Using Bayes' theorem to infer the material parameters of human tissue

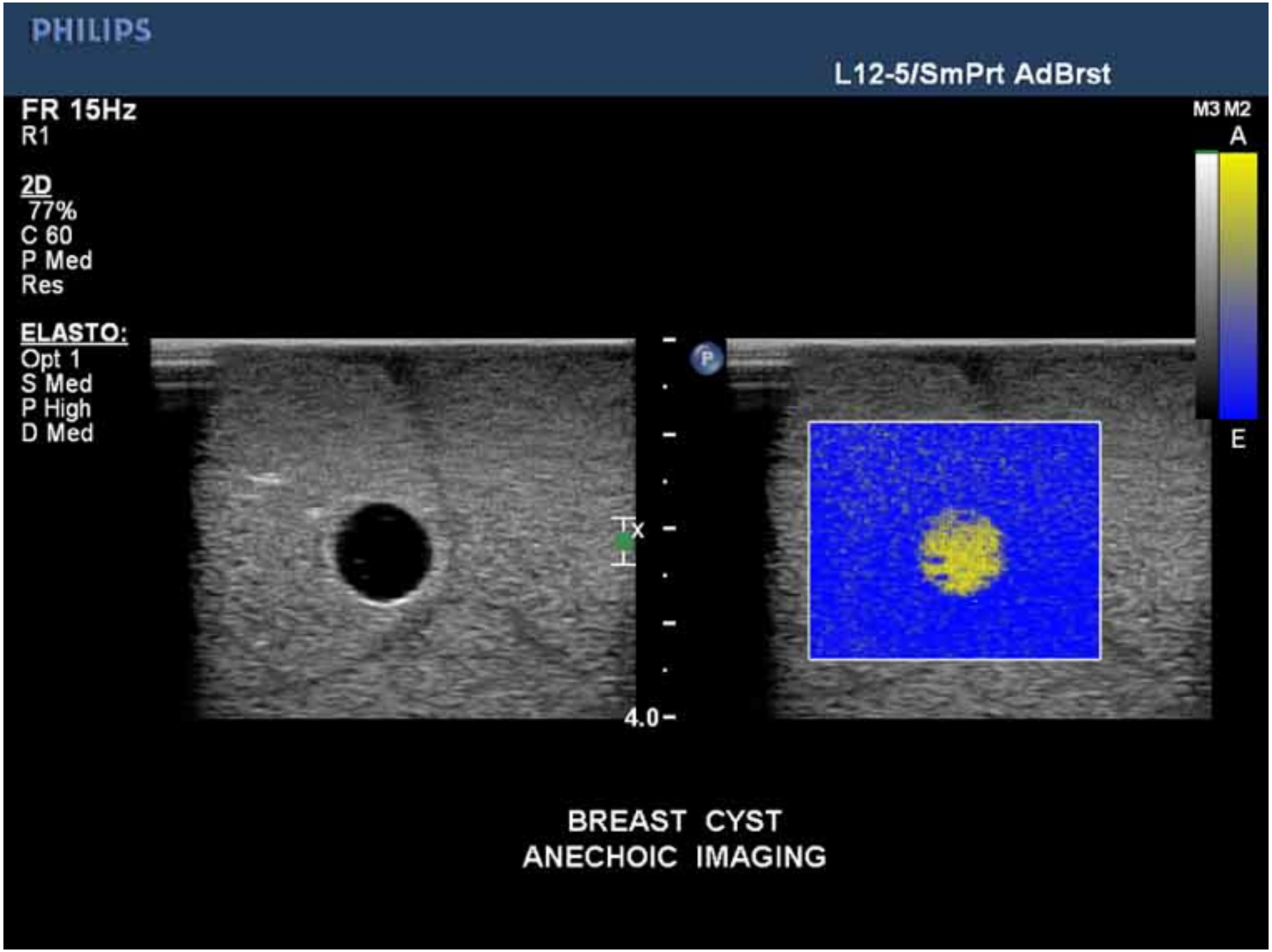
Jack S. Hale
Collaborators: Patrick E. Farrell, Stéphane P. A. Bordas

UNIVERSITÉ DU LUXEMBOURG
UNIVERSITY OF OXFORD
EPSRC Engineering and Physical Sciences Research Council
Fonds National de la Recherche Luxembourg
MARIE CURIE ACTIONS
European Research Council

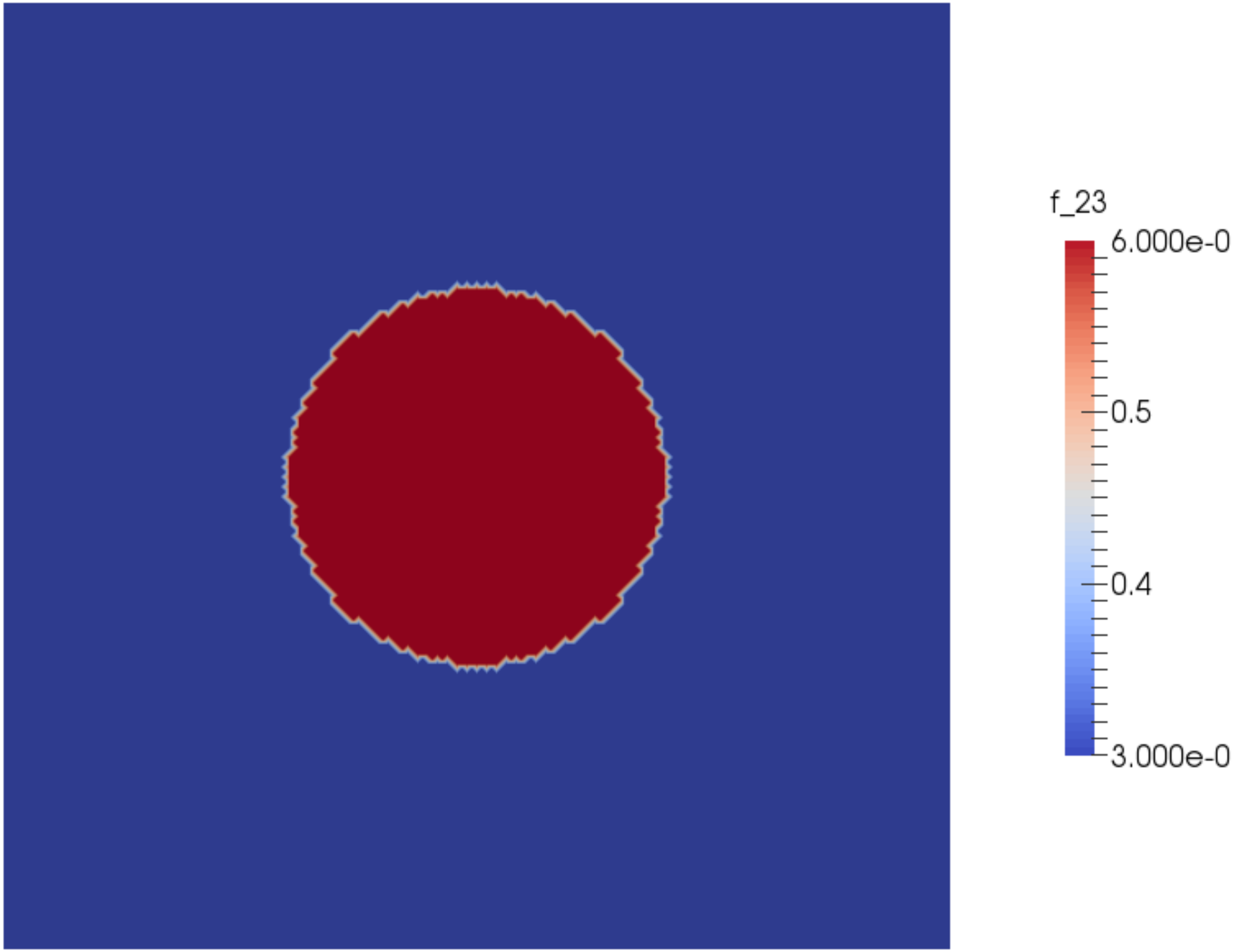
Overview

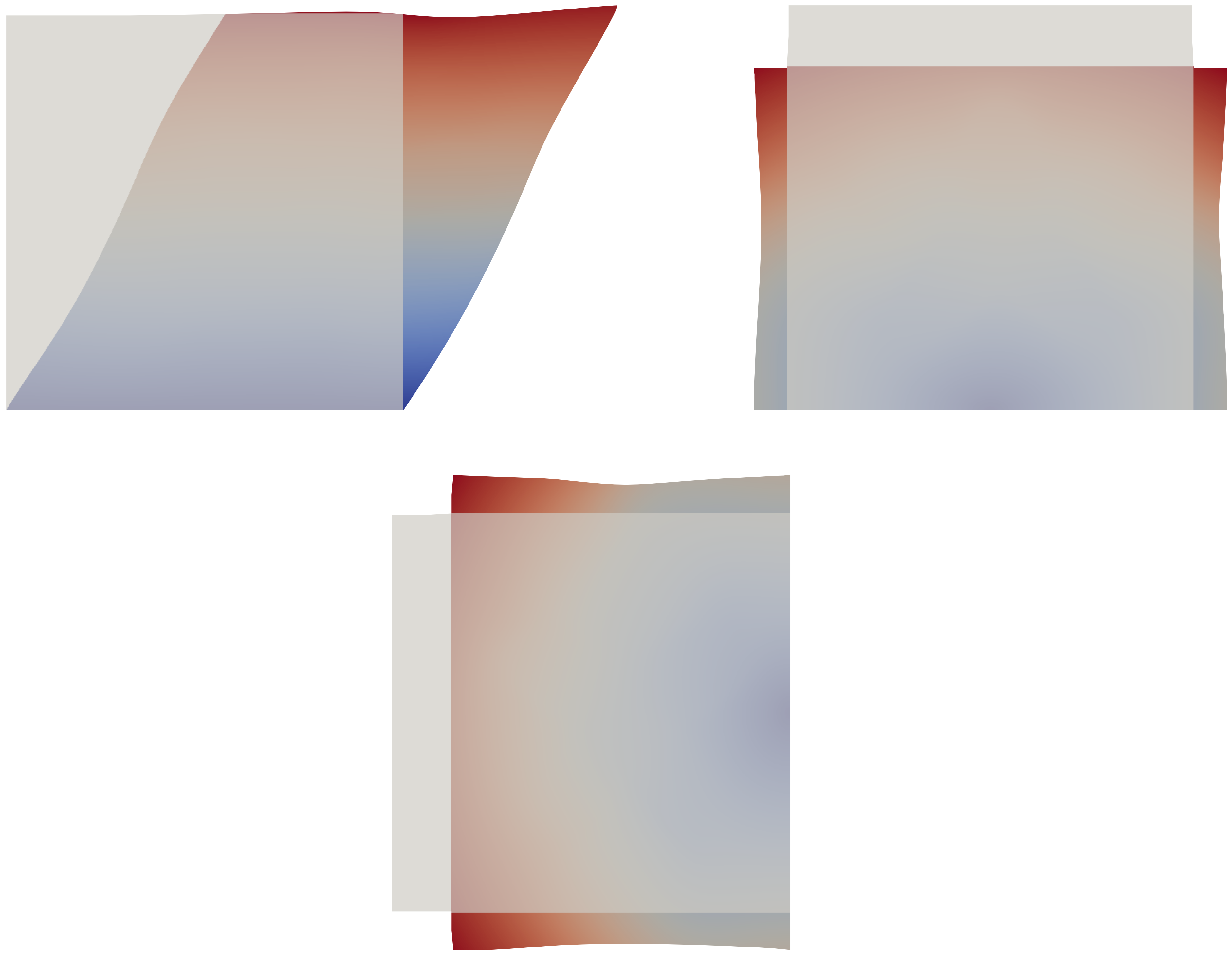
- Motivation.
- Bayesian approach to inversion.
- Relate classical optimisation techniques to the Bayesian inversion approach.
- Using a domain specific language for variational forms to solve the equations.
- Low-rank updates to deal with high-dimensional posterior covariance.
- Example problem: sparse surface observations of a solid block.

