.306189
σ_{et}	1.33207	1.24734	1.6459	1.64325
# fwd|bwd	26|22	17|12	24|22	21|20

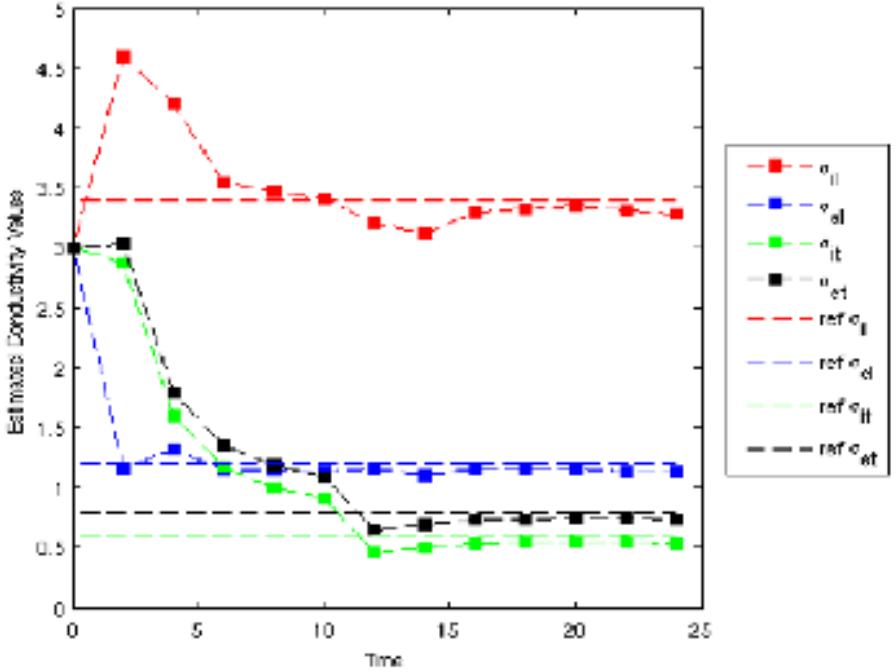
OD+CP

CPU time=13h
(4 processors)



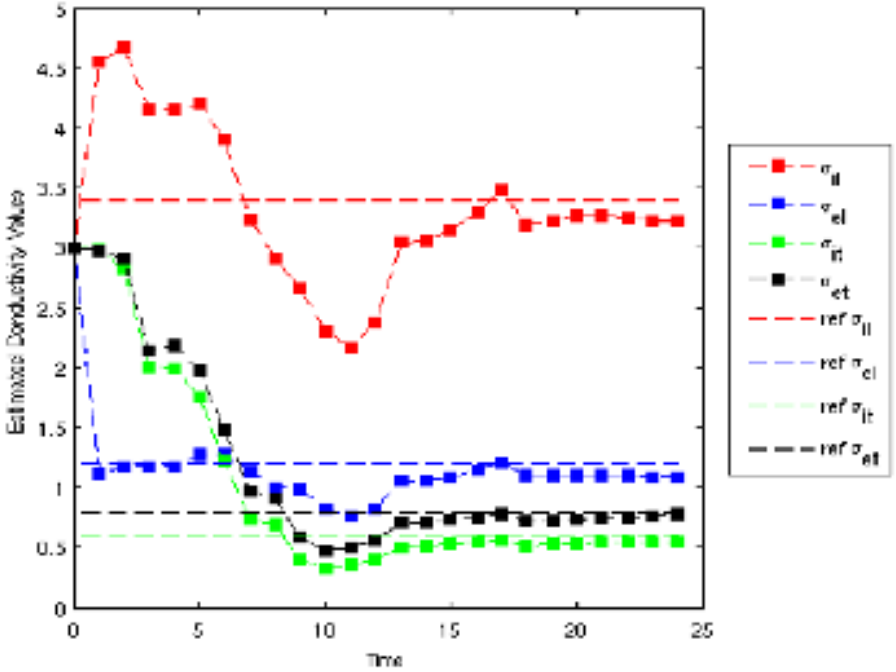


DO (Δt=2ms)



CPU time = 6h

DO (Δt=1ms)



CPU time = 5h

An example with a scar (region with a different conductivity)

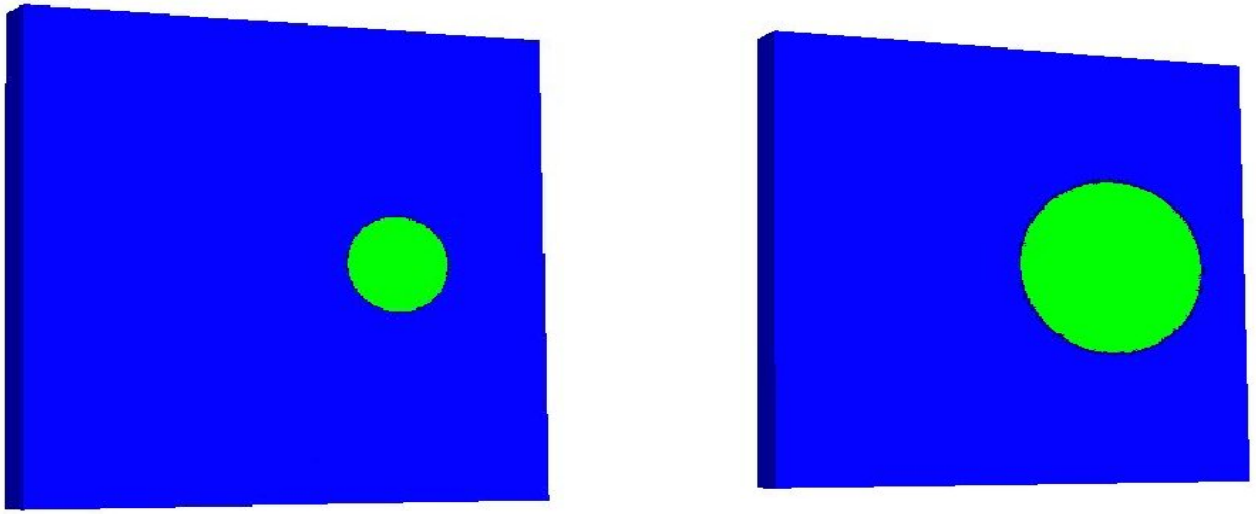
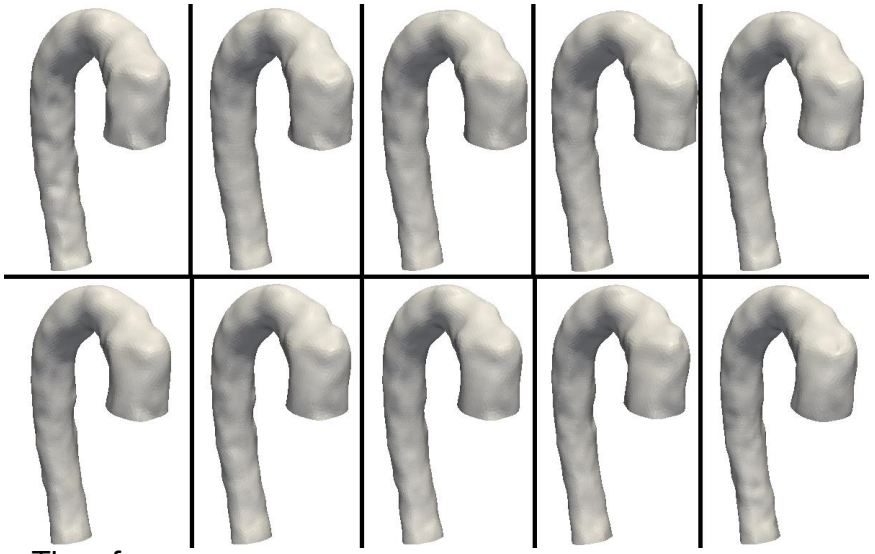


