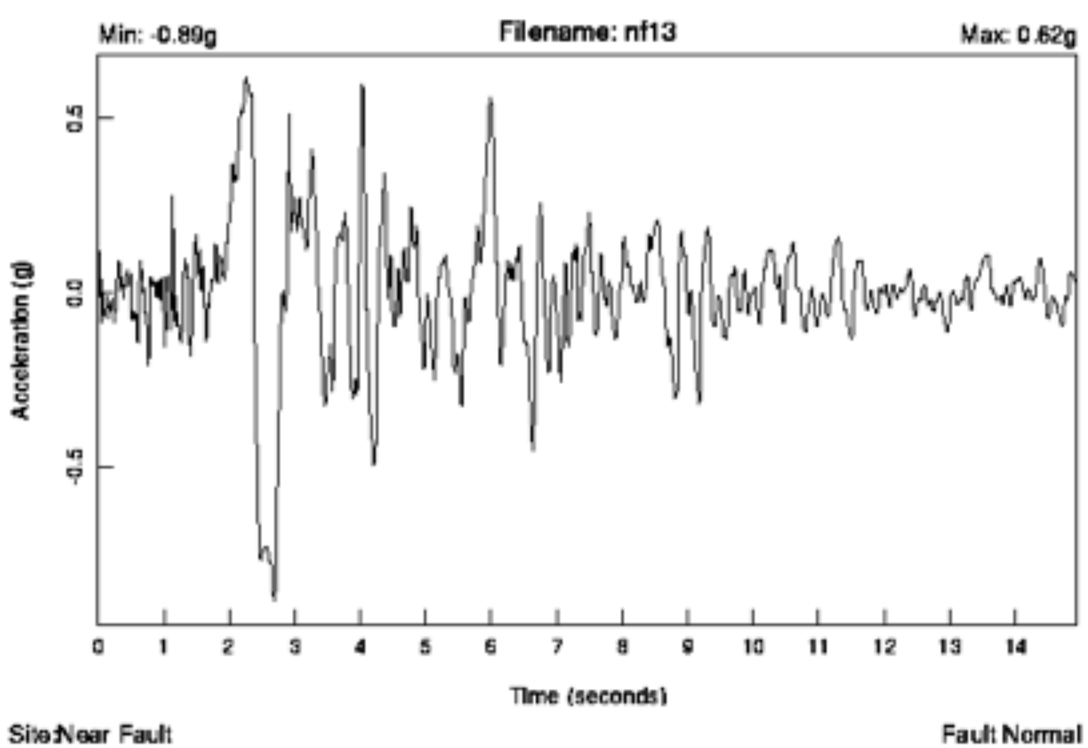
Lecture 12 - Dimensionality Reduction of Gaussian Random Fields

PREDICTIVE
SCIENCE LABORATORY

Objectives

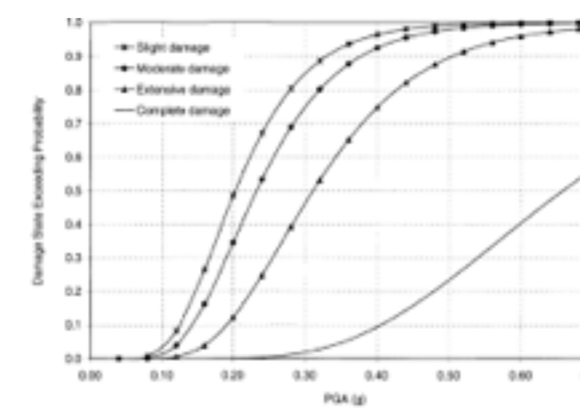
- Describe a Gaussian process with a finite number of uncorrelated random variables
- Use the reduced description to propagate uncertainties in functions through a model

Why do we need probability measures on functions?



Uncertainty in external forcing

Simulation



Fragility Curve

PREDICTIVE
SCIENCE LABORATORY

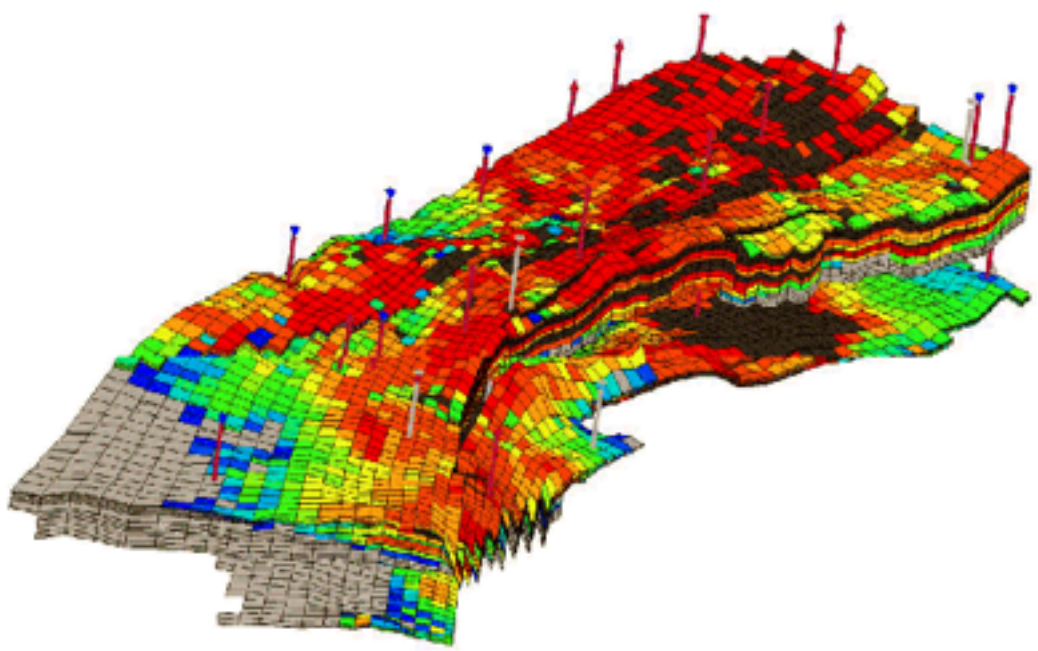
Why do we need probability measures on functions?

Portfolio Risk

Uncertainty in external forcing

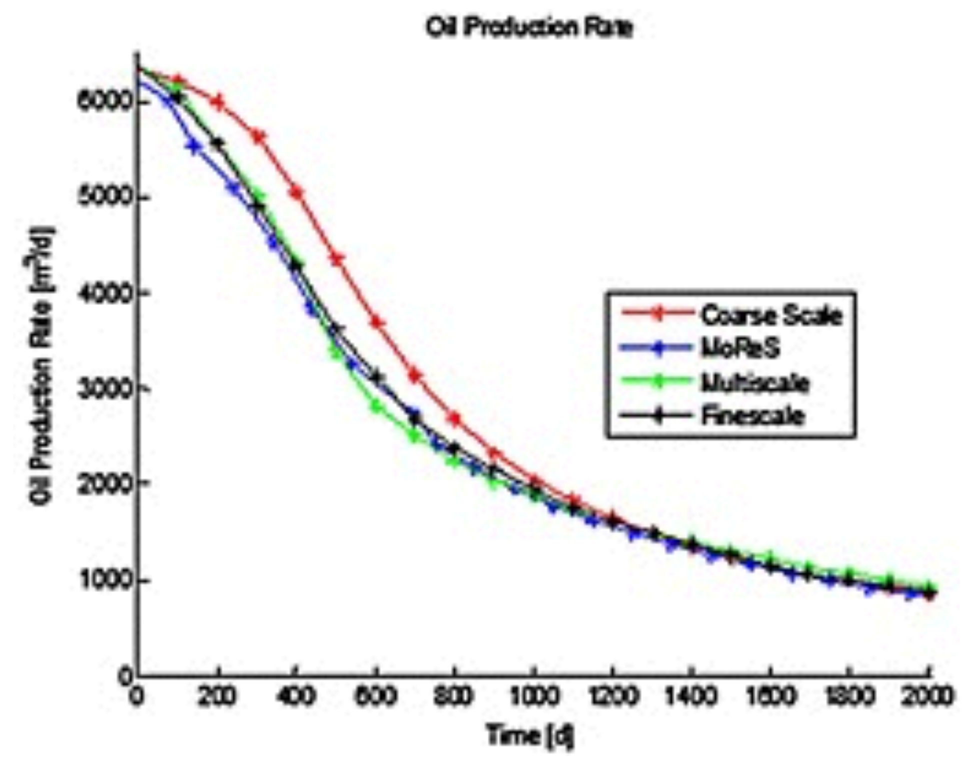
PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?