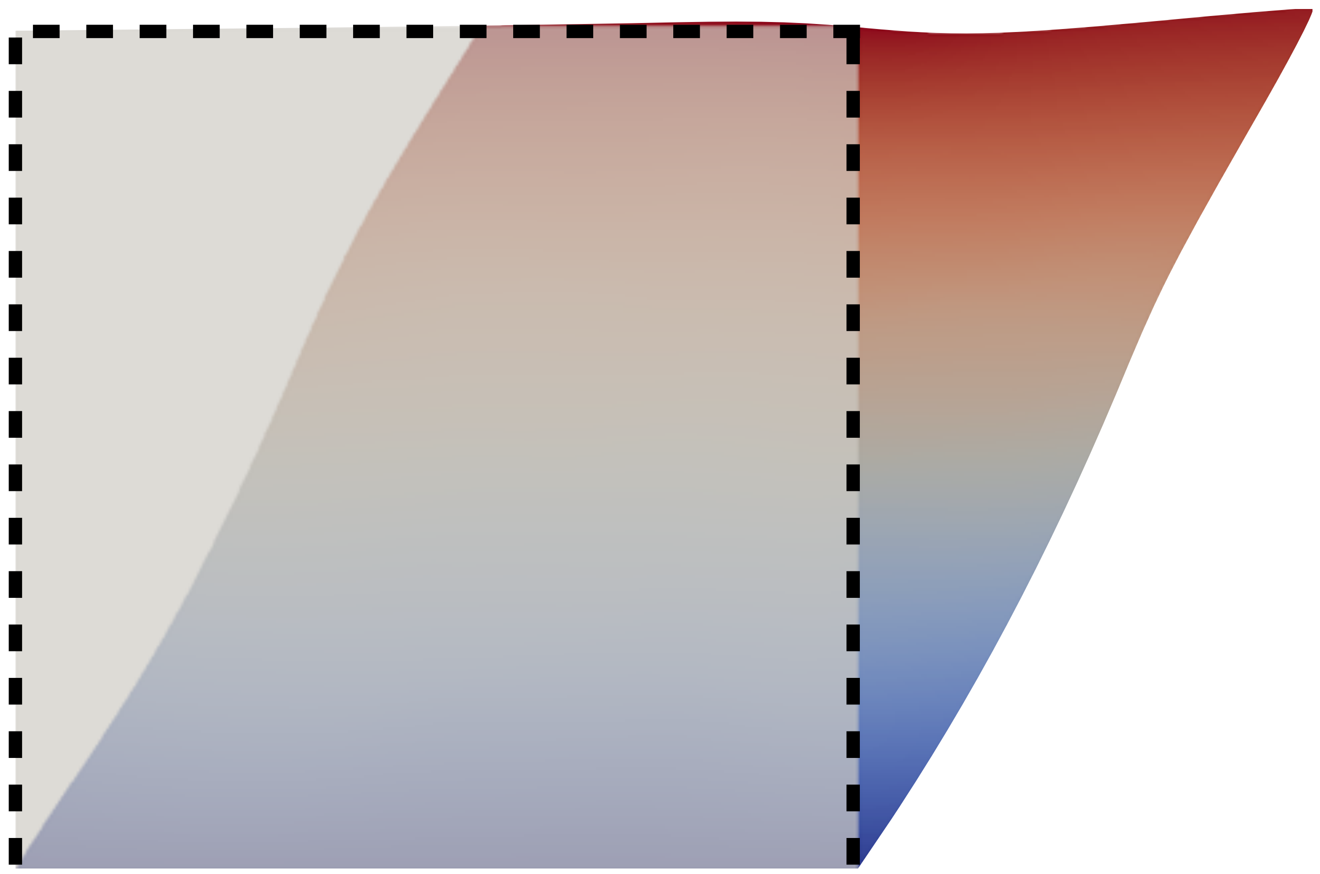
Motivation.



Source: Phillips


f_23
6.000e-01
0.5
0.4
3.000e-01

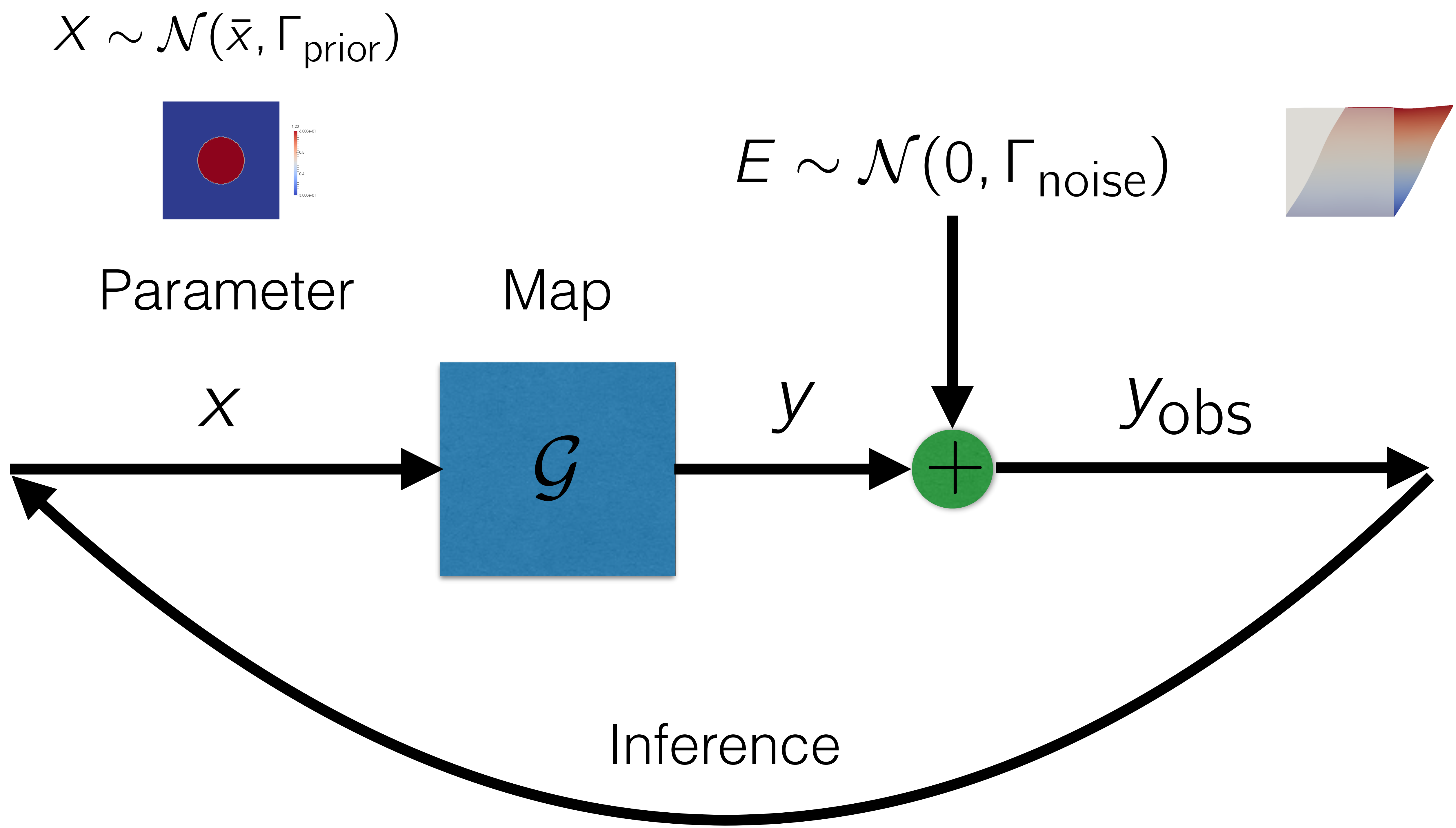




Q: What can we infer about the parameters inside the domain, just from displacement observations on the outside?

Q: Which parameters am I most uncertain about?

Framework



$$\pi_{\text{posterior}}(x \mid y) \propto \pi_{\text{likelihood}}(y \mid x)\pi_{\text{prior}}(x)$$

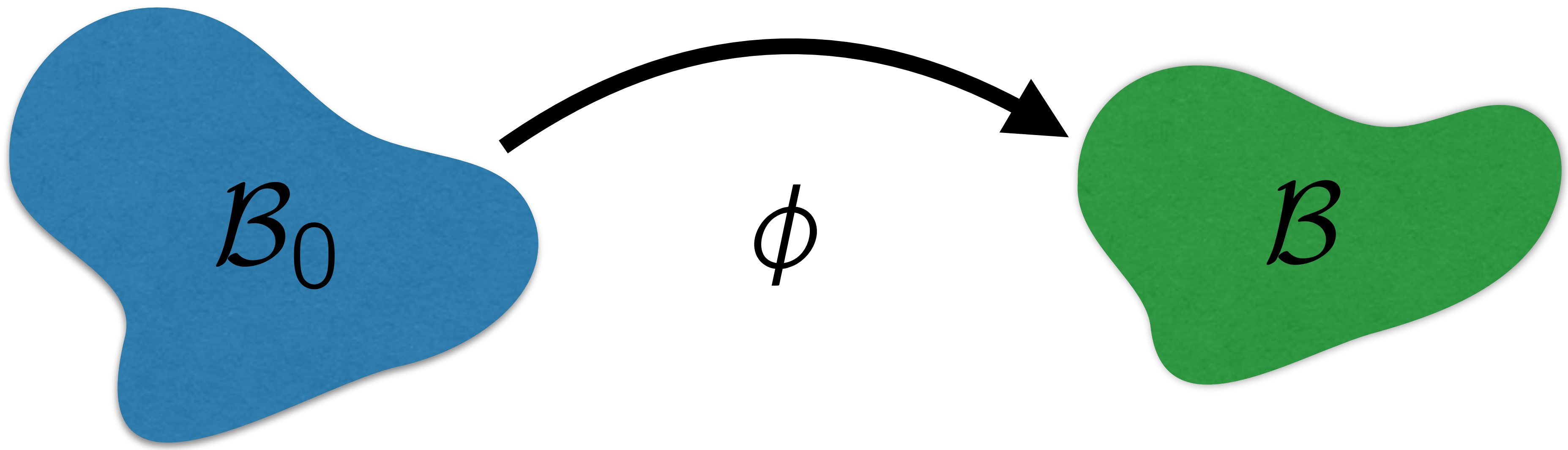
$$\pi_{\text{posterior}}(x|y) \propto \exp\left(-\frac{1}{2}||y - \mathcal{G}(u)||^2_{\Gamma^{-1}_{\text{noise}}} - \frac{1}{2}||x - \bar{x}||_{\Gamma^{-1}_{\text{prior}}}\right)$$

$$\mathcal{G} : \mathcal{X} \to \mathcal{Y}$$

Inverse Problems: A Bayesian perspective.

Stuart, Acta Numerica (2010).