Figure 4: Left: scar with radius 0.5cm, Right: scar with radius 1cm.

Table 8: Slab with Bidomain. $\alpha = 0, T = 20ms, \Delta t = 0.1ms$ with $\Delta t_{\text{snap}} = 2ms$, # sites = 400, noise = 10%.

r_{scar}	σ_{il}^{in}	σ_{el}^{in}	σ_{it}^{in}	σ_{et}^{in}	σ_{il}^{out}	σ_{el}^{out}	σ_{it}^{out}	σ_{et}^{out}	# fwd|bwd
0.25	1.577	1.219	0.2006	0.0547	3.288	1.213	0.6137	0.8341	284|144
0.5	1.871	0.8551	0.2257	0.09952	3.205	1.278	0.6337	0.8356	340|161
0.75	1.686	0.9116	0.2409	0.1216	3.332	1.201	0.6219	0.8094	137|76
1	1.991	0.9857	0.2061	0.1034	3.329	1.212	0.5881	0.8054	103|58
$\boldsymbol{\sigma}_{exact}$	2	1	0.2	0.1	3.4	1.2	0.6	0.8	
$\boldsymbol{\sigma}_{initial}$	1	1	1	1	1	1	1	1	

Example 2 - An Elastography-like approach for Arterial Compliance Estimation

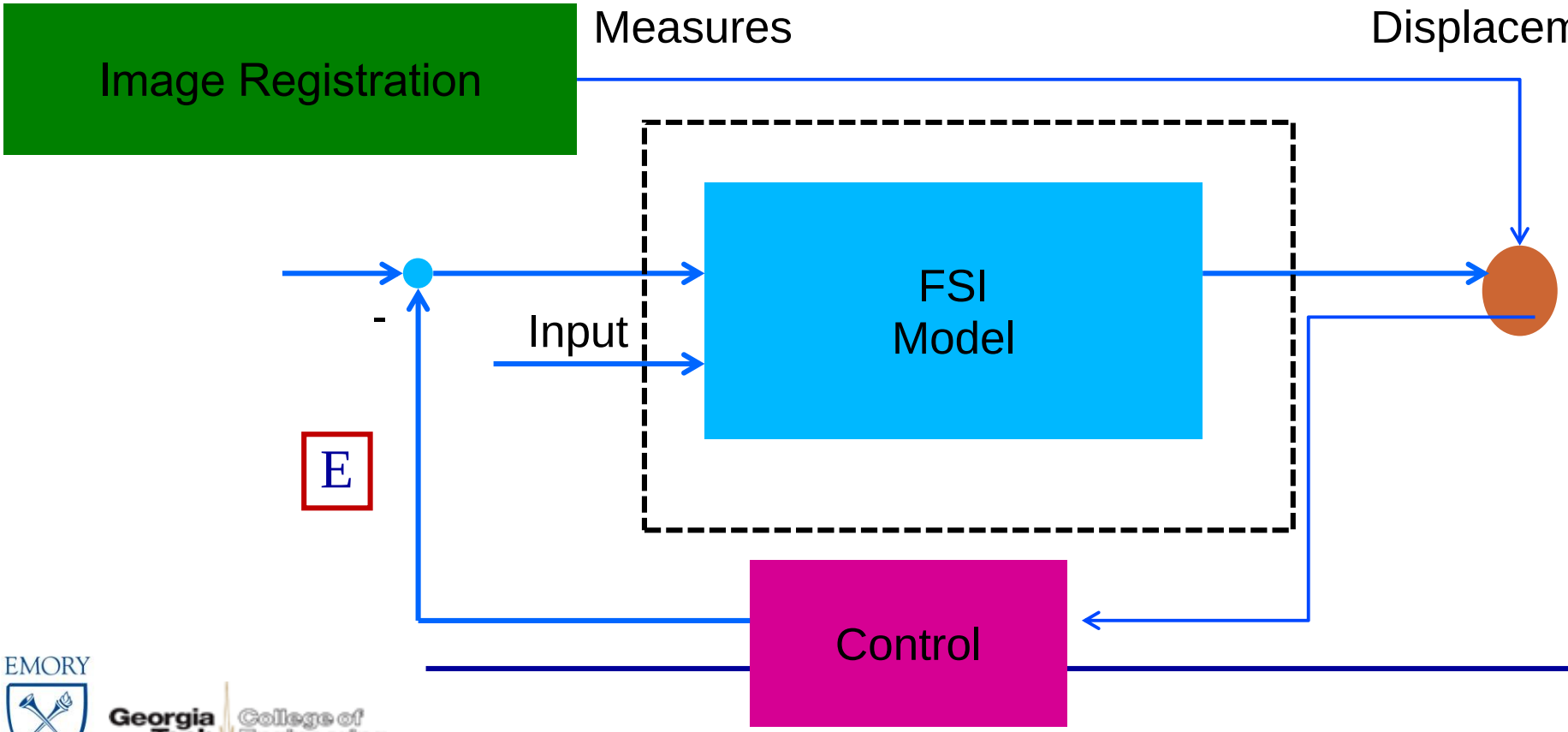


Time frames

GOAL: to estimate the vessel compliance by means of

1. Image registration
2. Control theory

Solution of an INVERSE FSI Problem



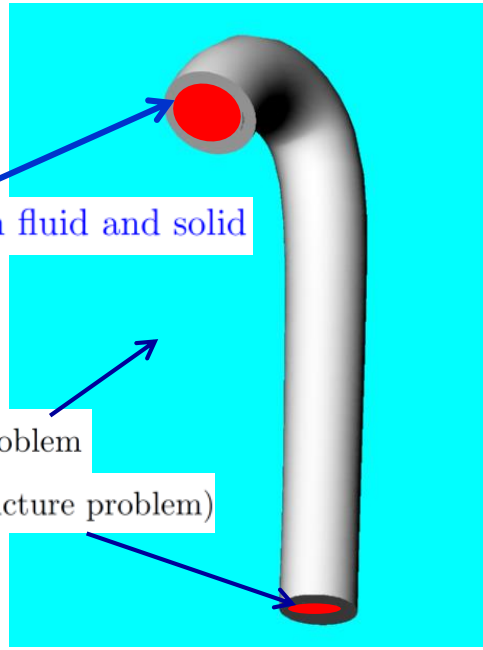
Inverse FSI Problem

Let $\boldsymbol{\eta}$ be the **computed displacement field** and $J(\boldsymbol{\eta})$ a functional measuring the "distance" data-simulations:

$$J(\boldsymbol{\eta}) = \int_{\Sigma} (\boldsymbol{\eta} - \boldsymbol{\eta}_{\text{meas}})^2 d\sigma$$

+ possible regularizing terms

Measured displacement (from the registration step)



Σ = interface between fluid and solid

Ω_s = domain of the structure problem

($\hat{\Omega}_s$ = *reference* domain of the structure problem)