Uncertainty in field parameters

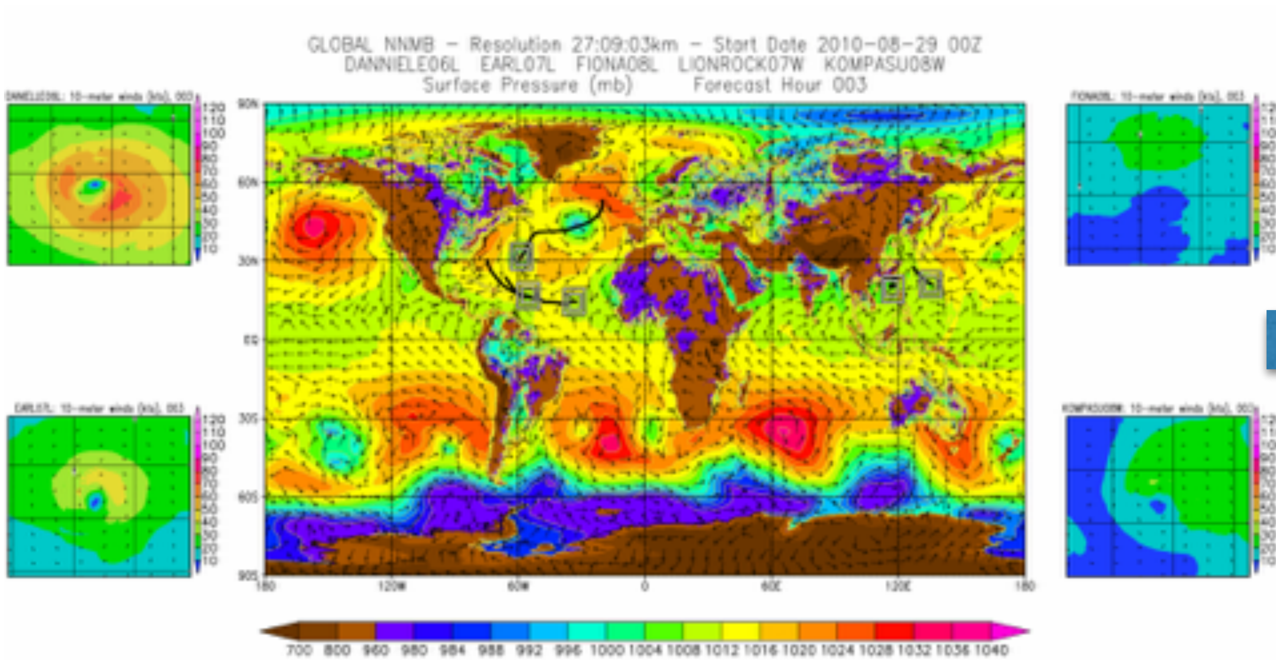
Simulation

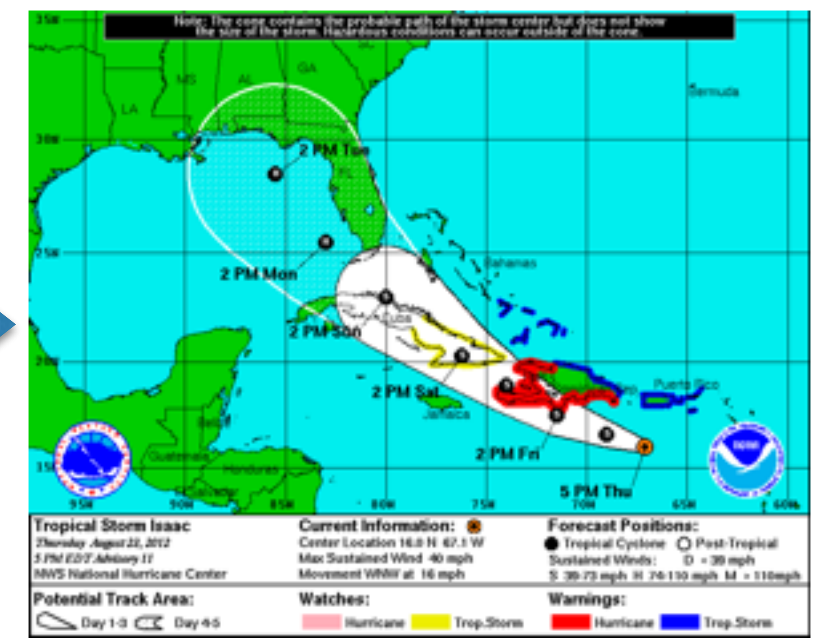


Oil produced over time

PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?





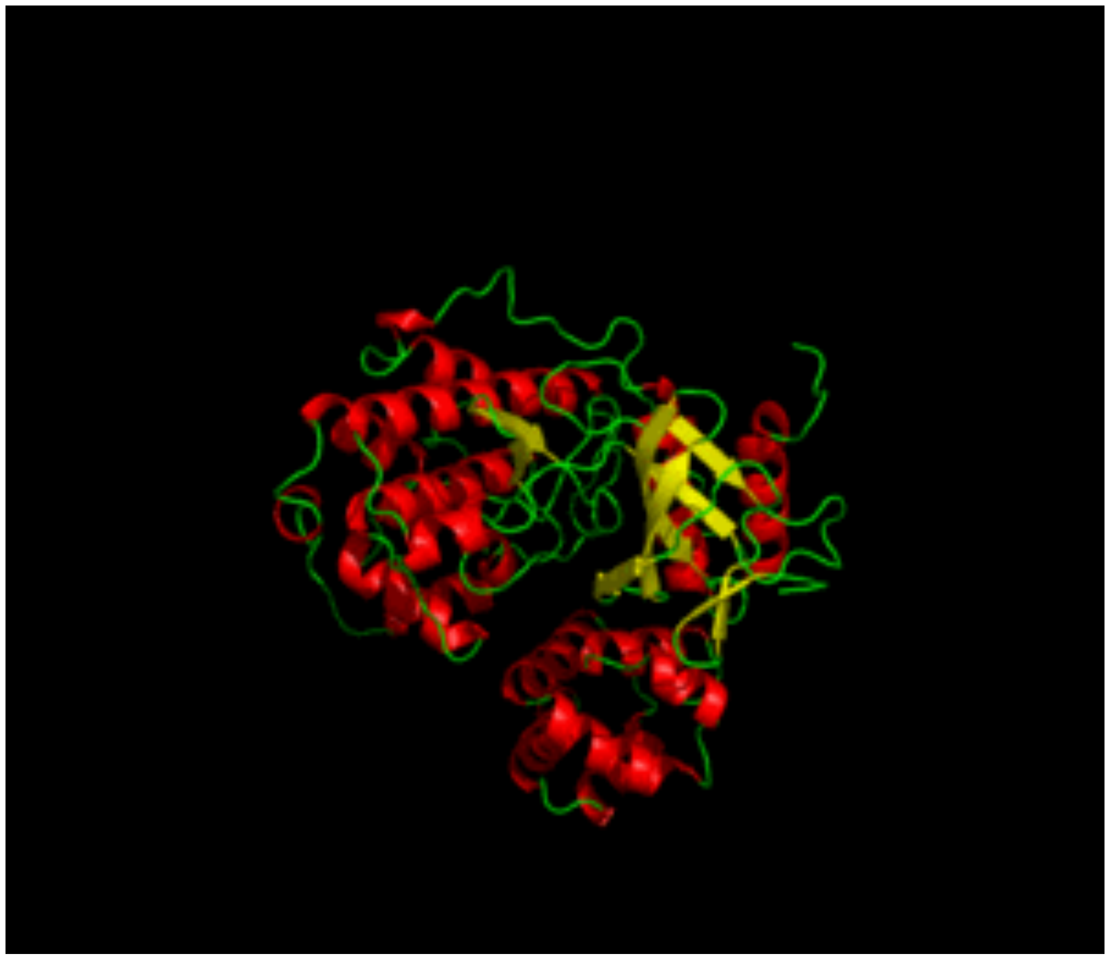
Hurricane path

Uncertainty in initial conditions

http://hwrf.aoml.noaa.gov/

PREDICTIVE
SCIENCE LABORATORY

Why do we need probability measures on functions?



Simulation of the interaction of two biomolecules

Statistical mechanics:

$$p(\mathbf{q}) \propto \exp\left\{-\frac{V(\mathbf{q})}{k_B T}\right\}$$

Positions of all atoms

Boltzmann constant

Temperature

Empirical potential. We are not exactly sure about its form…

PREDICTIVE
SCIENCE LABORATORY

For quantifying uncertainties in:

- External forcing
- Field parameters
- Initial conditions in PDEs
- Boundary conditions in PDEs
- Physical laws (e.g., constitutive relations, empirical force fields, etc.)
- …

PREDICTIVE
SCIENCE LABORATORY

Model Stochastic PDE Problem

Consider the elliptic partial differential equation:

$$\nabla \left(\alpha(\mathbf{x}) \nabla u(\mathbf{x}) \right) = \rho(\mathbf{x})$$

with boundary conditions:

$$u(\mathbf{x}) = g_D(\mathbf{x}), \text{ on } \mathbf{x} \in \Gamma_D$$

$$\mathbf{n} \cdot \nabla u(\mathbf{x}) = q_N(\mathbf{x}), \text{ on } \mathbf{x} \in \Gamma_N$$

Any of these could be uncertain…

PREDICTIVE
SCIENCE LABORATORY

Model Stochastic PDE Problem

Consider the elliptic partial differential equation:

$$\nabla \left(\alpha(\mathbf{x})\nabla u(\mathbf{x})\right) = \rho(\mathbf{x})$$

with boundary conditions:

$$u(\mathbf{x}) = g_D(\mathbf{x}),\ \text{on}\ \mathbf{x} \in \Gamma_D$$

$$\mathbf{n} \cdot \nabla u(\mathbf{x}) = g_N(\mathbf{x}),\ \text{on}\ \mathbf{x} \in \Gamma_N$$

Physical examples?