Contribution: Bayesian inverse problems in an infinite-dimensional setting. *When is a Bayesian inverse problem well-posed?*



The displacements y for a given material parameter x are defined by a the minimum point of the following Lagrangian:

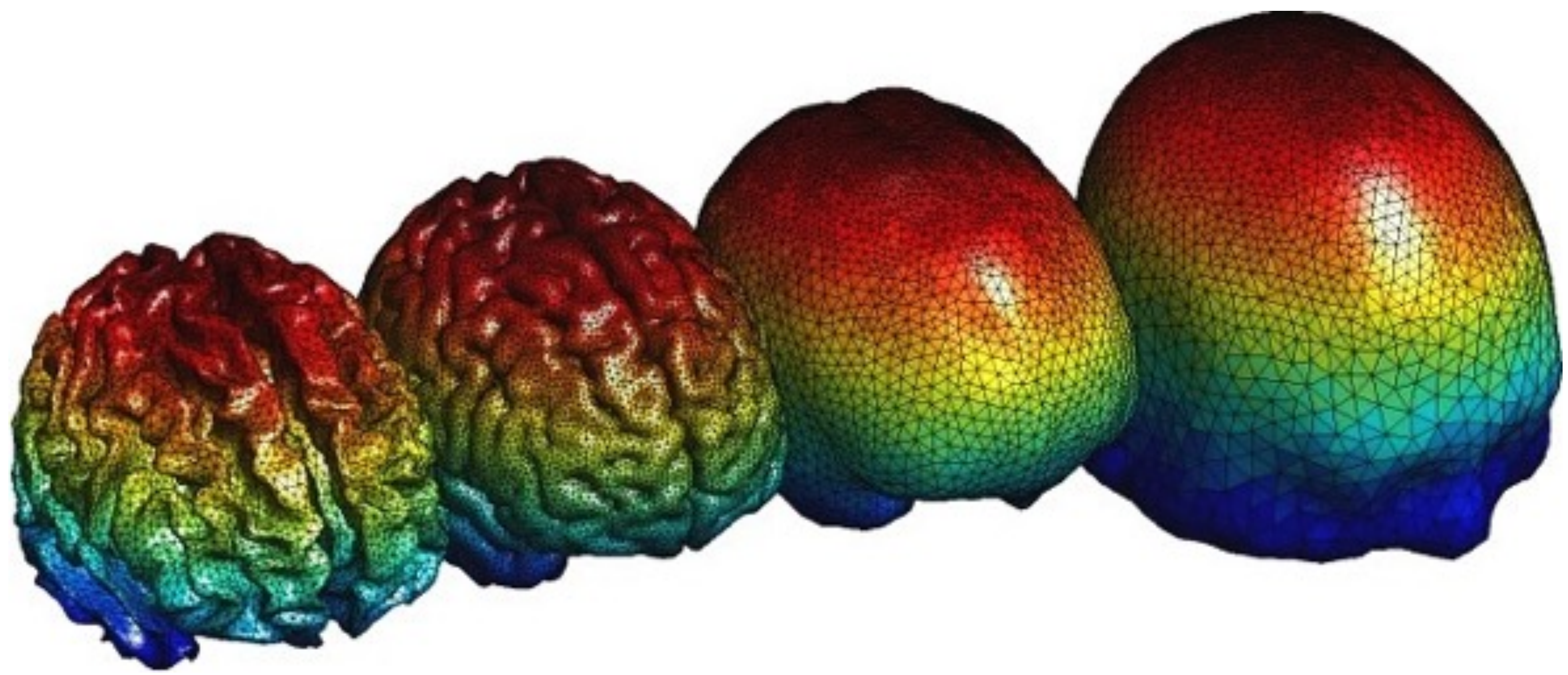
$$\mathcal{L}(y, x) = \int_{\Omega} \psi(y, x)\; dx - \int_{\Gamma} t \cdot y\; ds$$

where the energy density functional ψ is defined through the following equations:

$$\begin{aligned}
\psi(u, x) &= \frac{x}{2}(I_c - d) - x \ln(J) + \frac{\lambda}{2} \ln(J)^2, \\
\mathbf{F} &= \frac{\partial \phi}{\partial X} = \mathbf{I} + \nabla y, \\
\mathbf{C} &= \mathbf{F}^T \mathbf{F}, \\
\mathbf{I_C} &= \mathrm{tr}(C), \\
J &= \det \mathbf{F}.
\end{aligned}$$

Even once discretised (Finite Element Method)

$$\mathcal{G}_h : \mathbb{R}^n \to \mathbb{R}^m$$



Colin27 brain atlas

20% extension test, 16 Core Xeon, 1.12 million cells, ~29 secs.

$$n = 1,112,000$$

Problems

- Evaluating parameter-to-observable map is *very* expensive.
- Discretised parameter space can be *very* large.
- *Outcome:* Exploring posterior with ‘traditional sampling’ is not going to work.

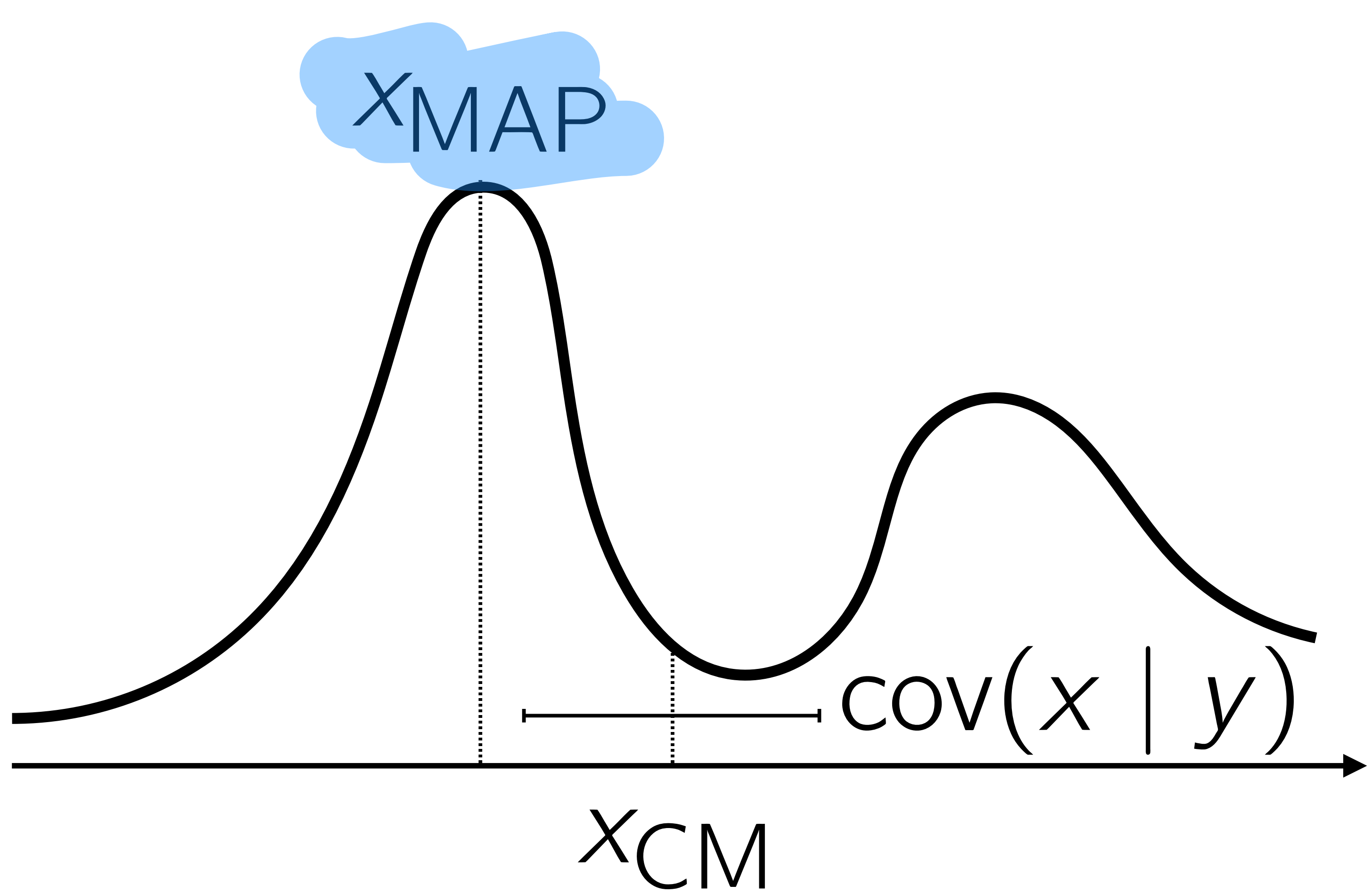
Solutions

1. Connect Bayesian approach to ideas from classical optimisation. Using derivatives of posterior in parameter-space (Girolami).

2. Exploiting low-rank structure of prior to posterior covariance updates (Flath 2012, Spantini 2015).

$$\pi_{\text{posterior}}(x \mid y) \propto \pi_{\text{likelihood}}(y \mid x)\pi_{\text{prior}}(x)$$

$$x_{\text{MAP}} = \arg\max_{x \in \mathbb{R}^n} \pi_{\text{posterior}}(x \mid y)$$



Linearise

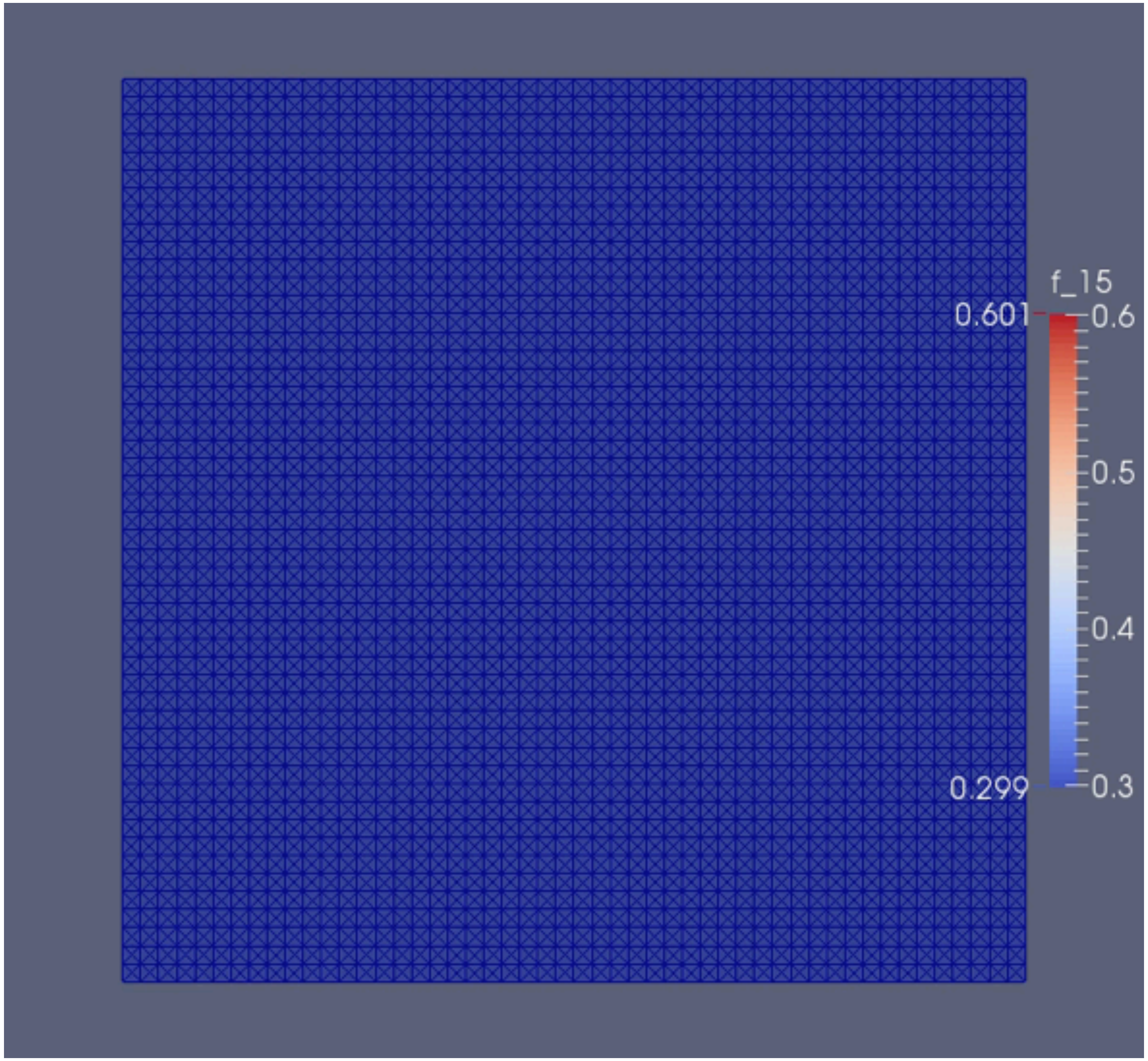
$$y = Ax$$

$$-\ln \pi_{\text{posterior}}(x|y) = \frac{1}{2}||y - Ax||^2_{\Gamma^{-1}_{\text{noise}}} + \frac{1}{2}||x - x_0||_{\Gamma^{-1}_{\text{prior}}}$$