Ω_f = domain of the fluid problem

The displacement field $\boldsymbol{\eta}$ is the solution of a FSI problem, depending on some parameters $\boldsymbol{\beta}$ to be used as control variables (and related to the compliance).

The FSI problem is regarded as a *constraint to the minimization process*, so that $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\beta}^\star)$.

Find the parameters $\boldsymbol{\beta}^\star$ such that

$$J(\boldsymbol{\eta}(\boldsymbol{\beta}^\star)) \leq J(\boldsymbol{\eta}) \qquad \forall \boldsymbol{\eta}$$

under the conditions

$$\begin{cases} \text{Fluid}\,(\boldsymbol{u}, p) = 0 & \text{Incompressible Navier-Stokes} \\ \text{Structure}(\boldsymbol{\eta}) = 0 & \text{Linear elasticity} \\ \text{Matching conditions}\,(\boldsymbol{u}, p, \boldsymbol{\eta}) = 0 & \text{Standard} \end{cases}$$

$$\boldsymbol{T}_s = \gamma_1(\nabla\boldsymbol{\eta} + (\nabla\boldsymbol{\eta})^T) + \gamma_2(\nabla\cdot\boldsymbol{\eta})\boldsymbol{I}$$

$$\gamma_1 = \frac{E}{1+\nu}, \qquad \gamma_2 = \frac{E\nu}{(1+\nu)(1-2\nu)} + \gamma_1,$$

Two parameters: E and ν

EMORY Georgia Tech College of Engineering

Simulations details:

- 49,146 velocity DOFs (P2), 2362 pressure DOFs (P1), 12,972 displacement DOFs (P2).
- inflow: sinusoidal pressure wave $p(t) = 1e3 + 5e2\sin(100\pi t)$.
- outflow: 0D-Windkessel ($R_p = 400$, $R_d = 6.2e3$, $C = 2.72e-4$)

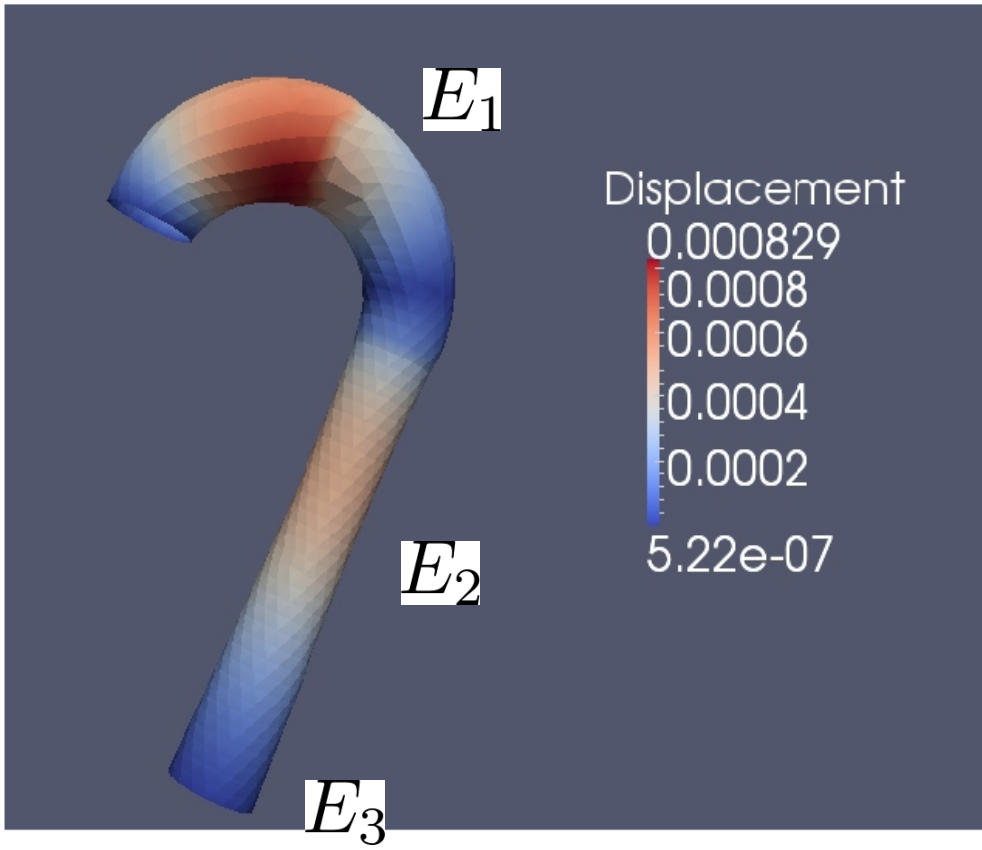
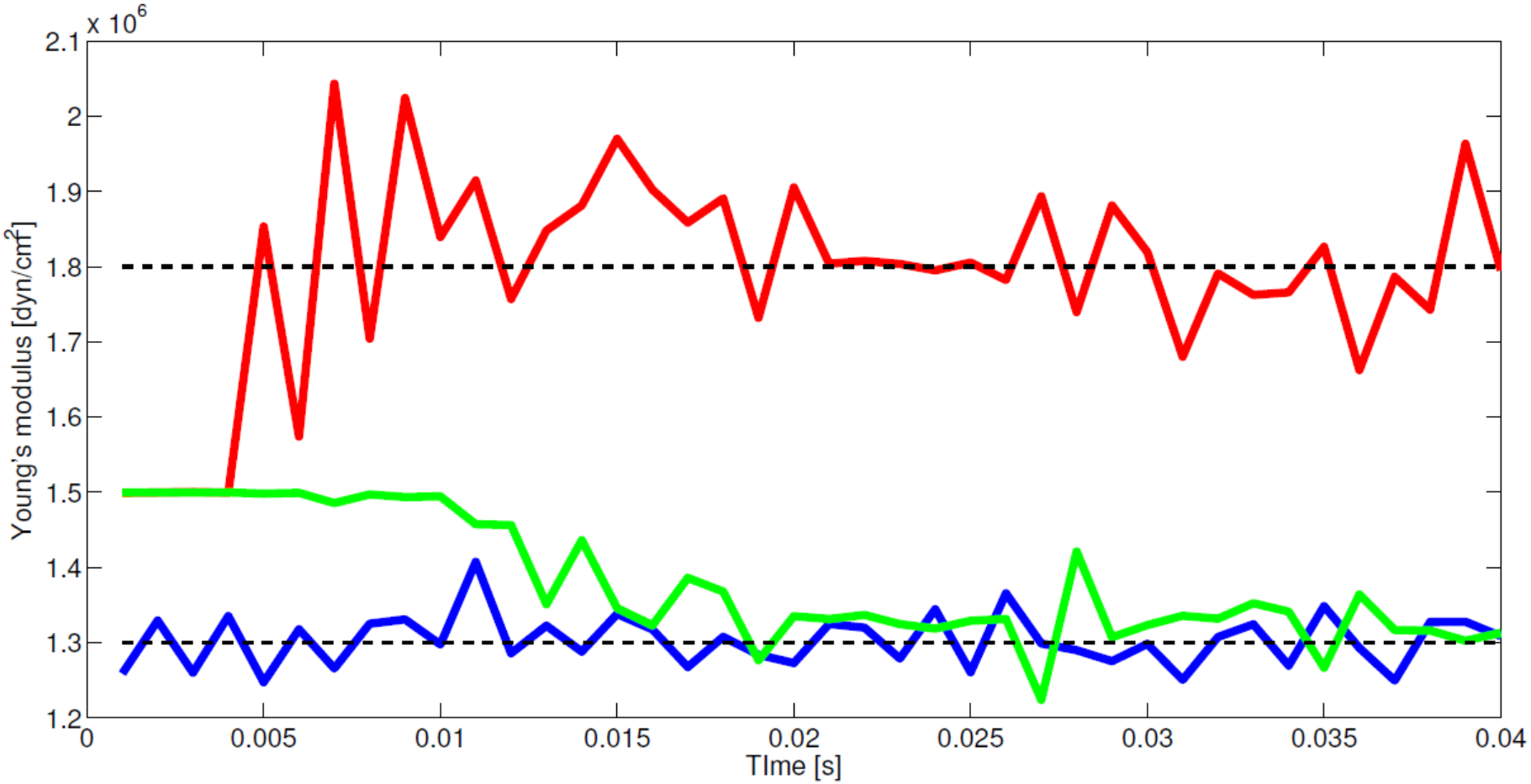


