GAUSSIAN PROCESS REGRESSION

- Noisy measurement y at location x

$$y(x) = f(x) + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma_n^2), f(x) \sim \mathrm{GP}(0, k(x, x'))$.

- Prediction at new locations x^*

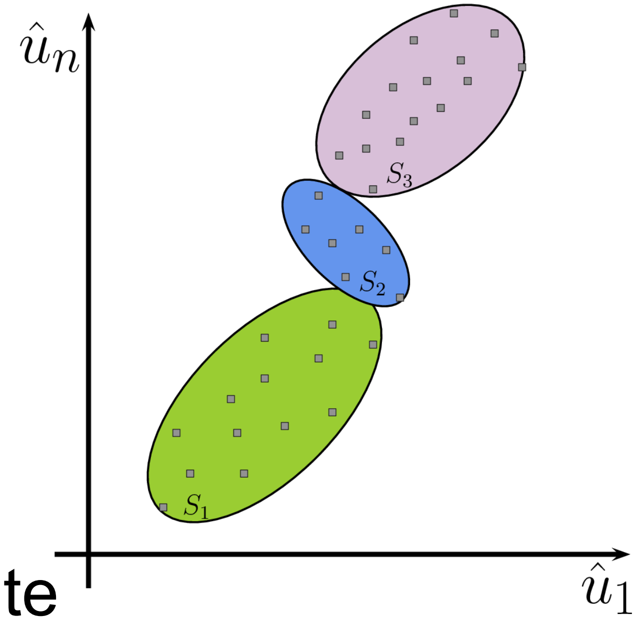
$$\mathbb{E}[f(x^*)|(x, y)] = k(x^*, x)k_{\sigma_n^2}(x, x)^{-1}y$$

- Prediction of uncertainty at x^*

$$\mathrm{Cov}[f(x^*)|(x, y)] = k(x^*, x^*) - k(x^*, x)k_{\sigma_n^2}(x, x)^{-1}k(x, x^*)$$

- Bottleneck: $\mathcal{O}(N^3)$ complexity for inverting the sample covariance matrix!