Take the derivative

$$
\begin{aligned}
g(x_{\text{MAP}}) &:= \nabla_x \left(\frac{1}{2} \|y - \mathbf{A}x\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{noise}}} + \frac{1}{2} \|x - x_0\|^2_{\boldsymbol{\Gamma}^{-1}_{\text{prior}}} \right) \Bigg|_{x = x_{\text{map}}} \\
&= \mathbf{A}^T \boldsymbol{\Gamma}^{-1}_{\text{noise}} (y - \mathbf{A} x_{\text{map}}) + \boldsymbol{\Gamma}^{-1}_{\text{prior}} (x_{\text{map}} - x_0) \\
&= 0
\end{aligned}
$$

$$
x_{\text{MAP}} = \left(\boldsymbol{\Gamma}^{-1}_{\text{prior}} - \mathbf{A}^T \boldsymbol{\Gamma}^{-1}_{\text{noise}} \mathbf{A} \right)^{-1} (\mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}} y + \boldsymbol{\Gamma}_{\text{prior}} x_0)
$$



MAP estimate. Bound-constrained Quasi-Newton BLMVM with More-Thuente line search and ‘correct’ Riesz map.

and the second derivative…

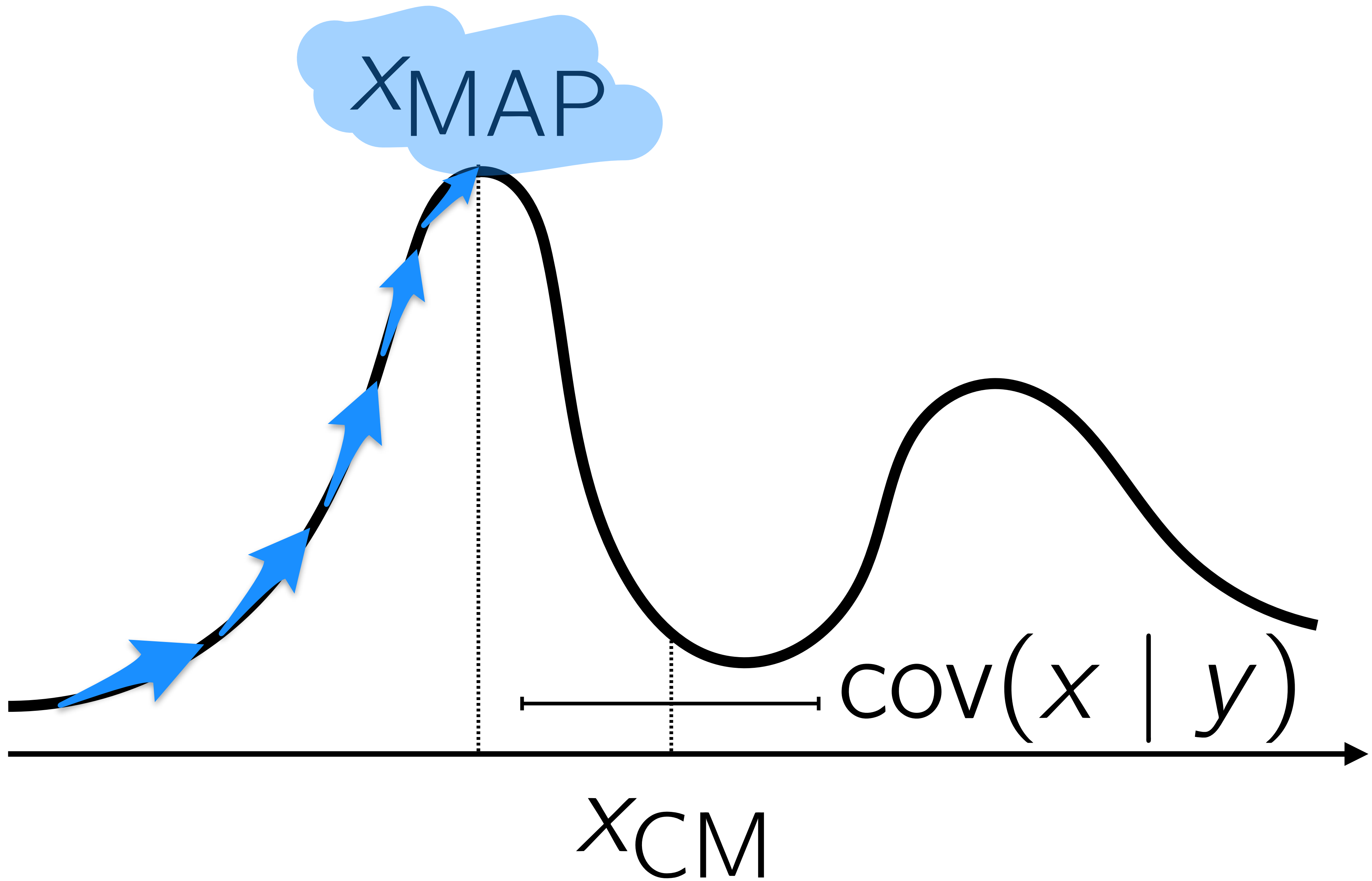
$$\mathbf{H} := \nabla_x g = \boldsymbol{\Gamma}_{\text{prior}}^{-1} - \mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$$

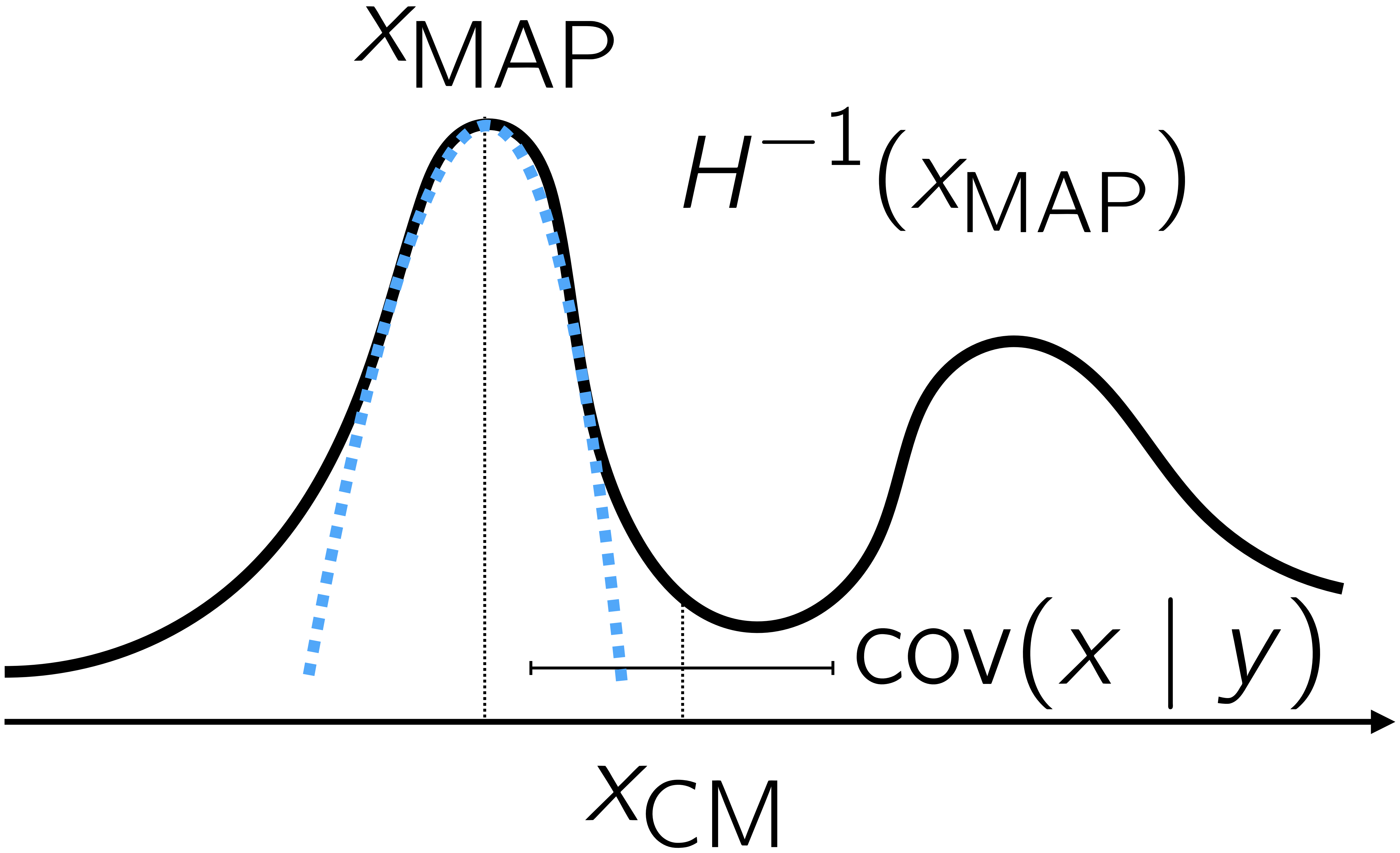
$$x_{\text{MAP}} = \left(\boldsymbol{\Gamma}_{\text{prior}}^{-1} - \mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{A}\right)^{-1} (\mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}} y + \boldsymbol{\Gamma}_{\text{prior}} x_0)$$