Figure: Displacement field at $t = 0.03s$



Time history of the Young's modulus estimate. SNR $= 5$. Tikhonov regularization parameter $\alpha = 10^{-2}$.

Variational DA may reduce Uncertainty but with High Computational Costs

Surrogate solution based on POD analysis

For the Inverse FSI problems and the compliance estimation

Proper Orthogonal Decomposition (POD) in a nutshell

1) **OFFLINE: compute the solution for different values of the Young modulus (chosen a priori)**
2) **Filter the Redundancy: compute the SVD of the snapshots and take the eigenfunctions corresponding to the largest SV**
3) **ONLINE: Use the eigenfunctions to construct a solution**

Snapshots based (POD) surrogate solution

$$\mathbf{u}(E) = \sum_{i=1}^{N_u} \alpha_i \mathbf{V}(E_i)$$

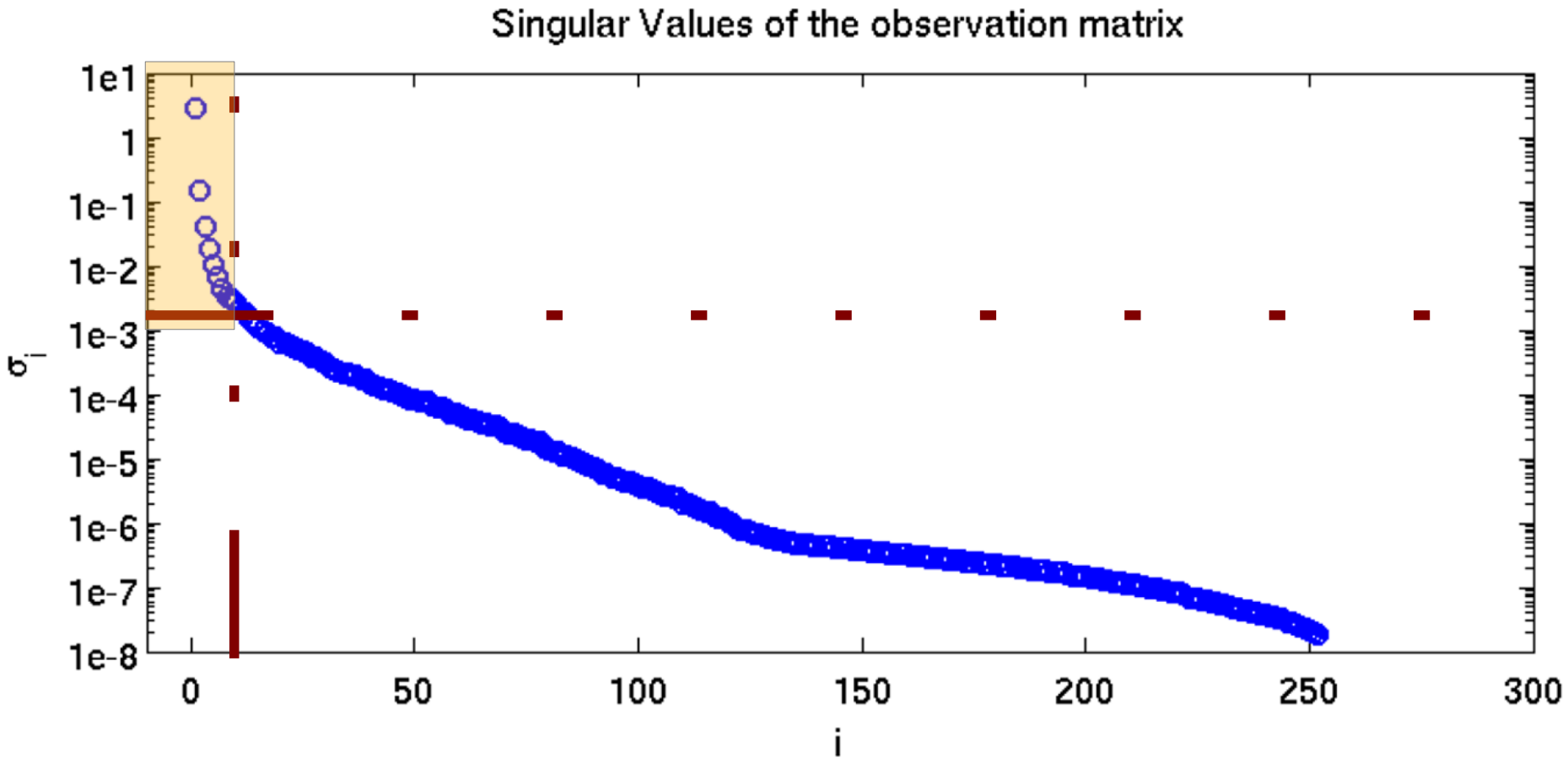
$$p(E) = \sum_{i=1}^{N_p} \beta_i Q(E_i)$$