- Electrostatics
- Steady state heat
- Subsurface flow
- …

PREDICTIVE
SCIENCE LABORATORY

Model Stochastic PDE Problem

Consider the elliptic partial differential equation:

$$\nabla\left(\alpha(\mathbf{x})\nabla u(\mathbf{x})\right) = \rho(\mathbf{x})$$

with boundary conditions:

$$u(\mathbf{x}) = g_D(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_D$$

$$\mathbf{n}\cdot\nabla u(\mathbf{x}) = g_N(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_N$$

Since it is positive:

$$\alpha(\mathbf{x}) = \exp\left\{f(\mathbf{x})\right\}$$

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

PREDICTIVE
SCIENCE LABORATORY

Gaussian process

We are uncertain about a function:

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

Anything we know (potentially observed input/output data)

Could be prior or posterior mean/covariance

PREDICTIVE
SCIENCE LABORATORY

Sampling from a GP

Pick a bunch of test points:

$$\mathbf{x}_{1:n} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$$

Look at the function values at these points:

$$\mathbf{f}_{1:n} = \{f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)\}$$

By definition, we have:

$$p\left(\mathbf{f}_{1:n}|\mathbf{x}_{1:n}, I\right) = \mathcal{N}\left(\mathbf{f}_{1:n}|\mathbf{m}_{1:n}, \mathbf{K}_n\right)$$

You need n random variables to sample from this…

PREDICTIVE
SCIENCE LABORATORY

Model Stochastic PDE Problem

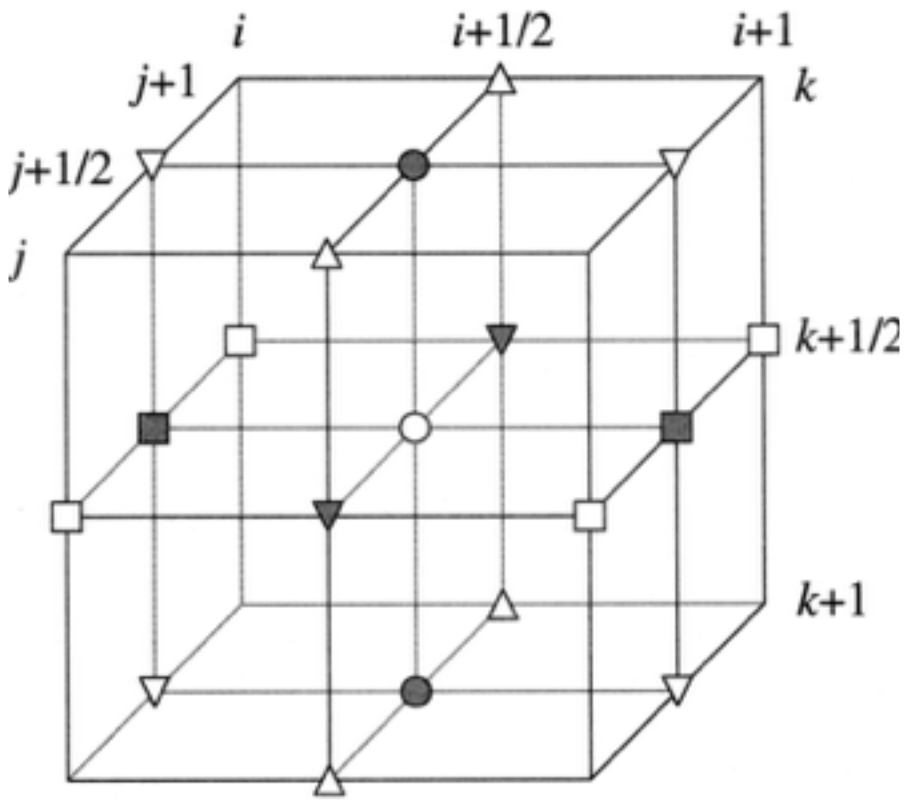
Where do you need to sample in order to solve this?

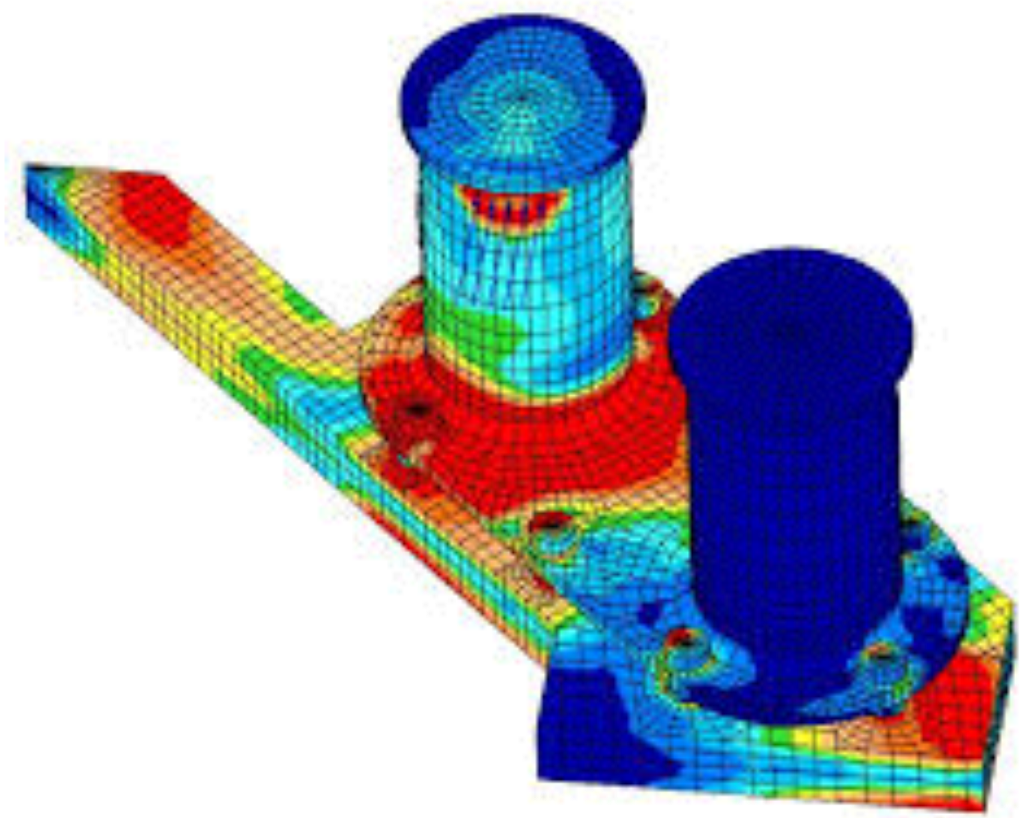
$$\nabla\left(\alpha(\mathbf{x})\nabla u(\mathbf{x})\right) = \rho(\mathbf{x})$$
$$u(\mathbf{x}) = g_D(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_D$$
$$\mathbf{n}\cdot\nabla u(\mathbf{x}) = g_N(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_N$$

Depends on the numerical method you employ…

PREDICTIVE
SCIENCE LABORATORY

Discretization of PDEs





You need to sample at:

- cell centers/faces
- quadrature points of each element
- …

PREDICTIVE
SCIENCE LABORATORY

The Curse of Dimensionality

Where do you need to sample in order to solve this?

$$\nabla\left(\alpha(\mathbf{x})\nabla u(\mathbf{x})\right) = \rho(\mathbf{x})$$
$$u(\mathbf{x}) = g_D(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_D$$
$$\mathbf{n}\cdot\nabla u(\mathbf{x}) = g_N(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_N$$

Suppose the domain is a box and that you discretize using a 100x100x100 regular grid.

n=1,000,000 random numbers to propagate through…

PREDICTIVE
SCIENCE LABORATORY

Dimensionality Reduction

Can you describe this:

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

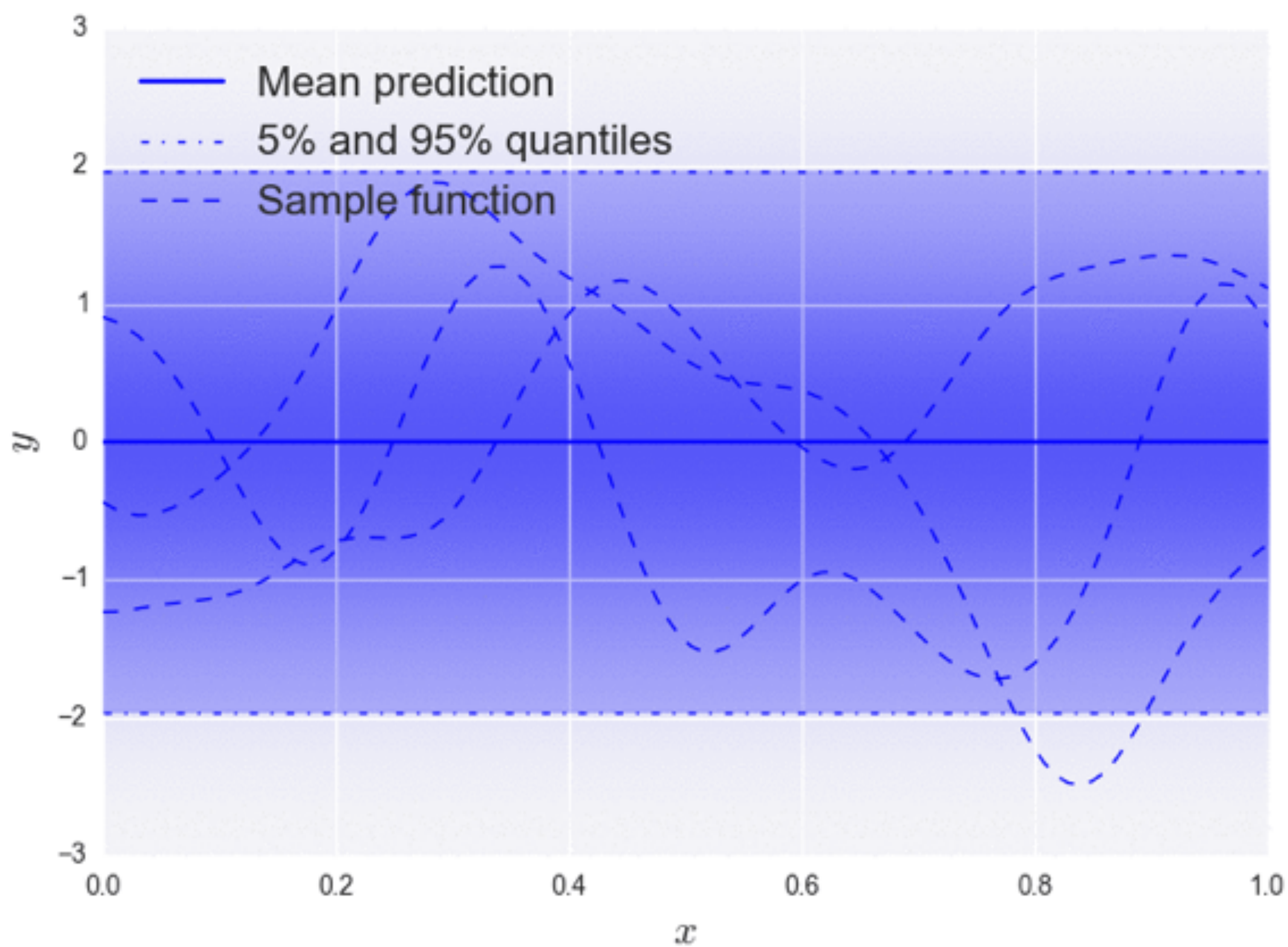
using just a few random variables?