(After a fair bit of manipulation…)

$$\pi_{\text{posterior}} \sim \mathcal{N}(x_{\text{MAP}}, \mathbf{H}^{-1})$$

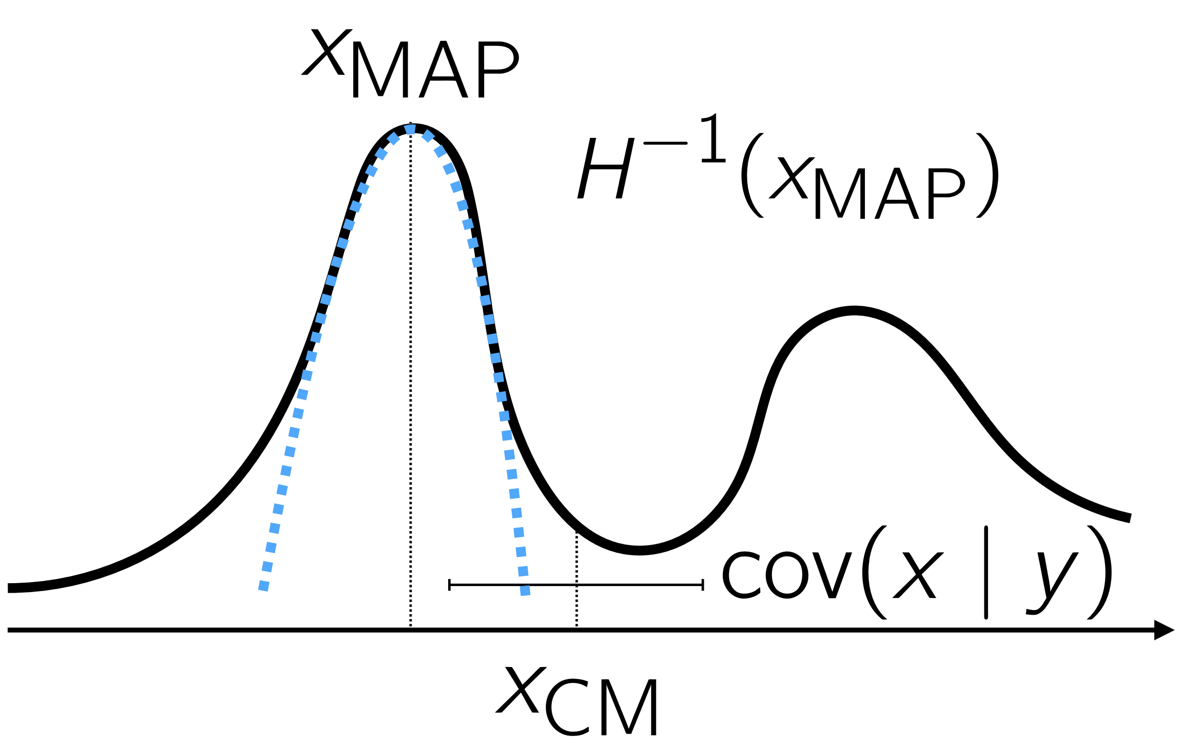
MAP estimate




x_{MAP}
$H^{-1}(x_{\mathrm{MAP}})$
$\mathrm{cov}(x \mid y)$
x_{CM}

$$\pi_{\text{posterior}} \sim \mathcal{N}(x_{\text{MAP}}, \mathbf{H}^{-1})$$

$$\pi^{\text{approx}}_{\text{posterior}} \sim \mathcal{N}(x_{\text{MAP}}, \mathbf{H}^{-1}(x_{\text{MAP}}))$$



Tools

- The FEniCS Project is collection of free software for the automated, efficient solution of differential equations using the finite element method.

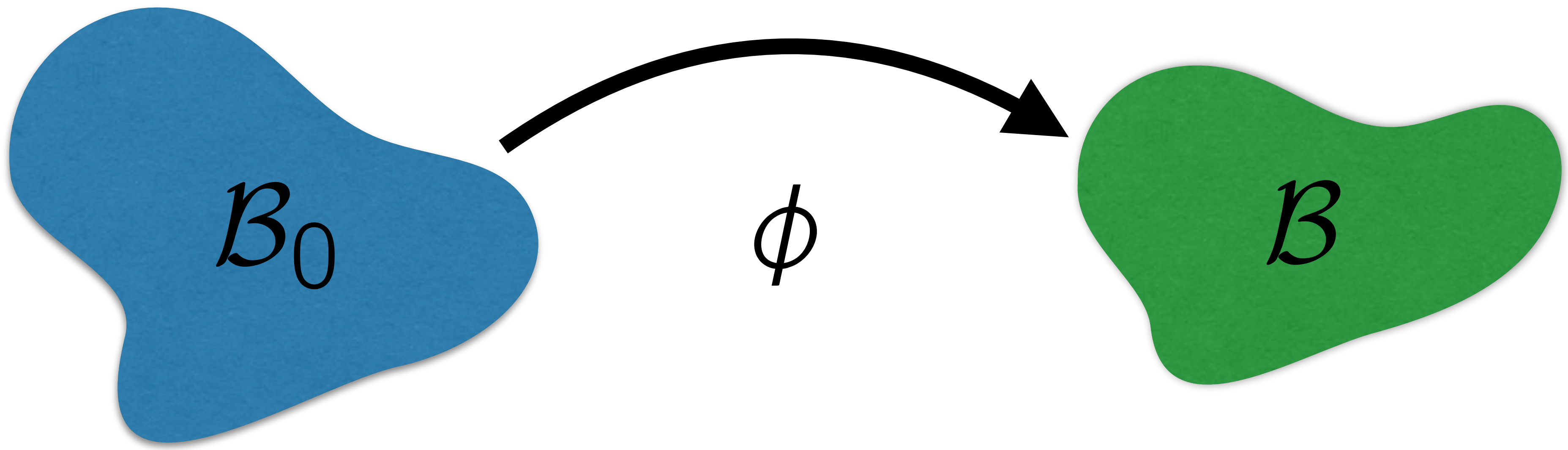
- dolfin-adjoint automatically derives the discrete adjoint, tangent linear and higher-order adjoint models from a high-level description of the forward model.

http://fenicsproject.org

Wells, Logg, Rognes, Kirby and many, many others…

http://www.dolfin-adjoint.org

Farrell, Funke, Ham and Rognes.
2015 Wilkinson Prize for Numerical Software.



The displacements y for a given material parameter x are defined by a the minimum point of the following Lagrangian:

$$\mathcal{L}(y, x) = \int_\Omega \psi(y, x)\, dx - \int_\Gamma t \cdot y\, ds$$

where the energy density functional ψ is defined through the following equations:

$$\begin{aligned}
\psi(u, x) &= \frac{x}{2}(I_c - d) - x \ln(J) + \frac{\lambda}{2} \ln(J)^2, \\
\mathbf{F} &= \frac{\partial \phi}{\partial X} = \mathbf{I} + \nabla y, \\
\mathbf{C} &= \mathbf{F}^T \mathbf{F}, \\
\mathbf{I_C} &= \mathrm{tr}(C), \\
J &= \det \mathbf{F}.
\end{aligned}$$

The displacements y for a given material parameter x are defined by a the minimum point of the following Lagrangian:

$$\mathcal{L}(y,x) = \int_\Omega \psi(y,x)\ dx - \int_\Gamma t \cdot y\ ds$$

where the energy density functional ψ is defined through the following equations:

$$\psi(u,x) = \frac{x}{2}(I_C - d) - x\ln(J) + \frac{\lambda}{2}\ln(J)^2,$$

$$\mathbf{F} = \frac{\partial\phi}{\partial X} = \mathbf{I} + \nabla y,$$

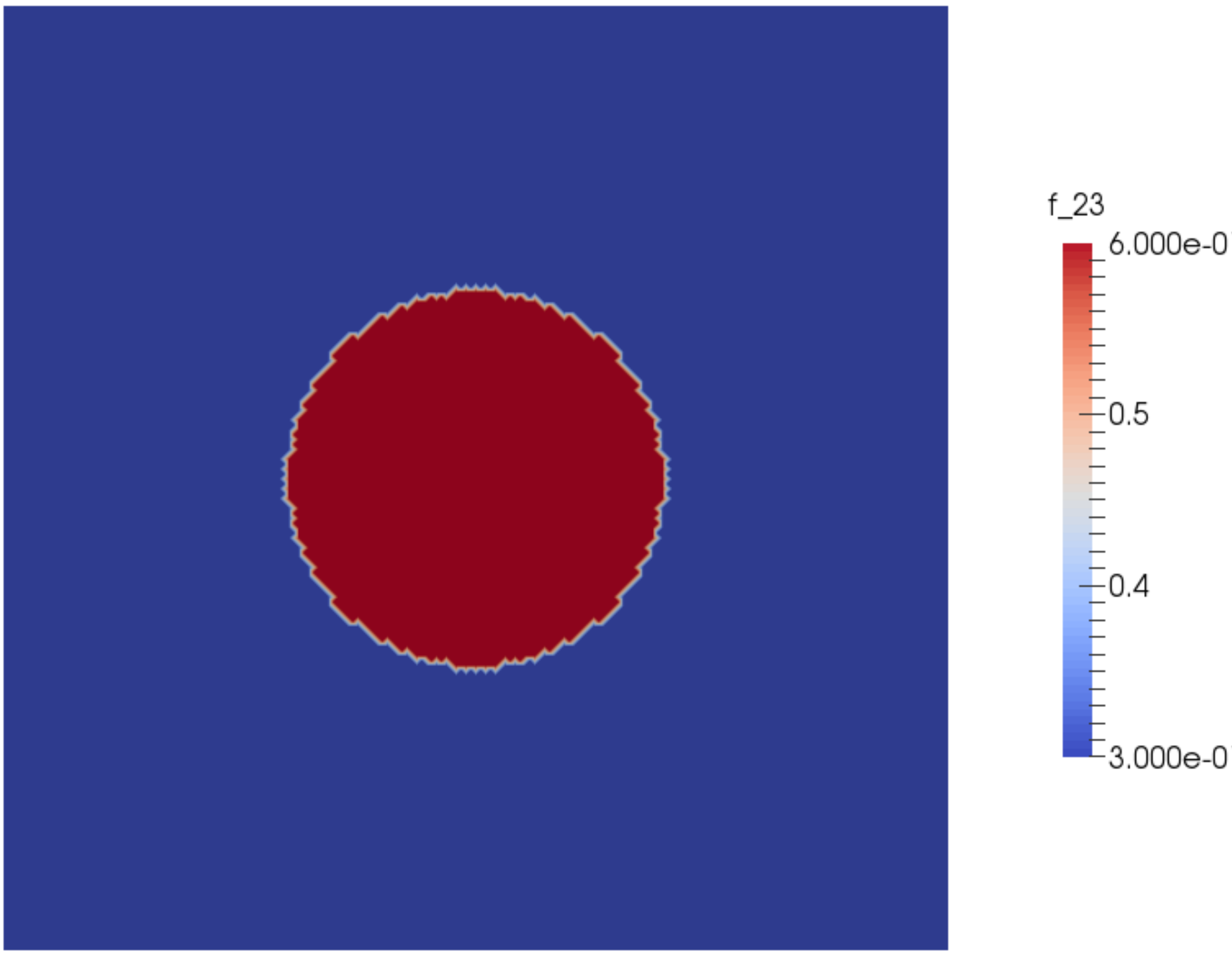
$$\mathbf{C} = \mathbf{F}^T\mathbf{F},$$