$$\eta(E) = \sum_{i=1}^{N_\eta} \beta_i \mathbf{H}(E_i)$$

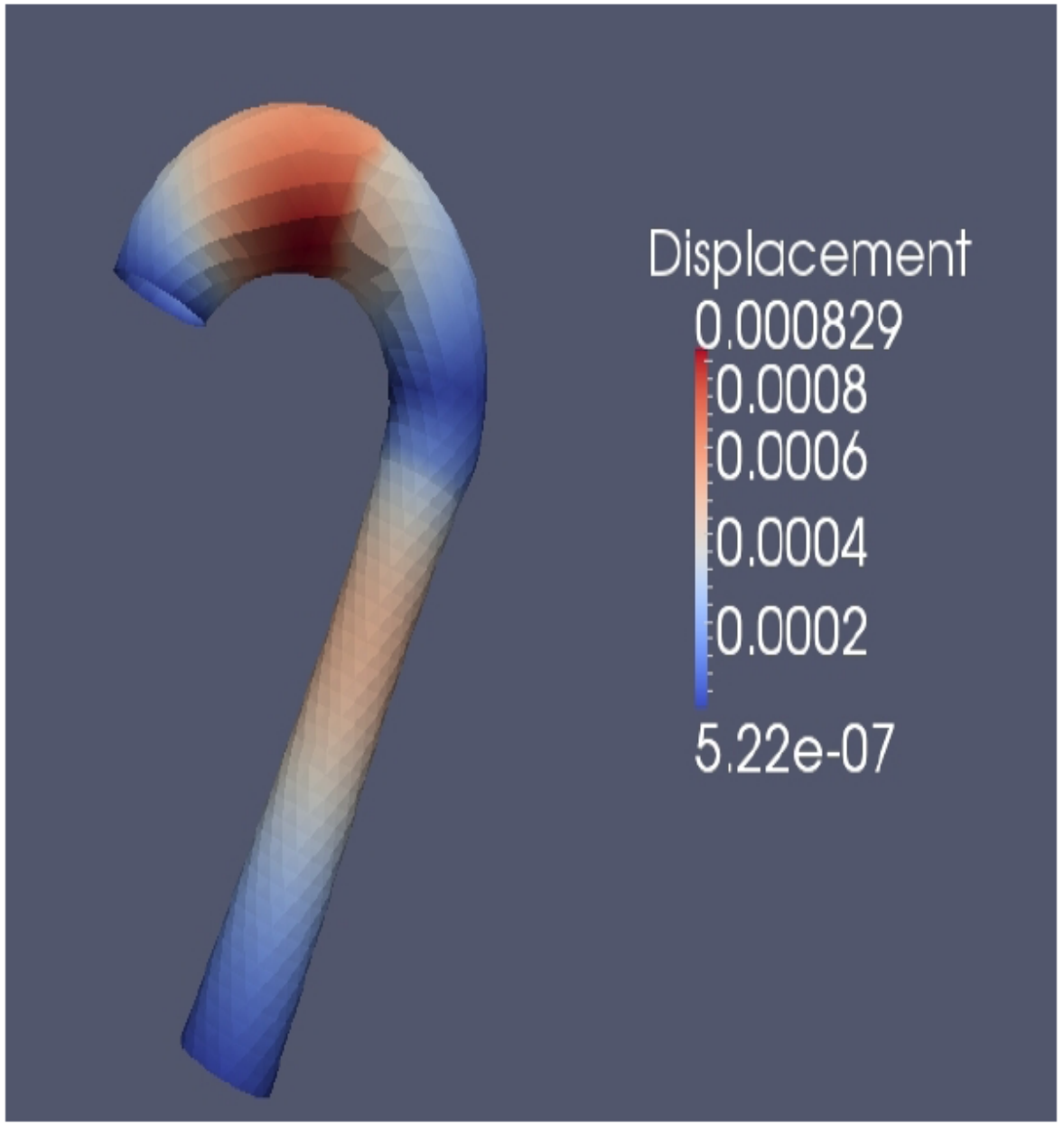
Small size

L.Bertagna, A.V., Inverse Problems (2014)

