SPARSE GP ON GRIDS $\mathcal{D}^{(q)}$

- Choose the sparsity parameter R
- Define the sparsity pattern

$$\mathcal{S}_R = \left\{(i,j) \in I^{(q)} \times I^{(q)} \,\middle|\, i \in J^{(k)}, j \in J^{(l)}, |x_i - x_j| \leq R2^{-k \wedge l}\right\}$$

- Define the sparse matrix based on $\mathcal{S}_R$

$$K_R^{(q)}(i,j) = \begin{cases} K^{(q)}(i,j) & \text{for } (i,j) \in \mathcal{S}_R \\ 0 & \text{otherwise,} \end{cases}$$

- Let $L_R^{(q)}$ the zero fill-in incomplete Cholesky factor of $K_R^{(q)}$
- Accuracy (exponentially small error): $\left\| K^{(q)} - L_R^{(q)} L_R^{(q),T} \right\| \leq p(N) e^{-\gamma R}$
- Efficiency (near linear instead of cubic): $\mathcal{O}(N \log^2(N) R^4)$

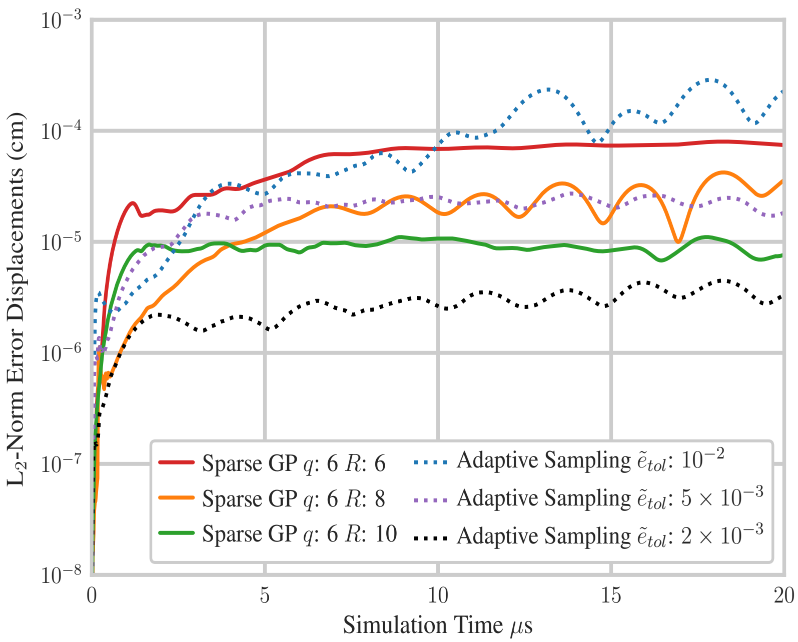
Algorithm 1 Sparse GP on an uniform grid of level q