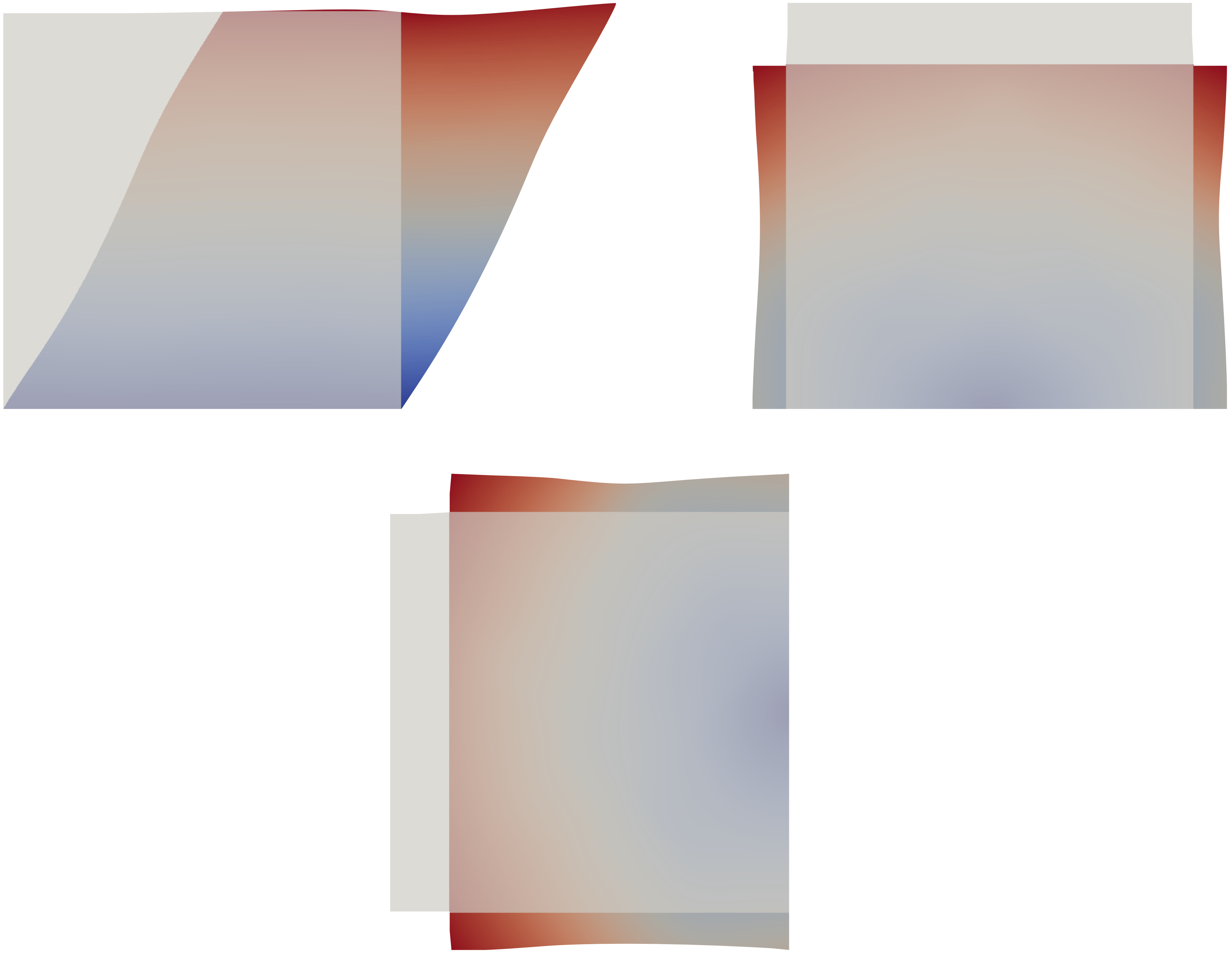
$$\mathbf{I_C} = \mathrm{tr}(C),$$

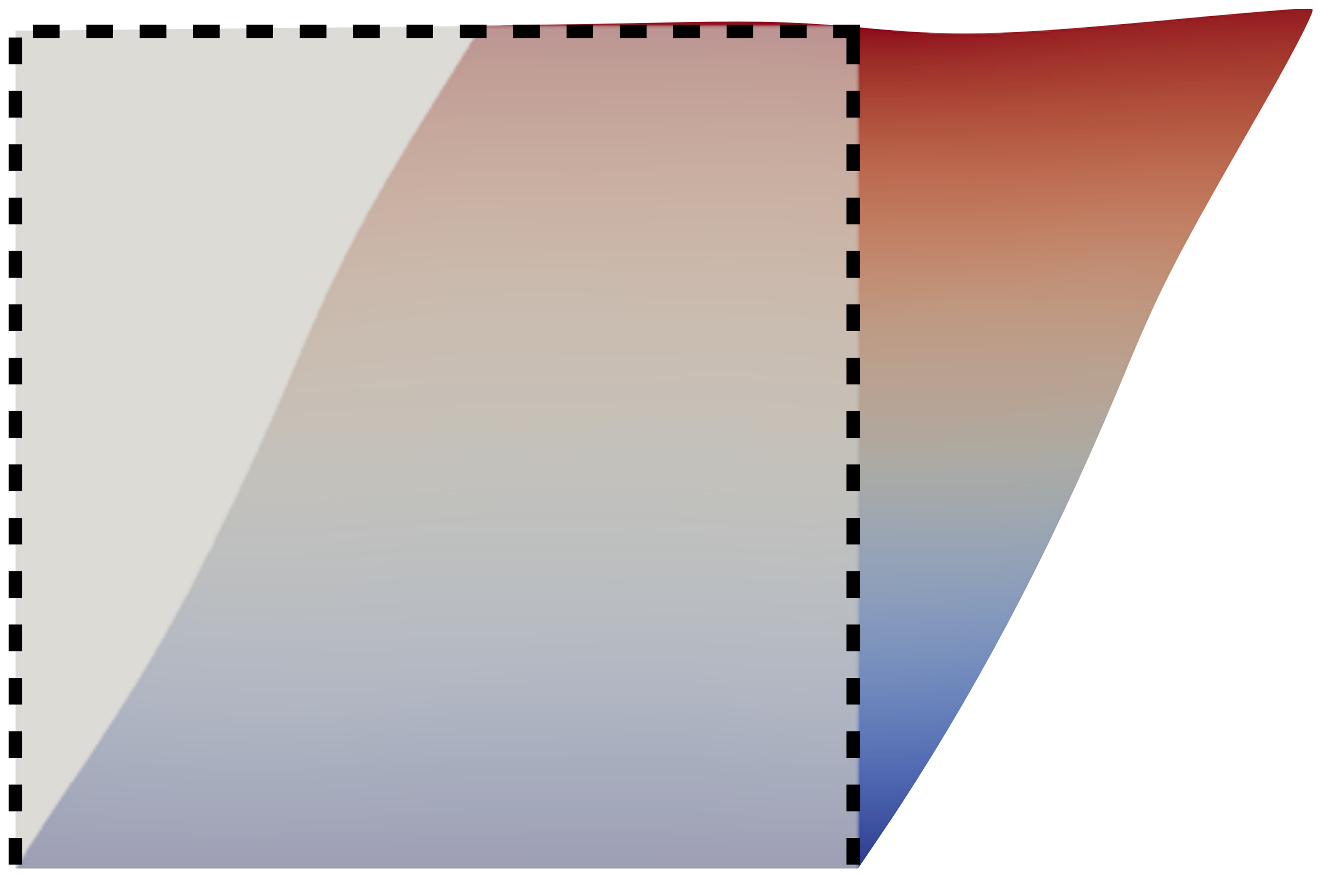
$$J = \det\mathbf{F}.$$

```
from dolfin import *
mesh = UnitSquareMesh(32, 32)

U = VectorFunctionSpace(mesh, "CG", 1)
V = FunctionSpace(mesh, "CG", 1)
# solution
u = Function(U)
# test functions
v = TestFunction(U)
# incremental solution
du = TrialFunction(U)
x = interpolate(Constant(1.0), V)
lmbda = interpolate(Constant(100.0), V)

dims = mesh.type().dim()
I = Identity(dims)
F = I + grad(u)
C = F.T*F
J = det(F)
Ic = tr(C)
```

```
phi = (x/2.0)*(Ic - dims) - x*ln(J) + (lmbda/
2.0)*(ln(J))**2
Pi = phi*dx
# gateux derivative with respect to u in direction v
F = derivative(Pi, u, v)
# and with respect to u in direction du
J = derivative(F, u, du)

u_h = Function(U)
F_h = replace(F, {u: u_h})
J_h = replace(J, {u: u_h})
solve(F_h == 0, u_h, bcs, J=J_h)
```

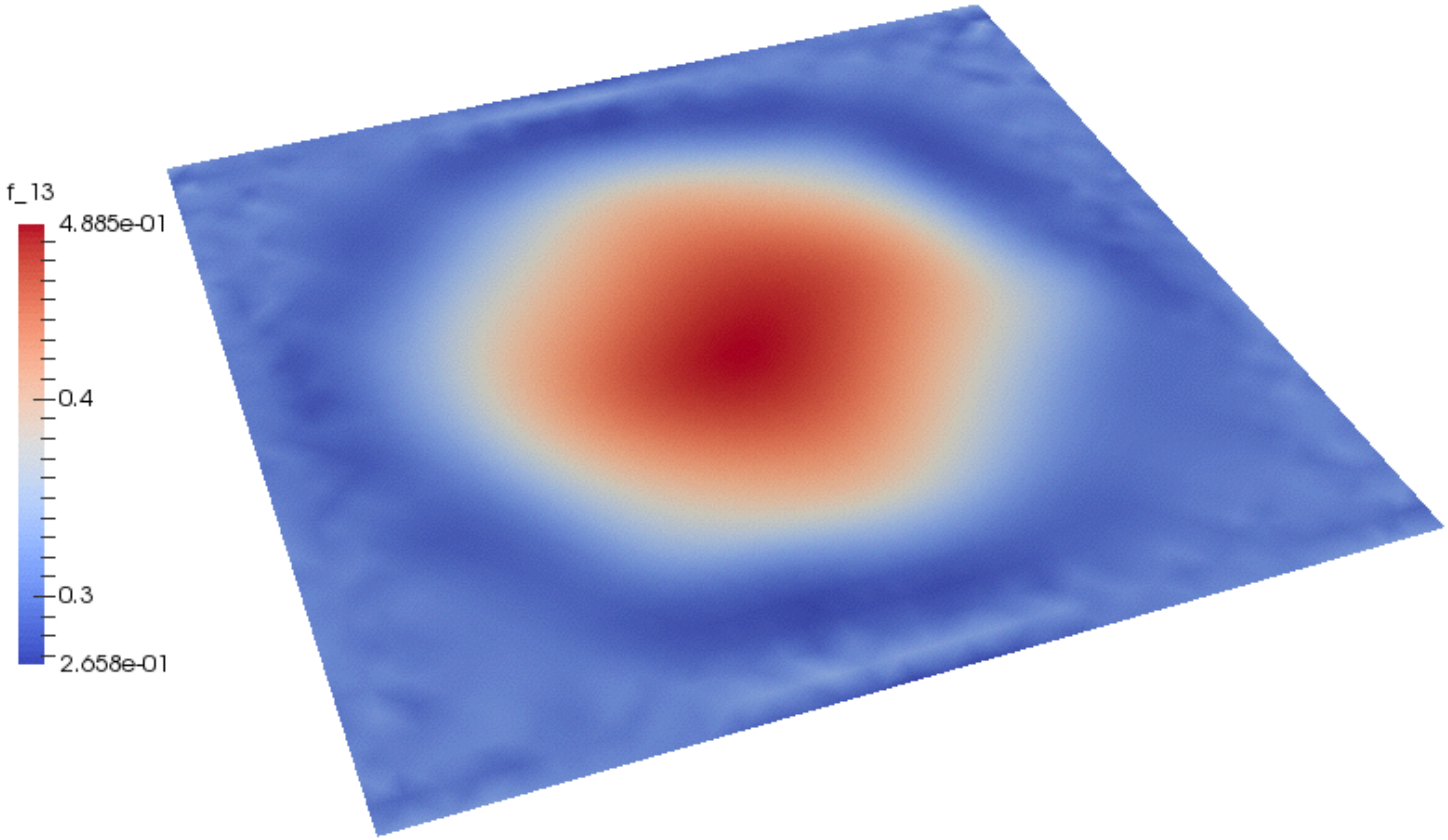

f_23
6.000e-01
0.5
0.4
3.000e-01





Q: What can we infer about the parameters inside the domain, just from displacement observations on the outside?

Q: Which parameters am I most uncertain about?