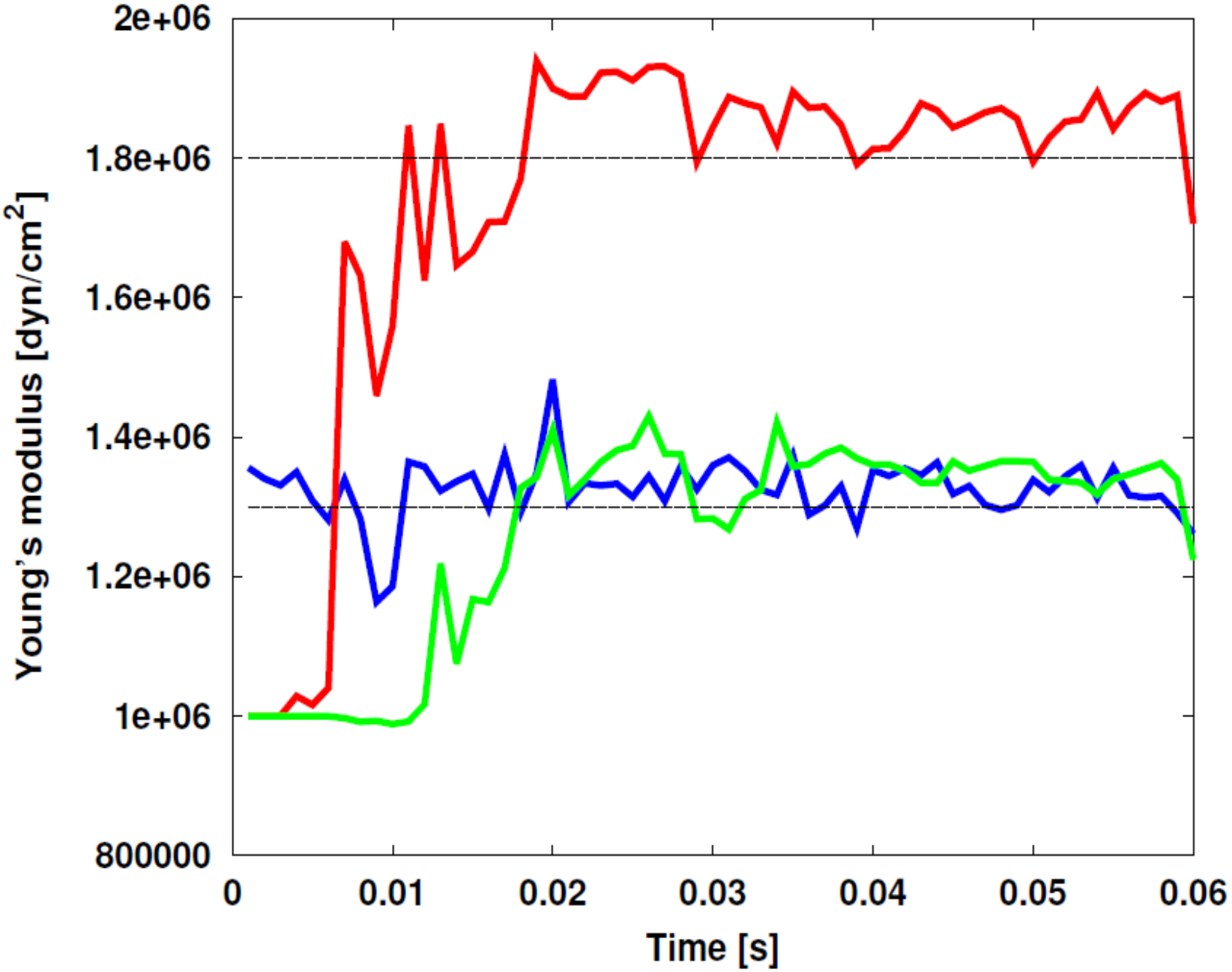
A “toy” problem on a 3D cylindrical domain



First 250 singular values of the correlation matrix $C = X^T X$, where the columns of X are snapshots of $\underline{\eta}(t)$ obtained with E=1e+06 and E=3e+06 and $0 \leq t \leq 0.2s$.



- 13,017 velocity DOFs (P2), 723 pressure DOFs (P1), 6,402 displacement DOFs (P2).
- inflow: sinusoidal pressure wave $p(t) = 1e3 \sin(100\pi t)$.
- outflow: 0D-Windkessel ($R_p = 400$, $R_d = 6.2e3$, $C = 2.72e-4$)


2e+06
1.8e+06
1.6e+06
1.4e+06
1.2e+06
1e+06
800000
Young's modulus [dyn/cm^2]
0
0.01
0.02
0.03
0.04
0.05
0.06
Time [s]

	FS	RS
$\underline{E}^*$	[1.33, 1.84, 1.31]*10^6	[1.34, 1.80, 1.33]*10^6
rel. error	1.91%	2.01%
exec. time	3176s	277s
NS solves	492	480 (offline)

Table: Comparison between Full Space (FS) and Reduced Space (RS) performance for the idealized aortic arch test case (for RS, $\tau = 0.95$).

	$\tau = 0.9$	$\tau = 0.95$	$\tau = 0.99$
$\underline{E}^*$	[1.37, 1.81, 1.32]*10^6	[1.34, 1.80, 1.32]*10^6	[1.30, 1.78, 1.29]*10^6
rel. error	2.83%	1.97%	0.87%

Table: Time average of the estimates and relative error for different values of the POD threshold for the idealized aortic arch test case.

MAJOR PROBLEM: How to use the offline part for several patients

POD has a(n expected) regularizing effect → ATLAS of Vascular Geometries

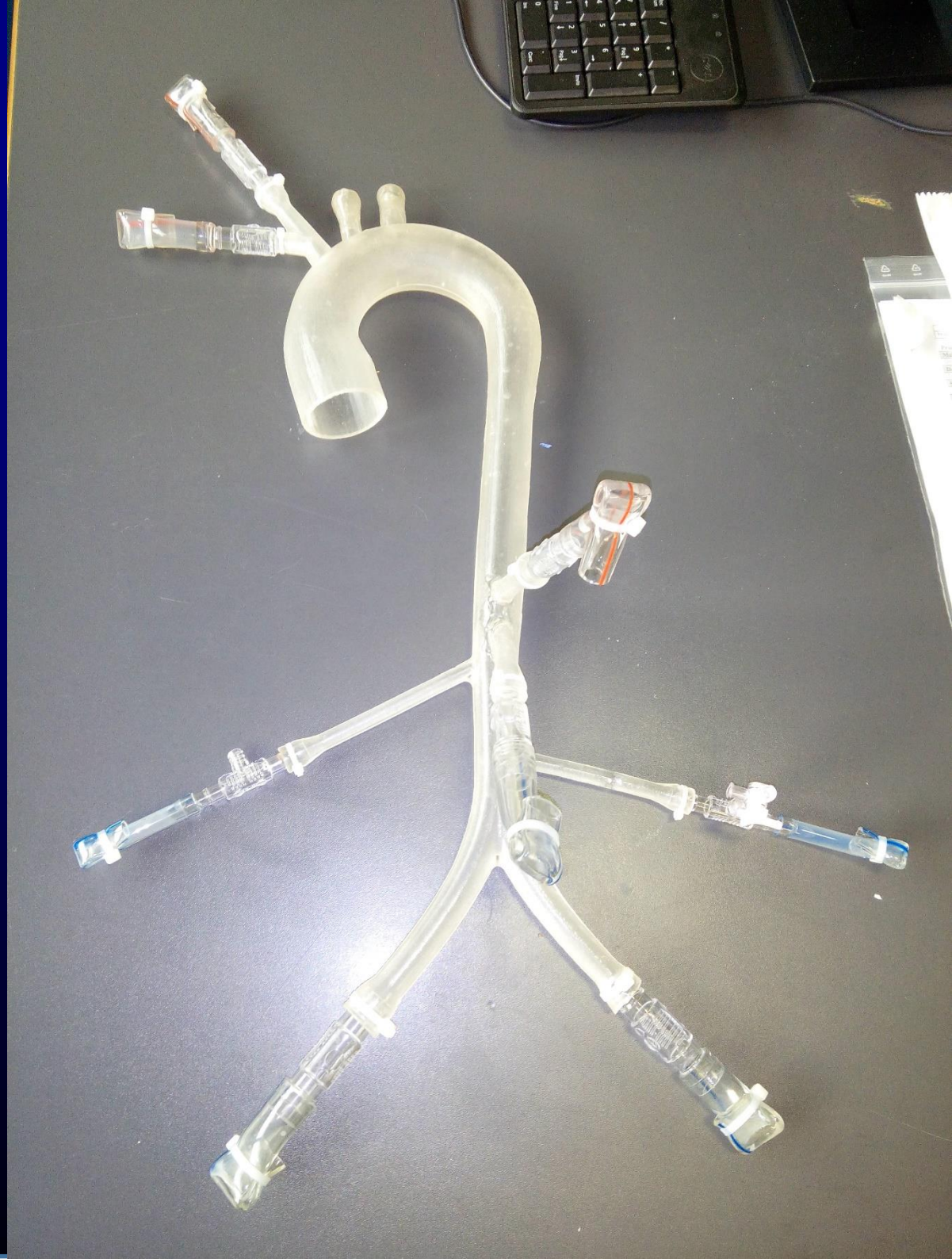
ATLAS of vascular geometries to take advantage of the offline phase with many patients