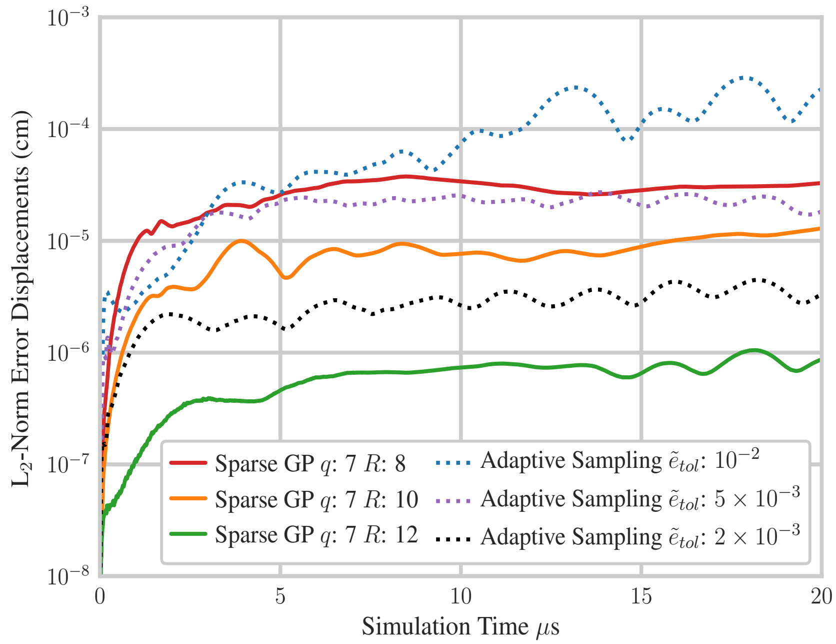
```
procedure SparseGP(D^(q))
    Order the data points from J^(1) to J^(q) (coarse to fine)
    Initialize K_R^(q)(i, j) = K^(q)(i, j) if (i, j) ∈ S_R and K_R^(q)(i, j) = 0 otherwise
    L_R^(q) ← IncompleteChol(K_R^(q))          ▷ Incomplete Cholesky decomposition
    α = L_R^(q),T \ L_R^(q) \ Y                ▷ Backward substitution and then forward substitution
    v = L_R^(q) \ k(X^(q), X*)
    E[f(X*)|D^(q)] = k(X^(q), X*)^T α          ▷ Update the predictive mean
    Cov[f(X*)|D^(q)] = k(X*, X*) − v^T v       ▷ Update the predictive variance
```

SIMULATION RESULT: ACCURACY

- Six-level grid (q = 6, N = 4225)
- Sparsity parameter R = 6, 8, 10

- Seven-level grid (q = 7, N = 16641)
- Sparsity parameter R = 8, 10, 12





SIMULATIONS RESULT: EFFICIENCY

- Sparsity: fraction of zero entries in the sparsified covariance matrix
- Wall-Clock Time: time for macro-scale model simulation and surrogate evaluation
- Compute Time: total of the wall-clock time and the time for sampling data from micro-scale model

q	R	Sparsity	Wall-Clock Time (hr)	Hyperparameter Opt. Time (hr)	Compute Time (hr)	# Microscale Model Evals
6	6	0.74	0.9	0.6	11,504	4,225
6	8	0.64	1.5	0.8	11,525	4,225
6	10	0.54	2.7	1.5	11,561	4,225
7	8	0.85	11.4	8.8	44,345	16,641
7	10	0.80	18.4	15.3	44,570	16,641
7	12	0.74	40.9	32.3	45,274	16,641

$\tilde{e}_{tol}$	Wall-Clock Time (hr)	Compute Time (hr)	# Microscale Model Evals
10^{-2}	2.3	7,546	833
5×10^{-3}	8.7	28,315	1,878
2×10^{-3}	156.3	506,333	27,827
No Surrogate Module	769.2	9,885,192	2,681,600

SUMMARY

- Hierarchical multiscale can produce high-fidelity material models
- Yet, the total cost of repeated evaluations of the lower-scale model can be often staggering, rendering the overall method impractical
- Surrogate modeling can substantially reduce the cost
- Gaussian process regression well-suited for construction of surrogate models
  - Prediction of function values accompanied by prediction of error
  - Cost of construction can be high
  - Hierarchical Cholesky decomposition can bring the cost down considerably

OPEN PROBLEMS

- Surrogate modeling
  - High-dimensional spaces
  - Space-time surrogate models
  - Adaptive, multi-fidelity surrogate models
  - Modeling of discontinuities
  - Surrogate models incorporating invariants
- Scale-bridging methods
  - Spatial scale-bridging for materials with evolving internal state
  - Temporal-scale bridging

OPEN PROBLEMS

- UQ for multiscale modeling
  - Propagation and aggregation of uncertainty across scales
  - Methods for design under uncertainty for complex materials
- Stochastic modeling and multiscale
  - The treatment of microstructure and its evolution may necessitate casting the multiscale modeling as a stochastic problem