X_{map}

Optimal low-rank updates

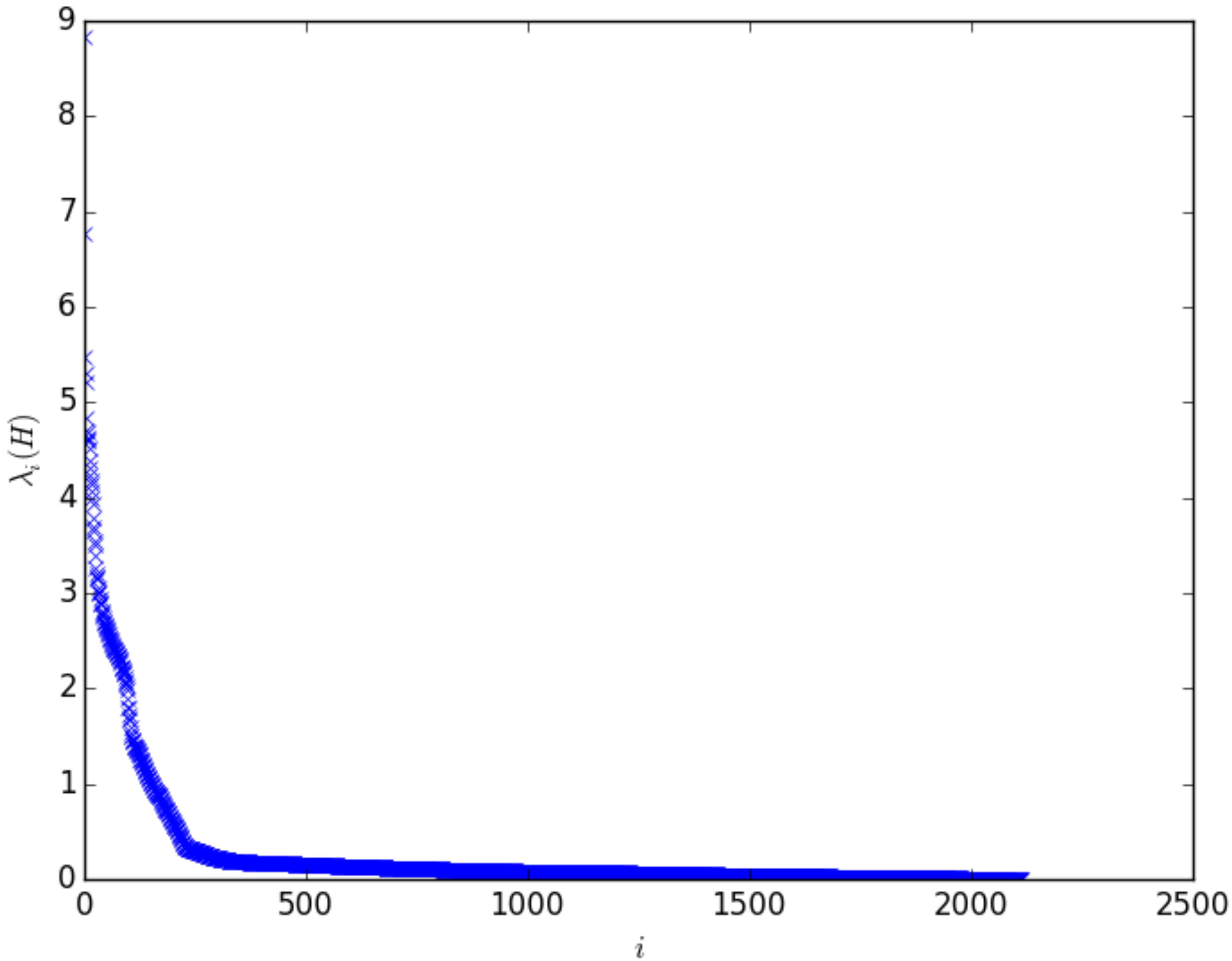
$$\Gamma_{\text{prior}} \xrightarrow{\quad y \quad} \Gamma_{\text{posterior}}$$

Critical idea: Observations are only informative on a low-dimensional subspace of the parameter space, *relative to the prior*. Spantini et al. (arXiV)

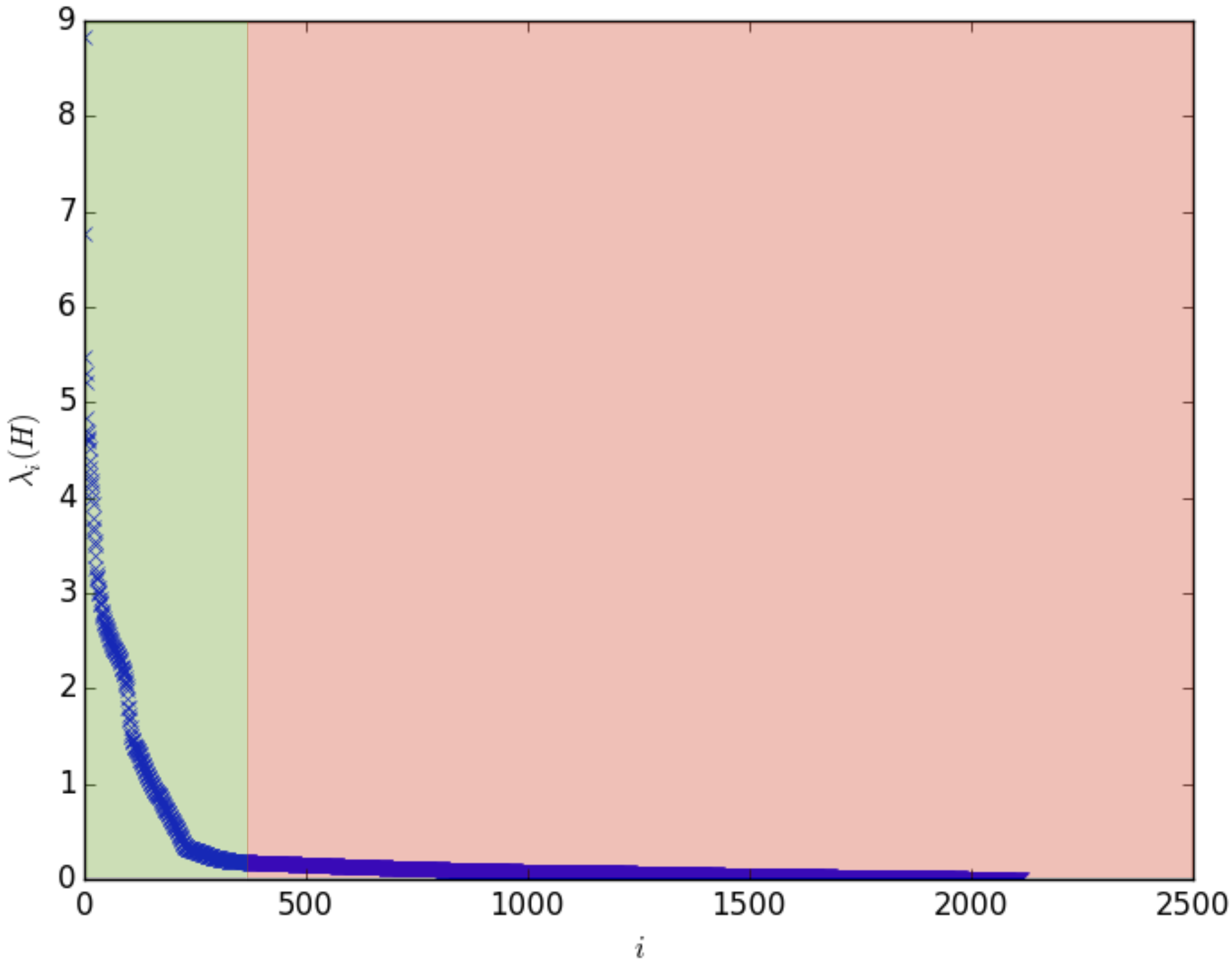
$$d^2_{\mathcal{F}}(A, B) = \sum_i \ln^2(\sigma_i)$$

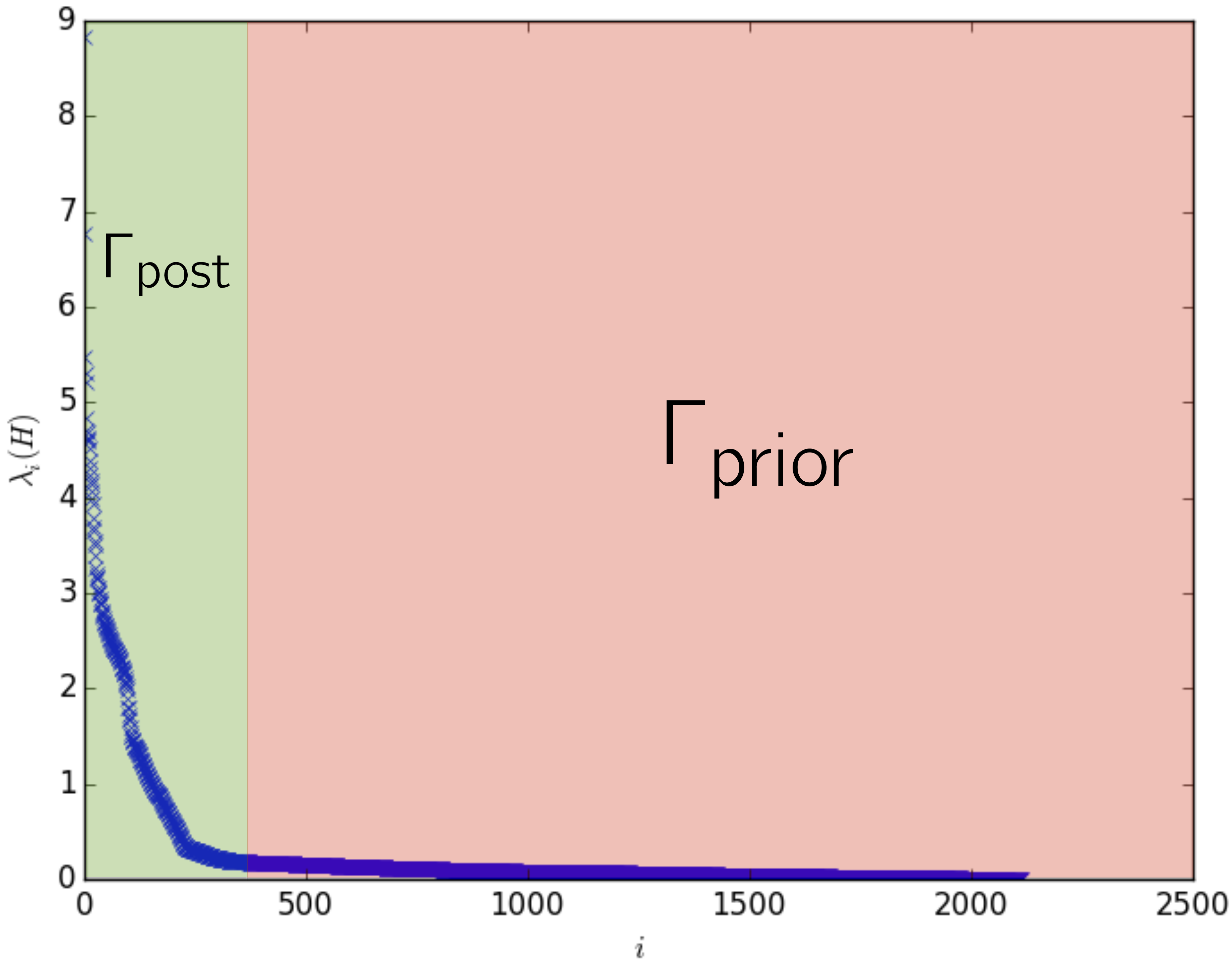
$$Aw_i = B\sigma_i w_i$$

Förstner metric



Matrix-free Krylov-Schur (SLEPc).