TO DO: In vitro/in vivo validation

TO DO - Validation

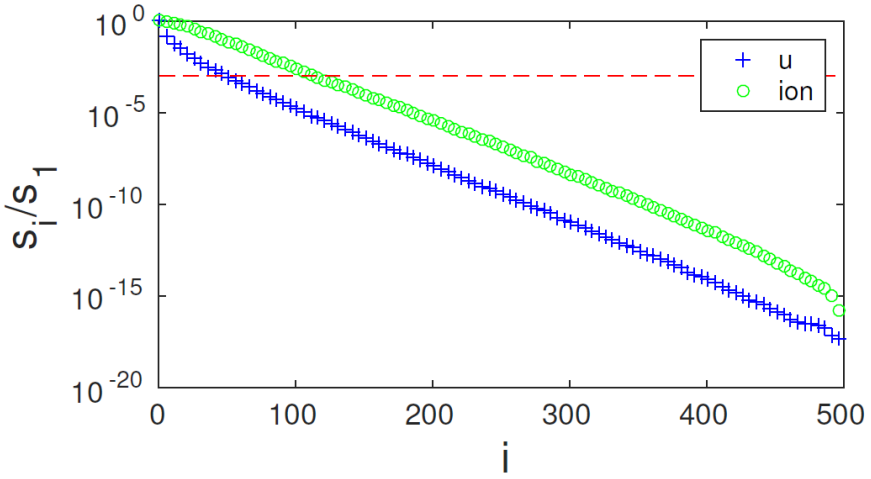
Cast obtained in Pavia (β-lab) with a 3D Printer

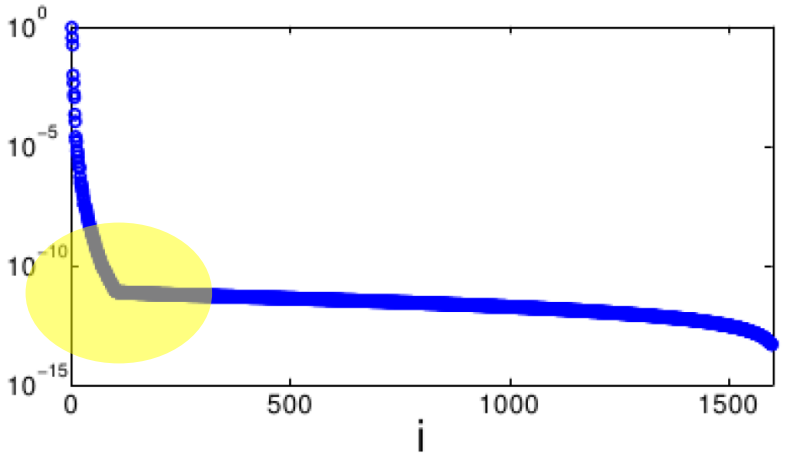
Material with controlled compliance



POD: Pain and Joy

1) Decay of Singular Values may be not fast – **<u>Cardiac conductivities!!!!</u>**