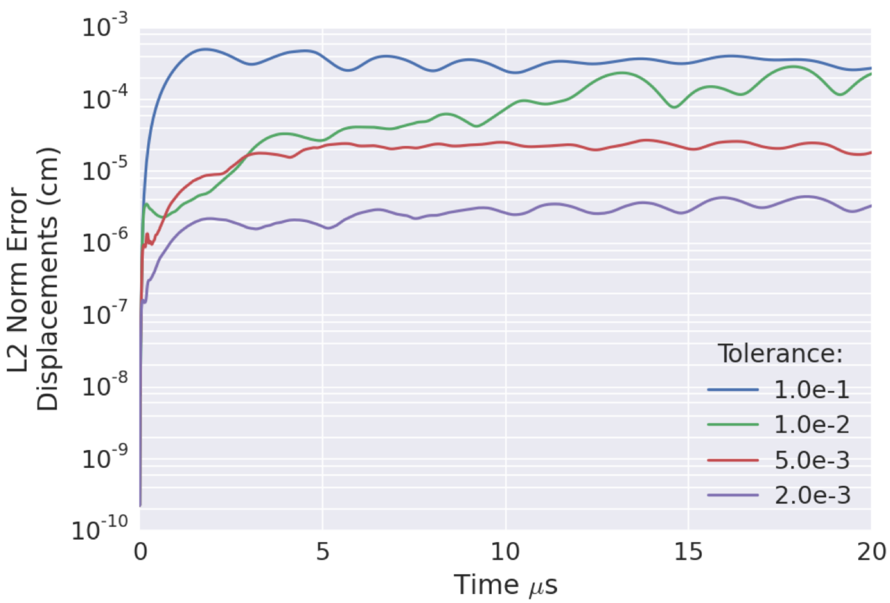
ADAPTIVE ONLINE GPR



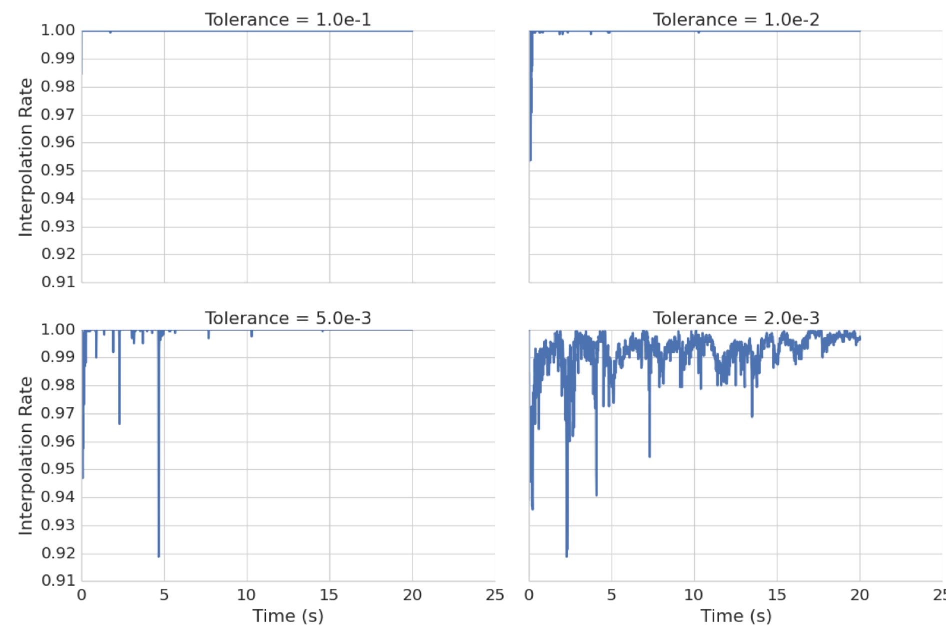
- Build surrogate models over subsets of the domain of f
- Select subsets along trajectories induced by F
- Adaptive online GPR
  - Impose an upper bound on the number of data points in a surrogate model
  - Continuously construct surrogate models from nearby data points
  - Employ the covariance to adaptively refine and construct new surrogate models given a target error of surrogate models

ADAPTIVE ONLINE GPR

- 2D (axisymmetric) model of impact with 1,600 elements
- 4 different tolerances



L2 error of solution

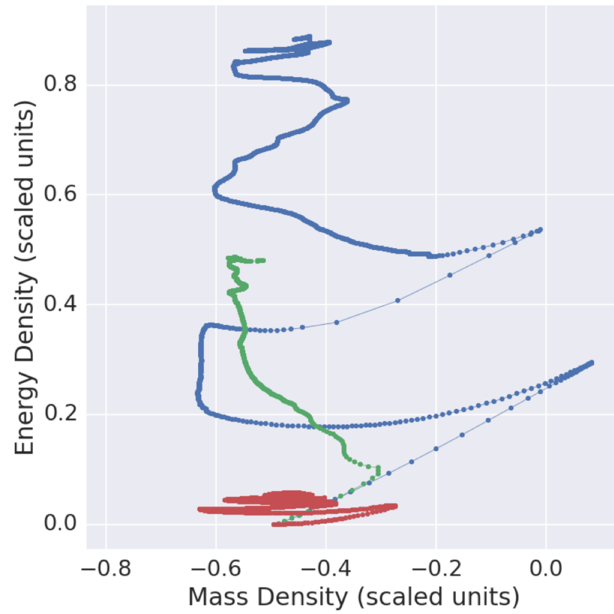


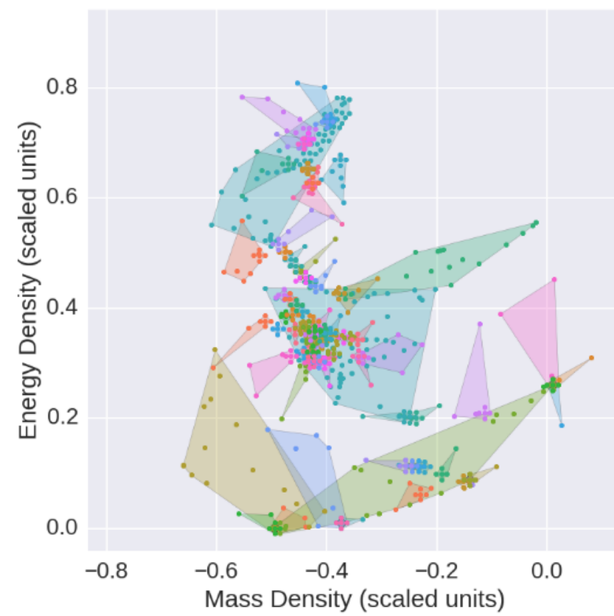
Interpolation rate

APPROVED FOR PUBLIC RELEASE

ADAPTIVE ONLINE GPR

- Structure of adaptive online GPR clearly induced by F
- Only a portion of the domain of f covered by surrogate models
- Adaptive online GPR substantially reduces the cost of computing f
- Individual surrogate models are continuous/smooth
- But global continuity/smoothness is lost
- Can the cost of constructing GPR be significantly reduced?