Yes!

PREDICTIVE
SCIENCE LABORATORY

If things are smooth, then yes…

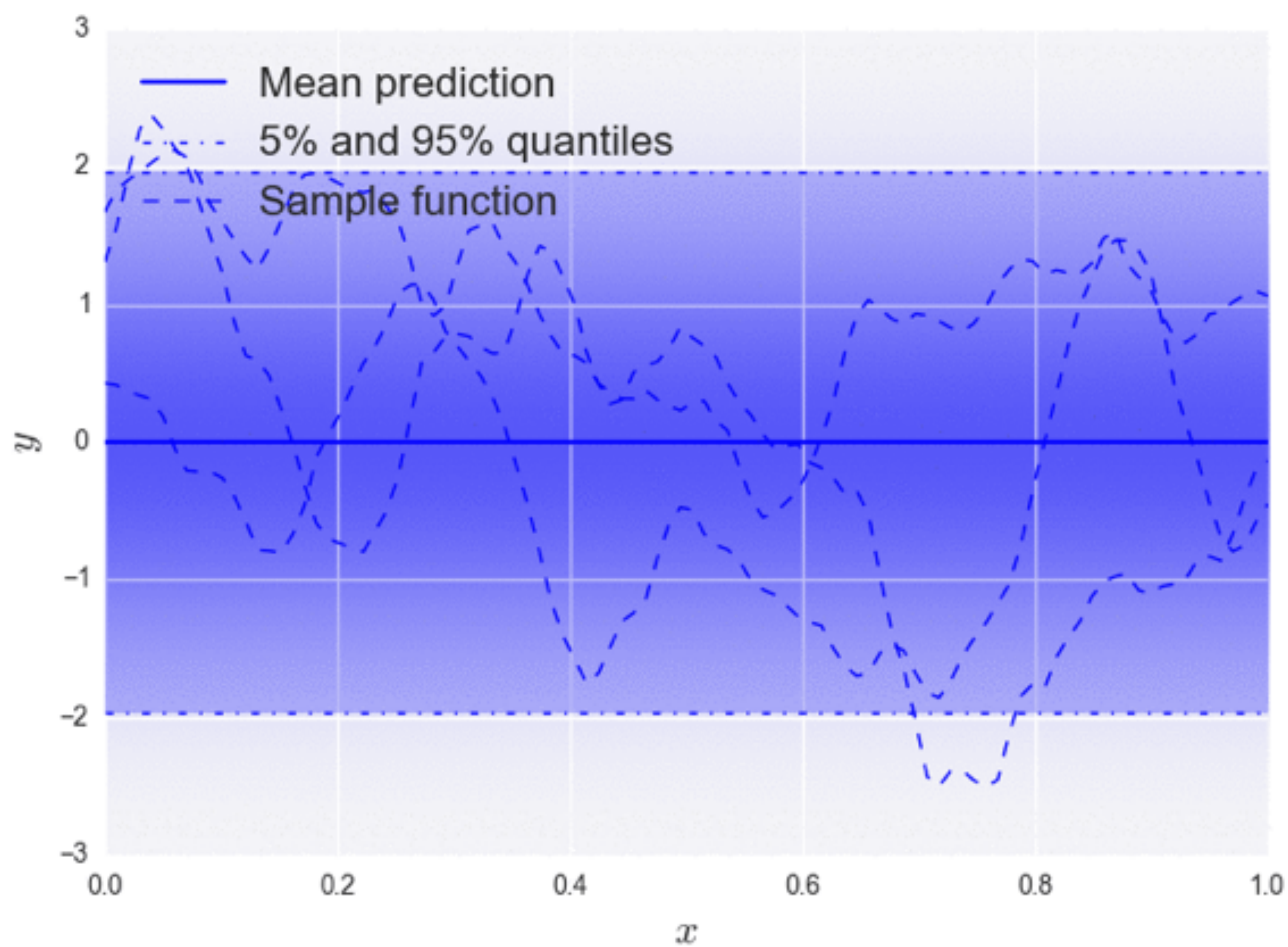
Infinitely smooth SE covariance



PREDICTIVE
SCIENCE LABORATORY

As smoothness decreases, we probably need more

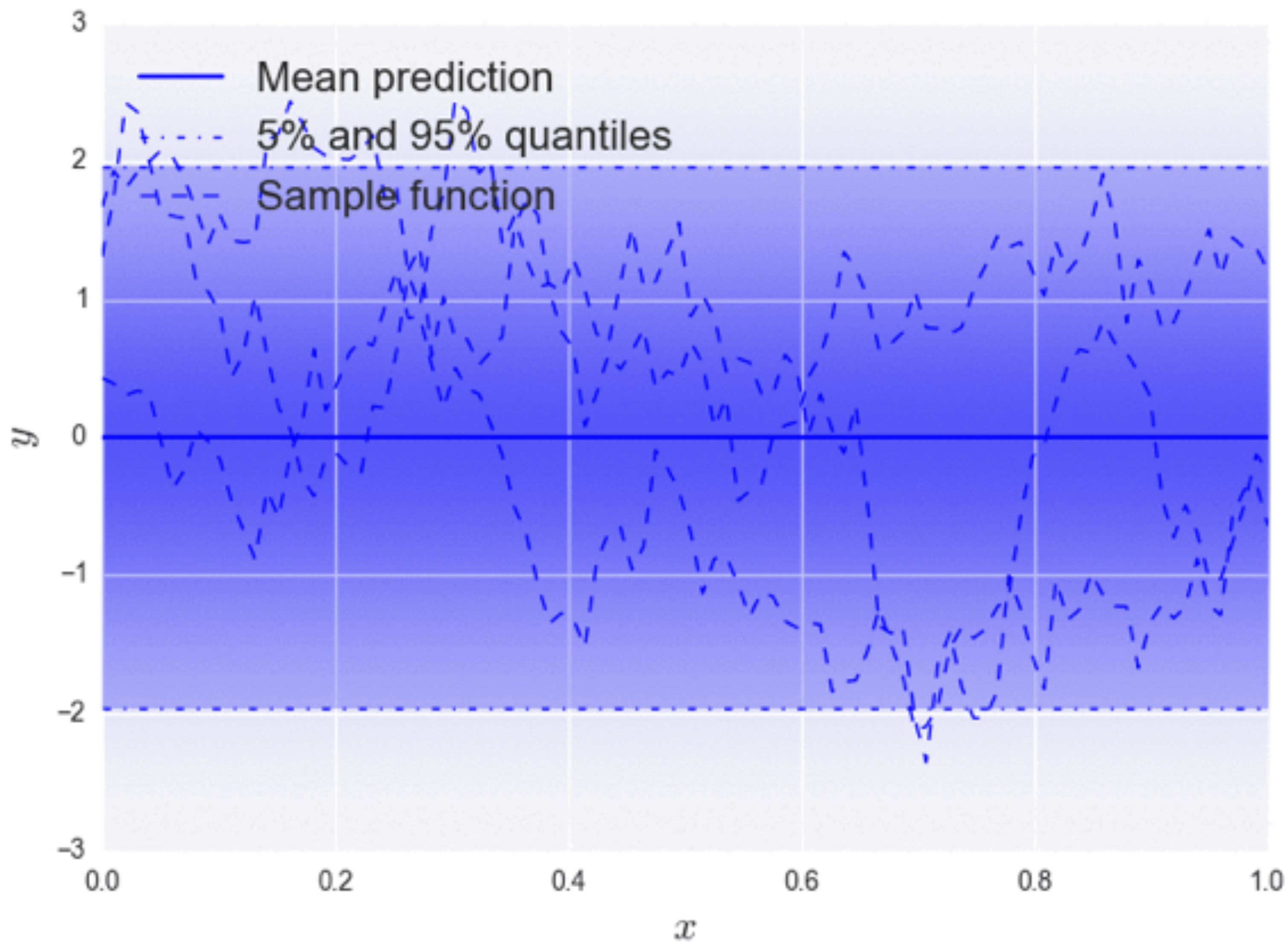
Matern 2-3, 2 times differentiable



PREDICTIVE SCIENCE LABORATORY

and more…

Exponential, continuous, nowhere differentiable



PREDICTIVE
SCIENCE LABORATORY

Karhunen-Loeve Expansion (KLE)

Consider a GP:

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot, \cdot)\right)$$

Then, you can actually express any sample from this as:

$$f(\mathbf{x}; \boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty} \xi_i \sqrt{\lambda_i} \phi_i(\mathbf{x})$$

$$\boldsymbol{\xi} = \{\xi_1, \xi_2, \dots\}$$

Standard normal

PREDICTIVE SCIENCE LABORATORY

Karhunen-Loeve Expansion (KLE)

Consider a GP:

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

Then, you can actually express any sample from this as:

$$f(\mathbf{x};\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

with

$$\xi_i \sim \mathcal{N}(0,1)$$

$$\int k(\mathbf{x},\mathbf{x}')\phi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i\phi_i(\mathbf{x})$$

Eigenvalues/functions of covariance function

PREDICTIVE SCIENCE LABORATORY

Proof of KLE

It suffices to show that:

$$f(\mathbf{x};\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

has the same mean and covariance as the GP.