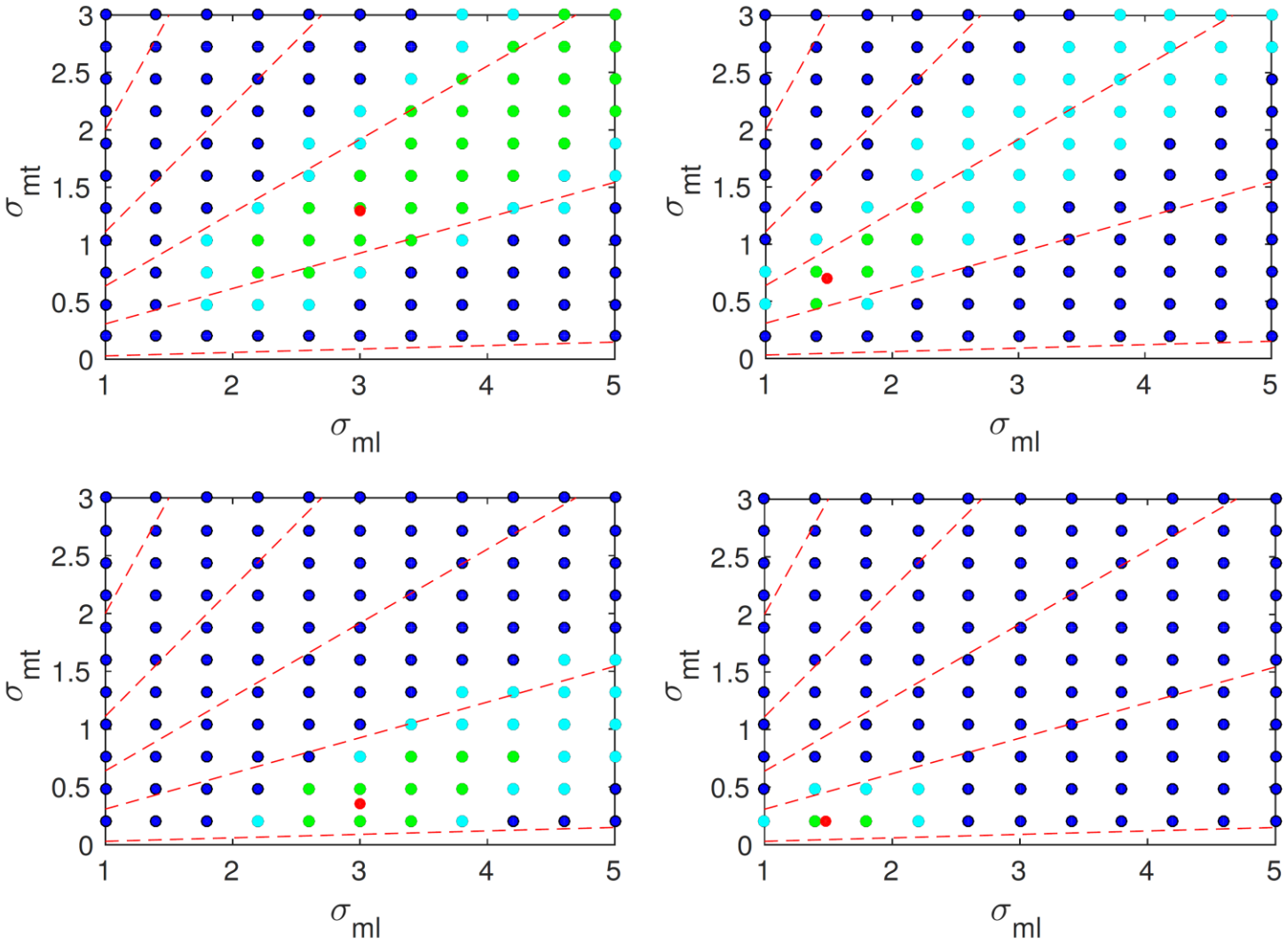


Left: SV for 1 parameter set of Monodomain problem (Cardiac conductivities)
Right: SV for Fluid-Structure Interaction problem (Young Modulus)

2) Nonlinear terms require special attention

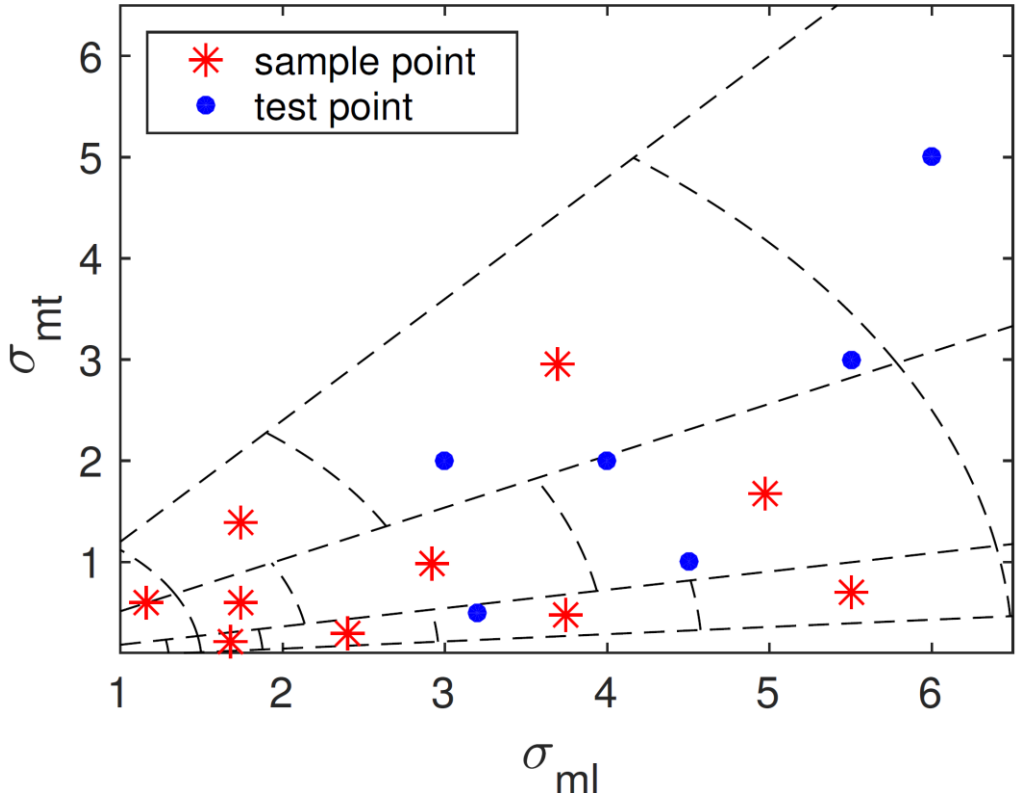
Discrete Empirical Interpolation Method

Sampling of the parameter space is crucial for POD (Monodomain)



Reconstructing a model from a parameter set (RED)
Errors:
GREEN: OK
CYAN: SO-AND-SO
BLUE: KO

CONCEPT: **DOMAIN OF EFFECTIVENESS** IN THE PARAMETER SPACE



With a judicious repartition of the parameter space by Domain of Effectiveness of the samples we can get good CPU reduction

	$\boldsymbol{\sigma}_{\text{exact}}$	$\boldsymbol{\sigma}_{\text{estimated}}$	# fwd | bwd	Total exe. time	Exe. time/solve
Full Order	[3.2, 0.5]	[3.237 0.625]	29 | 17	6211 s	135.02 s
POD+DEIM	[3.2, 0.5]	[3.161 0.393]	89 | 41	439.6 s	3.38 s
Full Order	[4.5, 1]	[4.535 1.099]	46 | 35	11160 s	137.78 s
POD+DEIM	[4.5, 1]	[4.596 0.974]	73 | 41	353.1 s	3.10 s
Full Order	[5.5, 3]	[5.528 3.064]	41 | 41	12780 s	155.85 s
POD+DEIM	[5.5, 3]	[5.384 3.053]	60 | 41	246.6 s	2.44 s
Full Order	[4, 2]	[4.045 2.076]	27 | 27	7525 s	144.71 s
POD+DEIM	[4, 2]	[3.764 1.968]	90 | 41	396.1 s	3.02 s
Full Order	[3, 2]	[3.054 2.077]	15 | 15	4199 s	139.97 s
POD+DEIM	[3, 2]	[2.490 2.434]	107 | 31	396.1 s	2.87 s
Full Order	[6, 5]	[5.854 5.121]	41 | 41	12870 s	156.95 s
POD+DEIM	[6, 5]	[4.820 5.836]	41 | 41	131.6 s	1.60 s