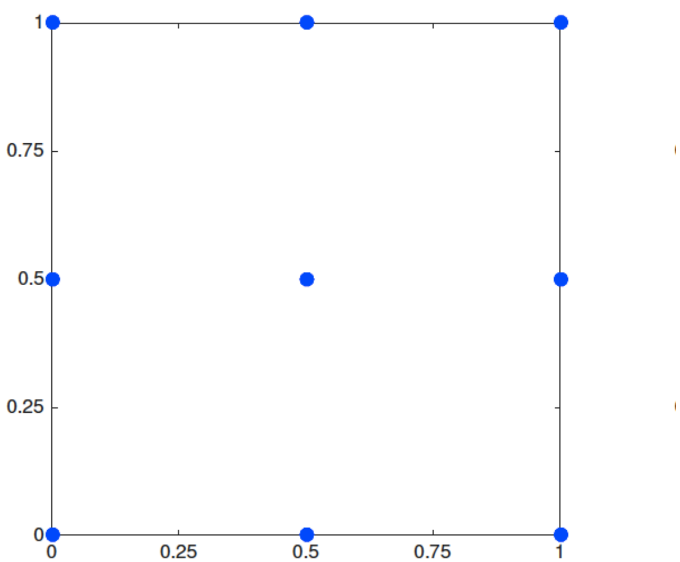


OFFLINE SPARSE GP

- Sub-sampling (Rasmussen and Williams, 2006): low rank approximation of the covariance matrix
- Inducing variables (Quinonero-Candela and Rasmussen, 2005): introduce sparsity into the covariance kernel by conditioning
- Predictive process approach (Banerjee at al., 2008): screen effects by Stein
- SPDE approach (Lindgren et al., 2011): exploit the link between Matern kernels and SPDE
- Hierarchical Cholesky decomposition (Schafer et al., 2017):
  - multi-resolution representation of the covariance matrix and the incomplete Cholesky decomposition
  - near linear complexity
  - exponentially small approximation error
  - Operates on a near-uniform grid

EXAMPLE: A TWO-LEVEL GRID $\mathcal{D}^{(2)}$

First Level Uniform Grid $X^{(1)}$



Index sets for one-level:

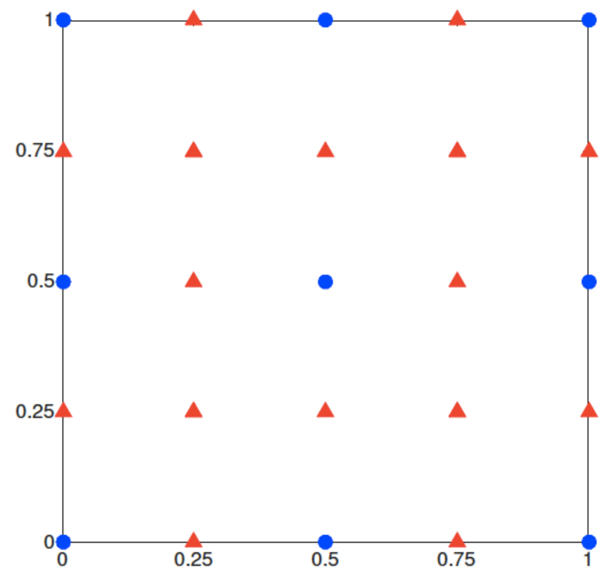
$$I^{(1)} = \{1, 2, \ldots, 9\}$$

$$J^{(1)} = \{1, 2, \ldots, 9\}$$

Covariance matrix for one-level:

$$K^{(1)} = k(X^{(1)}, X^{(1)}) \in \mathbb{R}^{9\times 9}$$

Second Level Uniform Grid $X^{(2)}$



Index sets in two-level:

$$I^{(2)} = \{1, 2, \ldots, 25\}$$

$$J^{(2)} = \{10, 11, \ldots, 25\}$$

Covariance matrix for two-level:

$$K^{(2)} = k(X^{(2)}, X^{(2)}) \in \mathbb{R}^{25\times 25}$$