Here we go:

$$\mathbb{E}_{\boldsymbol{\xi}}[f(\mathbf{x};\boldsymbol{\xi})] = m(\mathbf{x}) + \sum_{i=1}^{\infty}\overset{0}{\mathbb{E}_{\xi_i}[\xi_i]}\sqrt{\lambda_i}\phi_i(\mathbf{x}) = m(\mathbf{x})$$

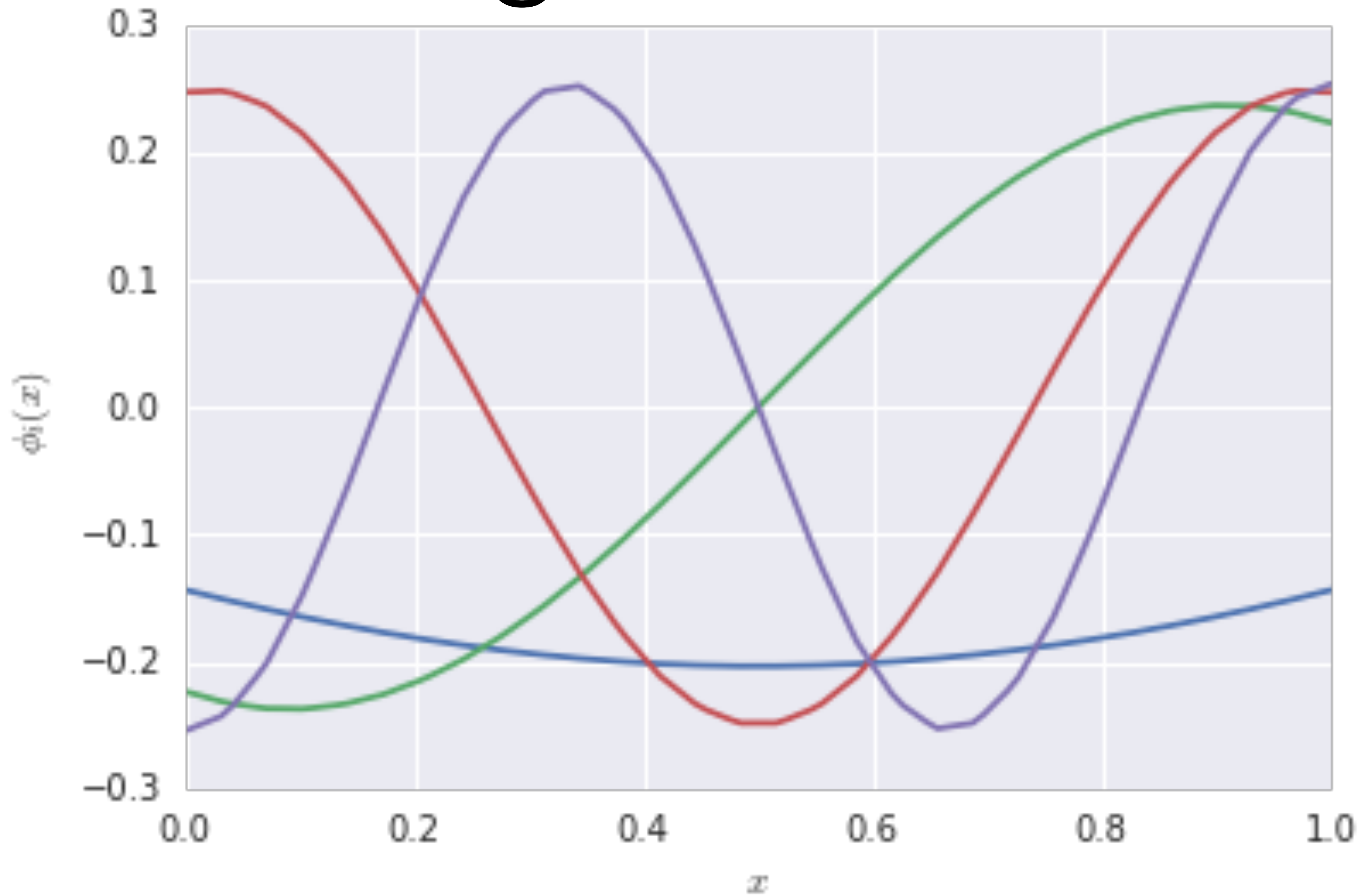
$$\mathbb{C}[f(\mathbf{x};\boldsymbol{\xi}), f(\mathbf{x}';\boldsymbol{\xi})] = \mathbb{E}\left[\left(f(\mathbf{x};\boldsymbol{\xi}) - m(\mathbf{x})\right)\left(f(\mathbf{x}';\boldsymbol{\xi}) - m(\mathbf{x}')\right)\right]$$

$$= \sum_{i=1}^{\infty}\lambda_i\phi_i(\mathbf{x})\phi_i(\mathbf{x}') = k(\mathbf{x},\mathbf{x}')$$