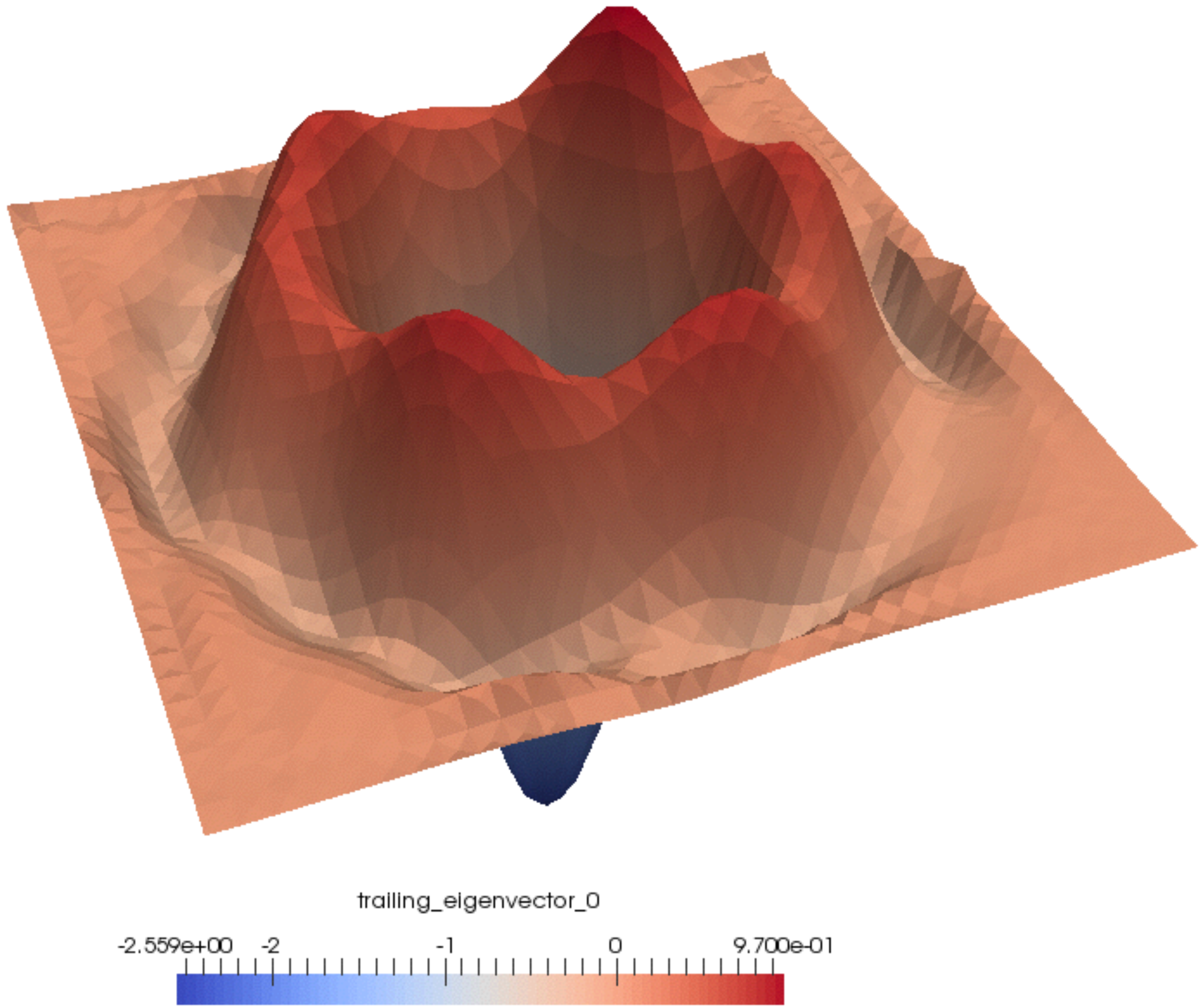

9
8
7
6
5
4
3
2
1
0
$\lambda_i(H)$
0
500
1000
1500
2000
2500
i

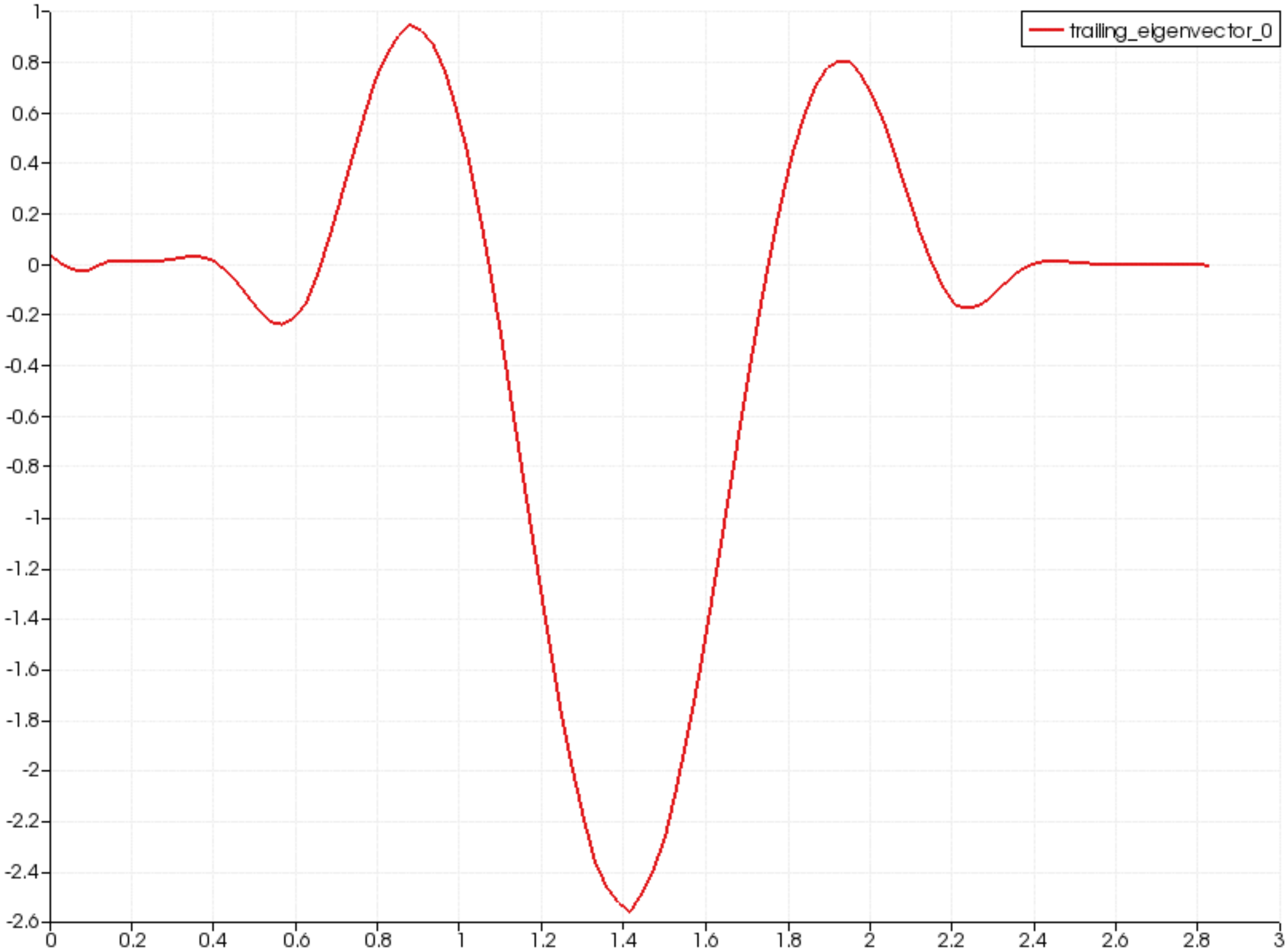


Matches trends from Flath et al. p424 for linear parameter to observable maps.

Trailing Eigenvector

Direction in parameter space *least* constrained by the observations


trailing_eigenvector_0
-2.559e+00
-2
-1
0
9.700e-01



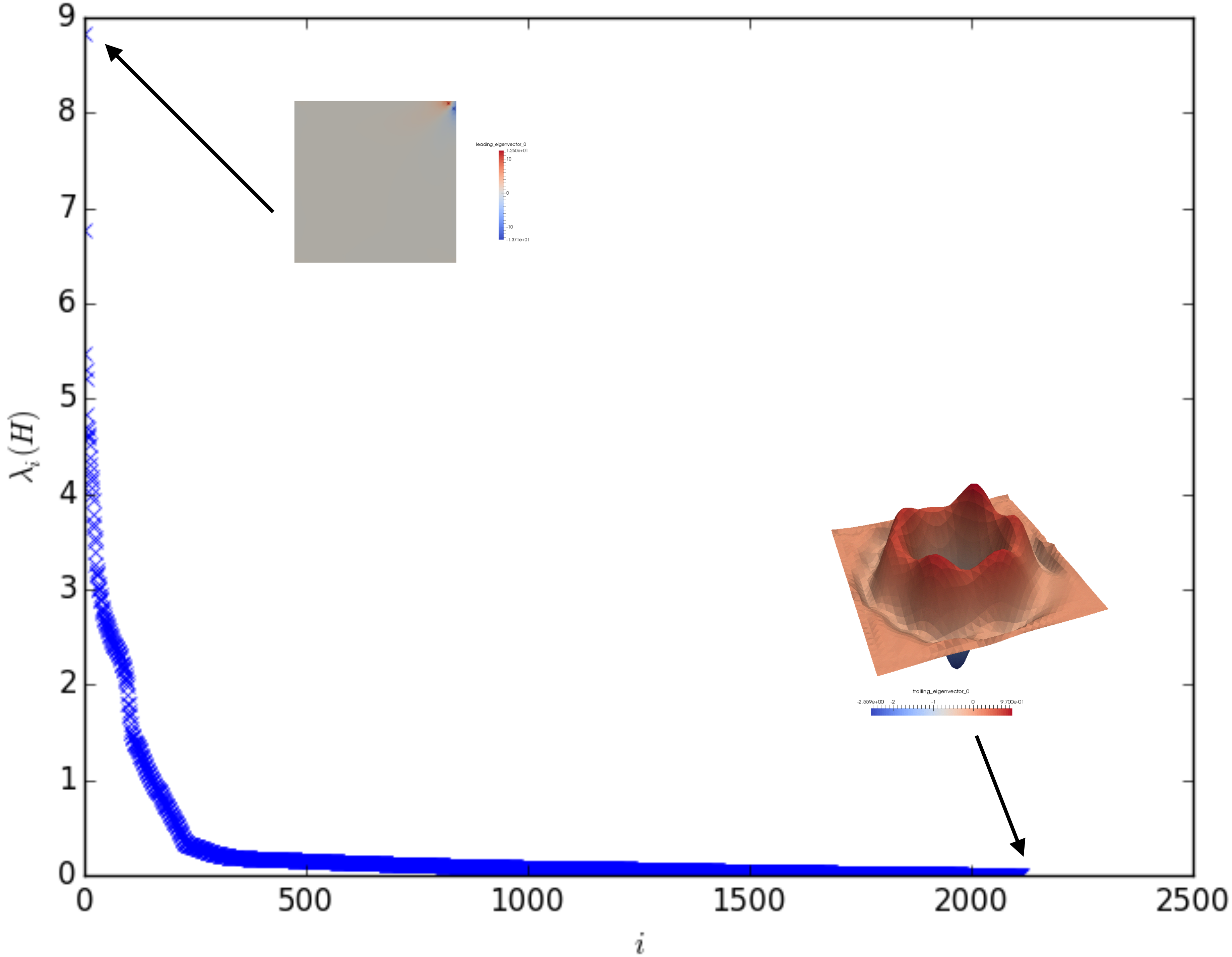
trailing_eigenvector_0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
-1.8
-2
-2.2
-2.4
-2.6
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3

Leading Eigenvectors

Direction in parameter space *most* constrained by the observations


leading_eigenvector_0
1.250e+01
10
0
-10
-1.371e+01


leading_eigenvector_1
1.405e+01
10
0
-10
-1.377e+01


9
8
7
6
5
4
3
2
1
0
0
500
1000
1500
2000
2500
$\lambda_i(H)$
i

Full Hessian.
4000+ actions.

Low-rank update.
292 actions.

Huge savings in computational cost.
Scales with model dimension because *observations* stay the same.

Summary

- We are developing methods to assess uncertainty in the recovered parameters in soft tissue.
- This is done within the framework of Bayesian inversion.
- FEniCS and dolfin-adjoint makes assembling the equations relatively easy, solving them is tougher!
- Next steps: exploring non-Gaussian nature of posterior using Hamiltonian MCMC. Requires derivatives.
- Paper soon on arXiv: infinite-dimensional setting, full derivations of equations, 3D problems.