Mercer's Theorem

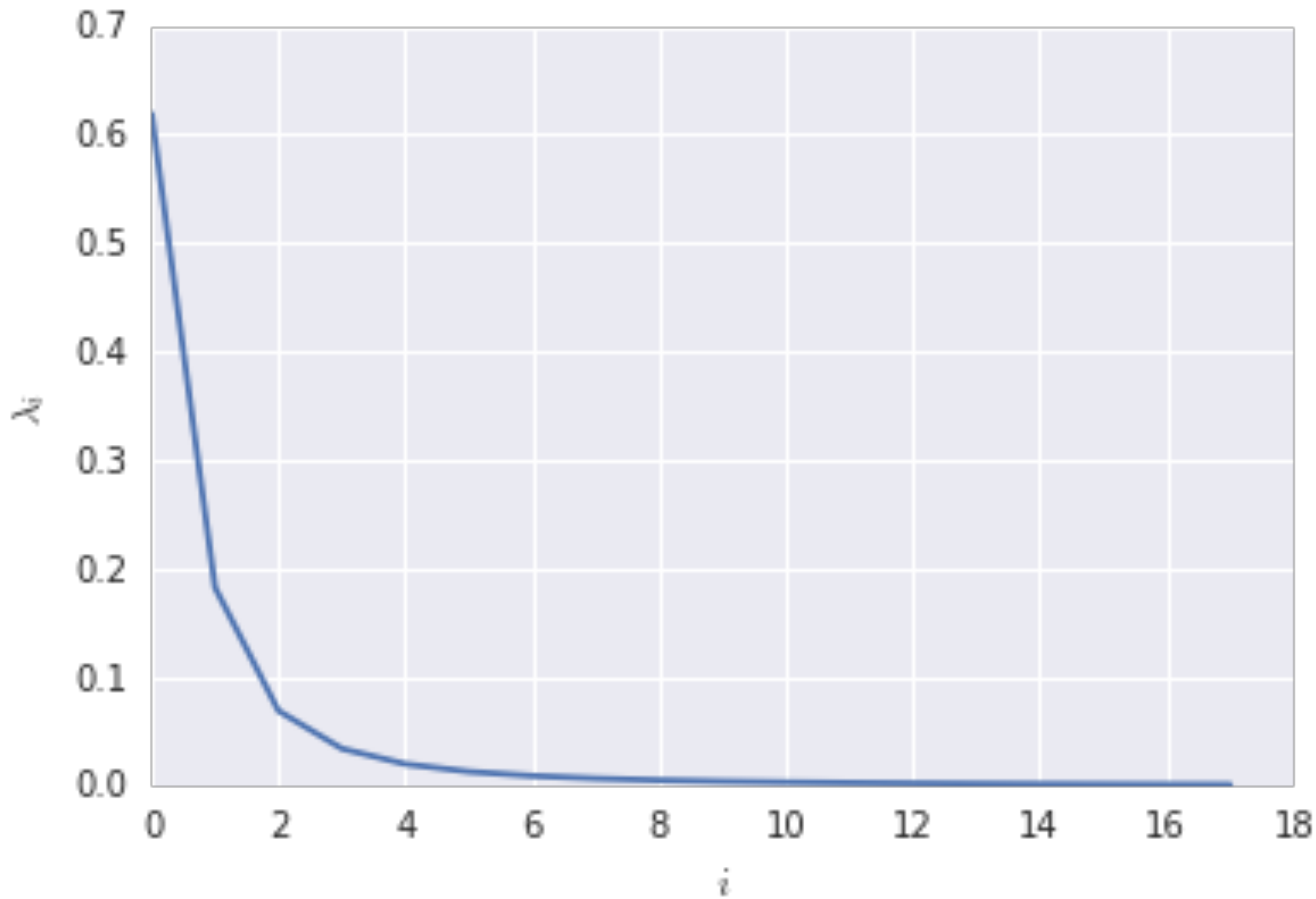
PREDICTIVE
SCIENCE LABORATORY

Example: SE Covariance, the eigenfunctions



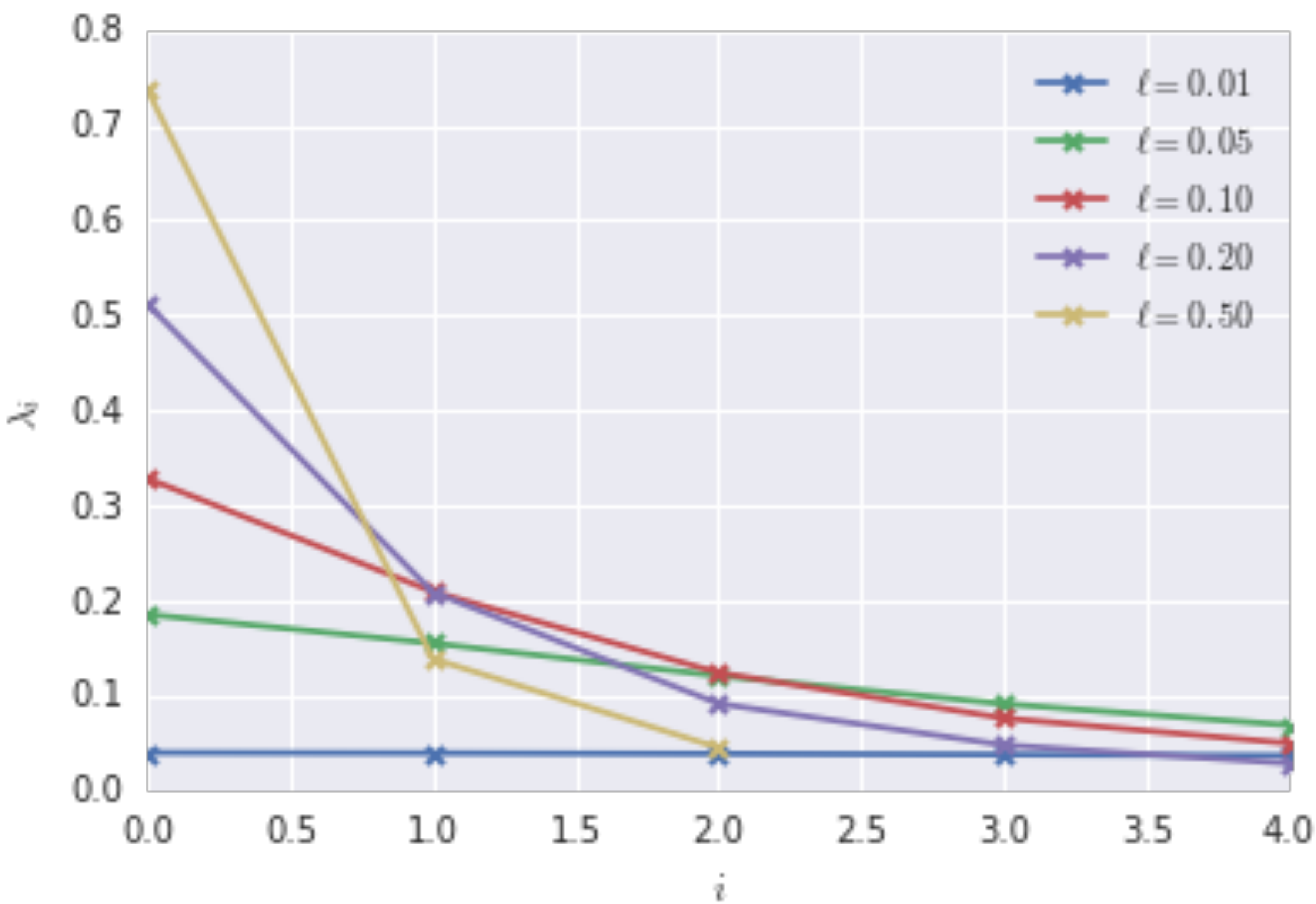
PREDICTIVE
SCIENCE LABORATORY

Example: SE Covariance, the eigenvalues



PREDICTIVE
SCIENCE LABORATORY

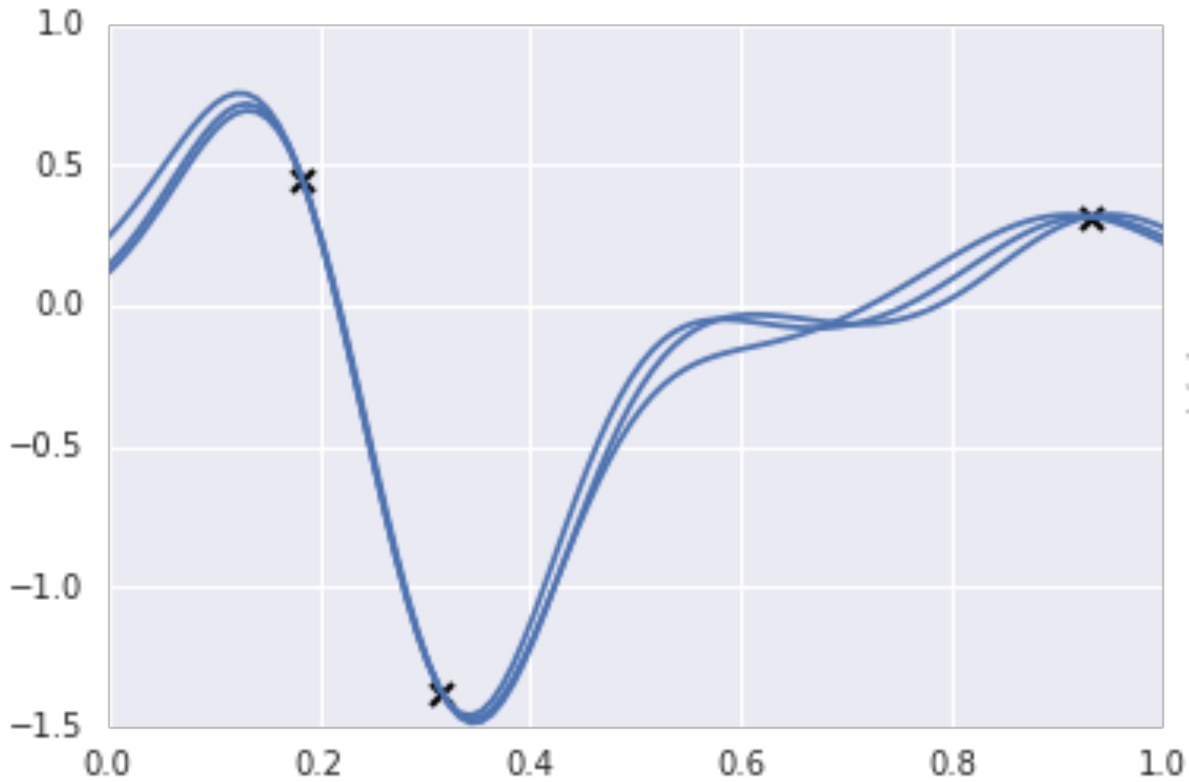
Example: SE Covariance, increasing the length scale

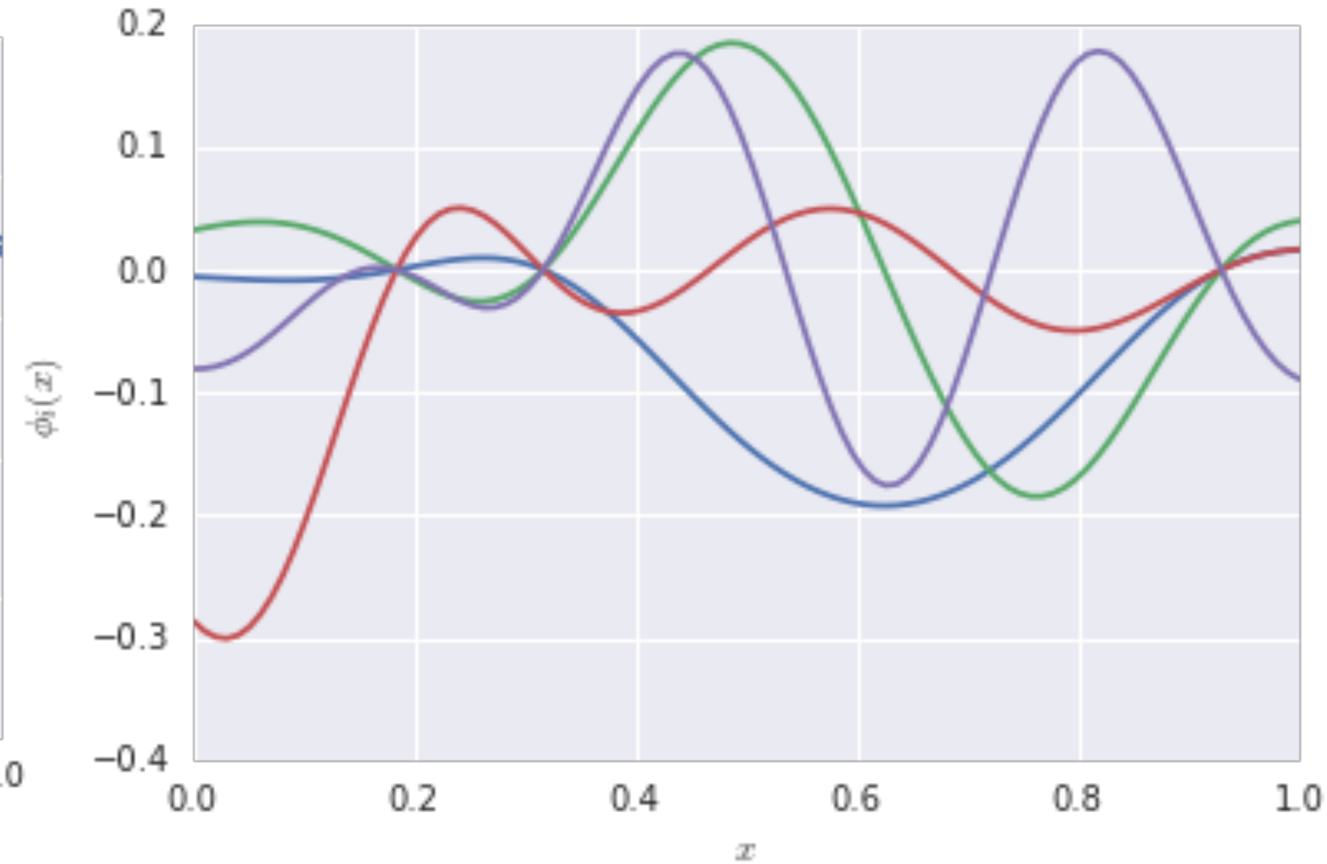


PREDICTIVE
SCIENCE LABORATORY

Posterior Covariance

$$f(\mathbf{x};\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty} \xi_i \sqrt{\lambda_i} \phi_i(\mathbf{x})$$

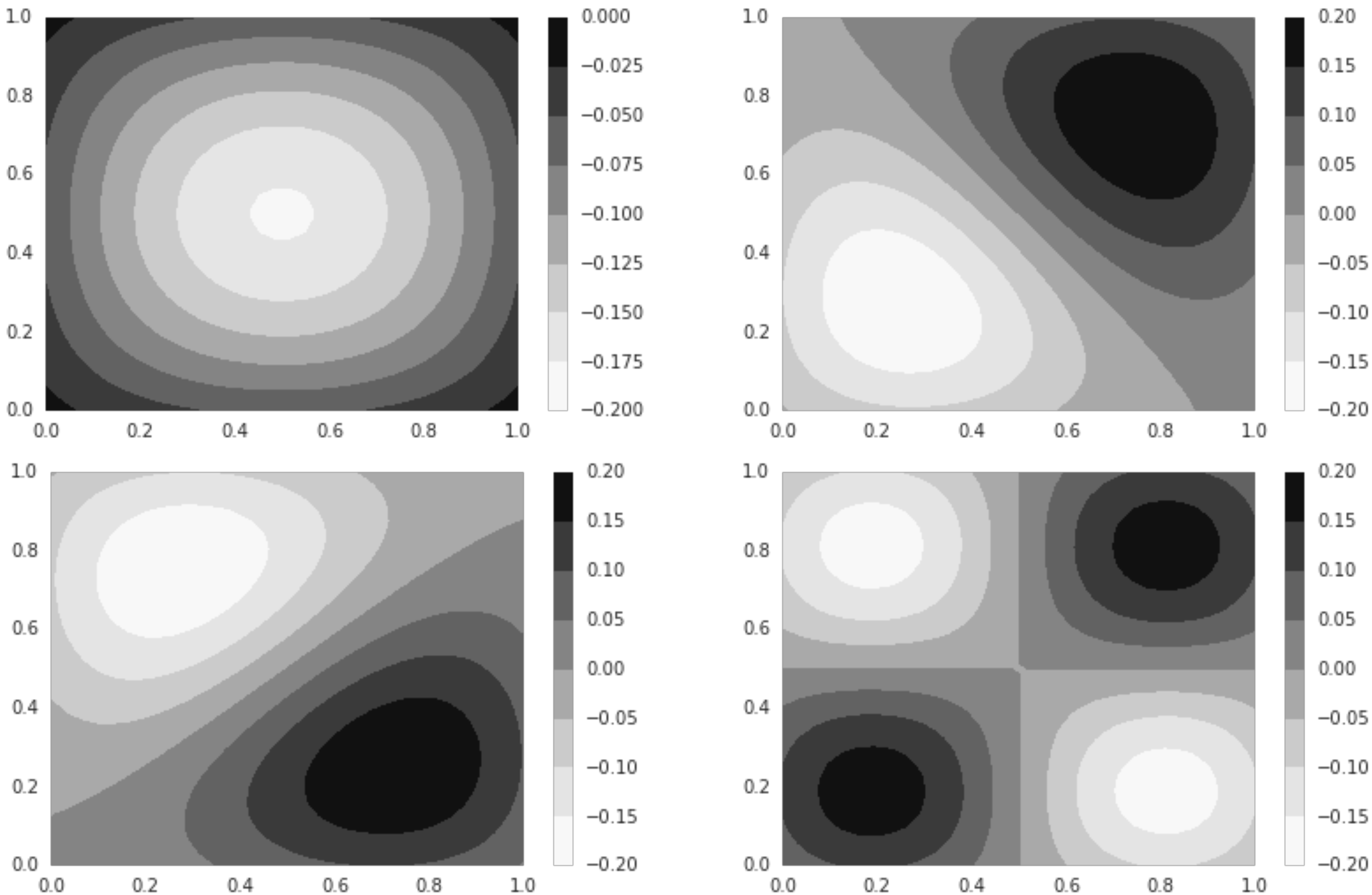




The eigenfunctions are zero at the points where we have observations!

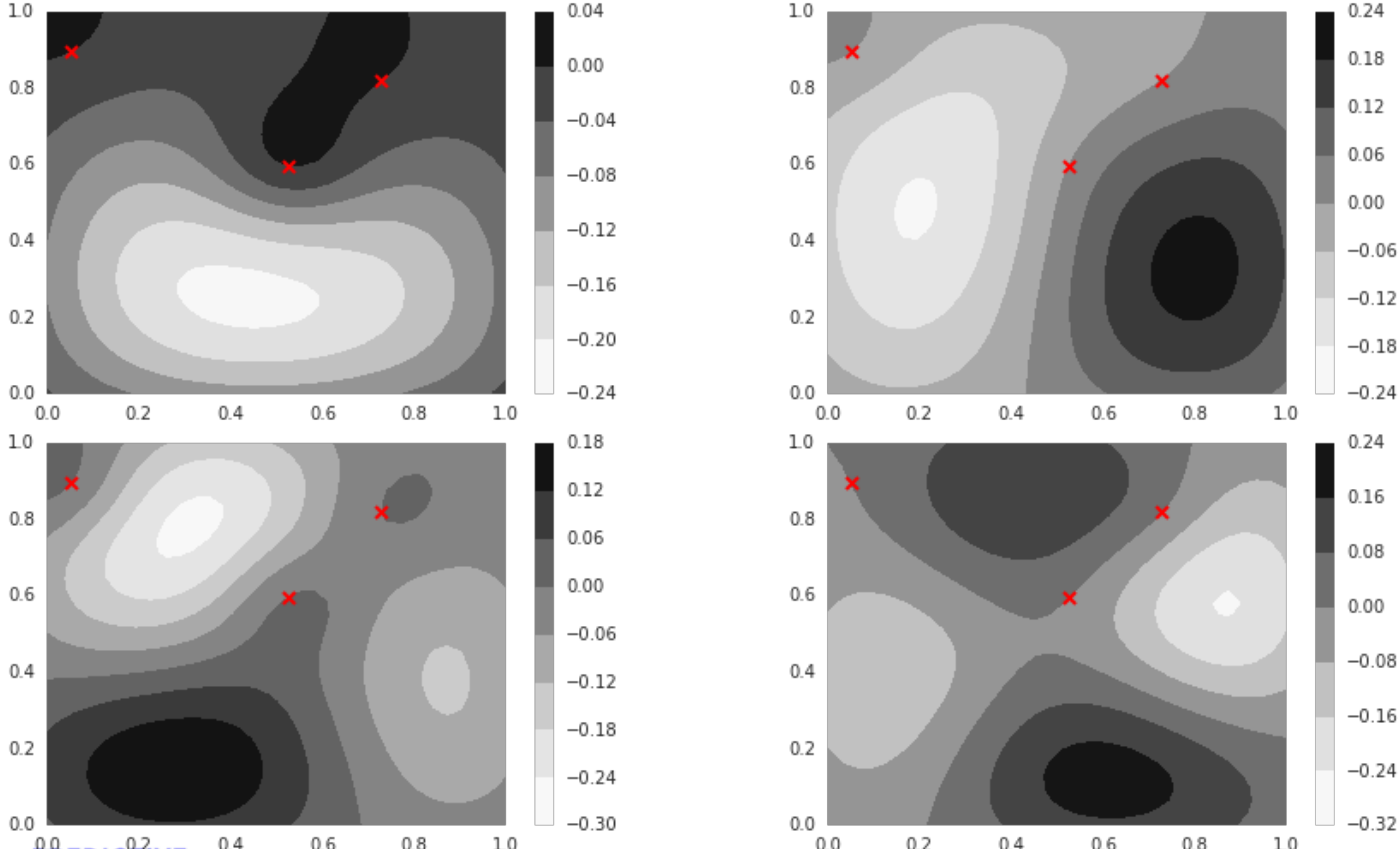
PREDICTIVE
SCIENCE LABORATORY

SE, in two-dimensions



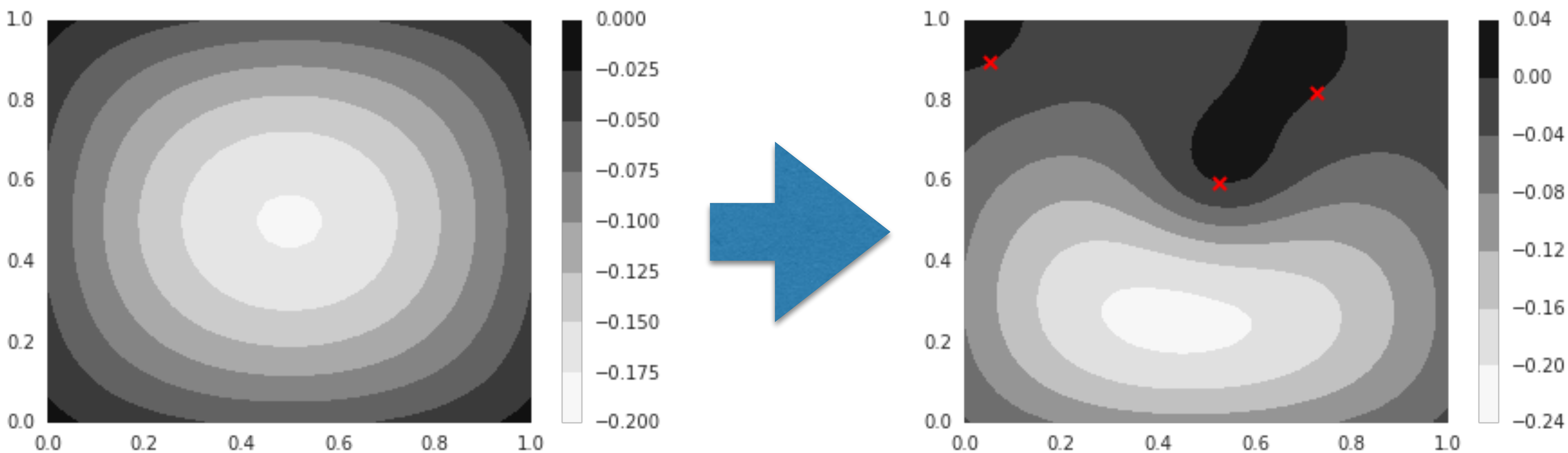
PREDICTIVE
SCIENCE LABORATORY

SE, in two-dimensions with observations



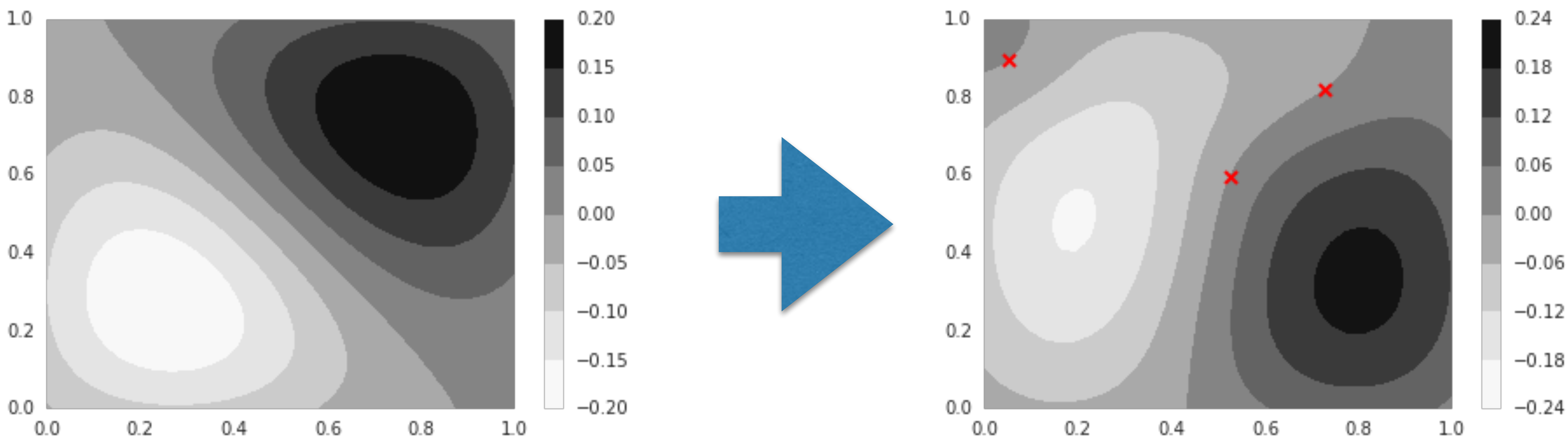
PREDICTIVE
SCIENCE LABORATORY

How did the eigenfunctions change?

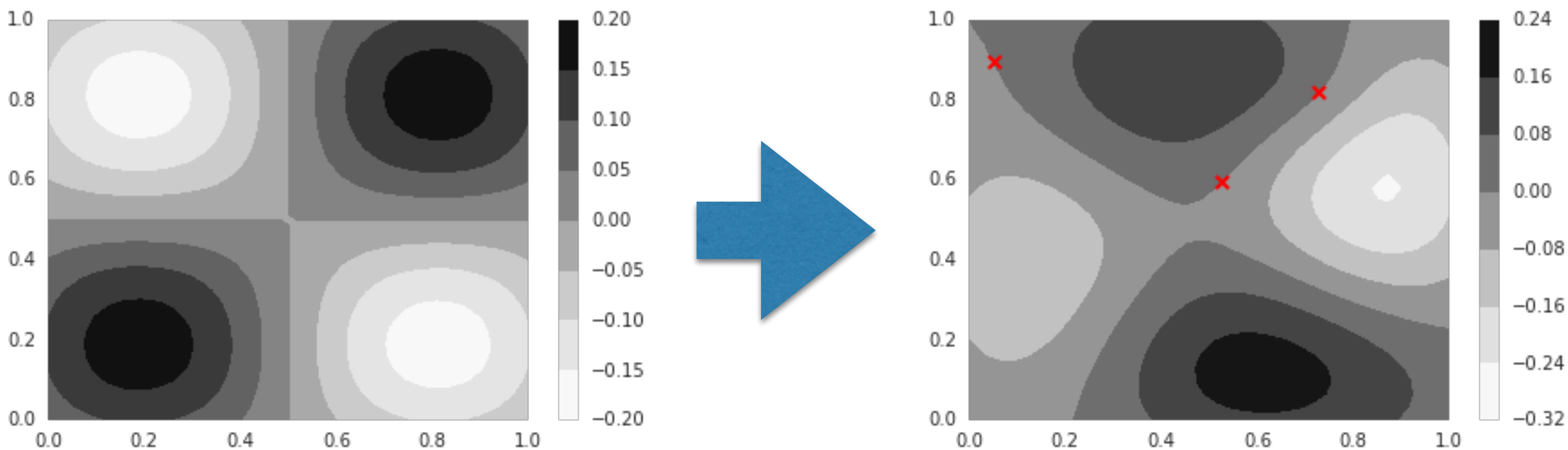


PREDICTIVE
SCIENCE LABORATORY

How did the eigenfunctions change?



How did the eigenfunctions change?



PREDICTIVE
SCIENCE LABORATORY

Karhunen-Loeve Expansion (KLE)

Consider a GP:

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

Then, you can actually express any sample from this as:

$$f(\mathbf{x};\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

After some terms, these are zero!

with

$$\xi_i \sim \mathcal{N}(0,1) \qquad \int k(\mathbf{x},\mathbf{x}')\phi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i\phi_i(\mathbf{x})$$

Where is the dimensionality reduction?

PREDICTIVE
SCIENCE LABORATORY

Truncated Karhunen-Loeve Expansion (TKLE)

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

$$f(x;\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

$$\Downarrow$$

$$f(x;\xi) = m(\mathbf{x}) + \sum_{i=1}^{d}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

How do we select d?

PREDICTIVE
SCIENCE LABORATORY

Truncated Karhunen-Loeve Expansion (TKLE)

$$p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot), k(\cdot,\cdot)\right)$$

$$f(x;\boldsymbol{\xi}) = m(\mathbf{x}) + \sum_{i=1}^{\infty} \xi_i \sqrt{\lambda_i}\phi_i(\mathbf{x})$$

$$f(x;\xi) = m(\mathbf{x}) + \sum_{i=1}^{d} \xi_i \sqrt{\lambda_i}\phi_i(\mathbf{x})$$

$$\sum_{i=1}^{d} \lambda_i = 0.95 \times \sum_{i=1}^{\infty} \lambda_i$$

Energy or Variance of the field

PREDICTIVE SCIENCE LABORATORY

“OK, I’m sold!
How can I compute the KLE?”

The Nystrom Approximation

We need to solve a Fredholm integral equation:

$$\int k(\mathbf{x}, \mathbf{x}')\phi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i \phi_i(\mathbf{x})$$

Approximate the left hand side with a quadrature:

$$\sum_{j=1}^{n_q} w_j k(\mathbf{x}, \mathbf{x}_j)\phi_i(\mathbf{x}_j) \approx \lambda_i \phi(\mathbf{x})$$

Assume that the equation holds at the quadrature points:

$$\sum_{j=1}^{n_q} w_j k(\mathbf{x}_r, \mathbf{x}_j)\phi_i(\mathbf{x}_j) \approx \lambda_i \phi(\mathbf{x}_r)$$

PREDICTIVE
SCIENCE LABORATORY

The Nystrom Approximation

We need to solve a Fredholm integral equation:

$$\int k(\mathbf{x},\mathbf{x}')\phi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i\phi_i(\mathbf{x})$$

$$\sum_{j=1}^{n_q} w_j k(\mathbf{x}_r,\mathbf{x}_j)\phi_i(\mathbf{x}_j) \approx \lambda_i\phi(\mathbf{x}_r)$$

$$\mathbf{K}_q\mathrm{diag}(w_1,\dots,w_{n_q})\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Eigenvalue problem

PREDICTIVE
SCIENCE LABORATORY

The Nystrom Approximation

We need to solve a Fredholm integral equation:

$$\int k(\mathbf{x},\mathbf{x}')\phi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i\phi_i(\mathbf{x})$$

$$\mathbf{K}_q\mathrm{diag}(w_1,\dots,w_{n_q})\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Eigenvalue problem

$$\phi_i(\mathbf{x}) = \lambda_i^{-1}\sum_{j=1}^{n_q} w_j k(\mathbf{x},\mathbf{x}_j)\phi_i(\mathbf{x}_j) = \lambda_i^{-1}\sum_{j=1}^{n_q} w_j k(\mathbf{x},\mathbf{x}_j)v_{ij}$$

PREDICTIVE
SCIENCE LABORATORY

Model Stochastic PDE Problem

Consider the elliptic partial differential equation:

$$\nabla\left(\alpha(\mathbf{x})\nabla u(\mathbf{x})\right) = \rho(\mathbf{x})$$

with boundary conditions:

$$u(\mathbf{x}) = g_D(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_D$$

$$\mathbf{n}\cdot\nabla u(\mathbf{x}) = g_N(\mathbf{x}),\ \text{on } \mathbf{x}\in\Gamma_N$$

Since it is positive:

$$\alpha(\mathbf{x}) = \exp\left\{f(\mathbf{x})\right\}$$

$$\cancel{p(f(\cdot)|I) = \mathrm{GP}\left(f(\cdot)|m(\cdot),k(\cdot,\cdot)\right)}$$

$$f(x;\xi) = m(\mathbf{x}) + \sum_{i=1}^{d}\xi_i\sqrt{\lambda_i}\phi_i(\mathbf{x})$$

PREDICTIVE
SCIENCE LABORATORY