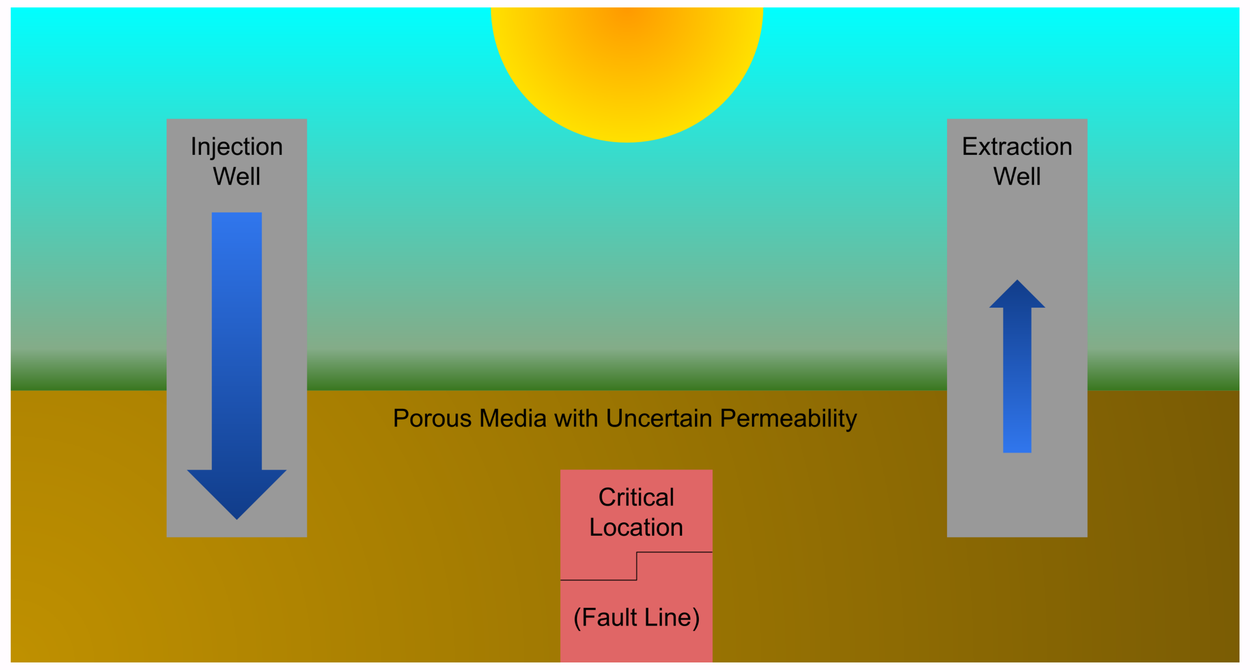
Probabilistic Models for PDEs with Random Coefficients

Aleksei G. Sorokin[1,2] Aleksandra Pachalieva[1] Dan O'Malley[1] Mac Hyman[1] Fred J. Hickernell[2] Nick Hengartner[1]

[1]LANL. [2]Illinois Institute of Technology, Department of Applied Mathematics.

LA-UR 23-28107

Subsurface Diagram



Darcy's Equation

$\nabla.(G(x).\nabla p(E,G,x)) = f(E,x)$ models pressure in porous media

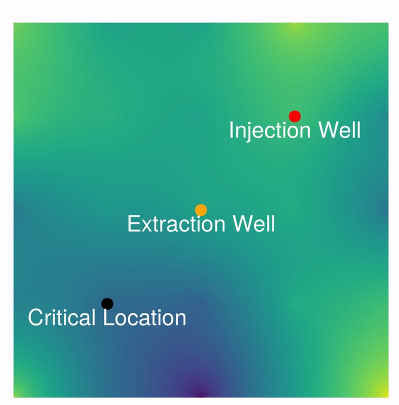
- $G(x)$ a **random** *Gaussian permeability field*
- $p(x)$ the **random** *pressure solution*

Following [4], we model the *flow rate* as

$$f(x,E) = \begin{cases} I, & x = x_{\text{injection}} \\ -E, & x = x_{\text{extraction}} \\ 0, & \text{else} \end{cases}.$$

- I, the **fixed** *injection rate*
- E, the **variable** *extraction rate*

Birds-Eye View



Problem Outline

- Given critical location x_{critical} e.g. fault line
- Given *pressure threshold* $\bar{p}$ at the critical location
- $p^c(E,G) := p(E,G,x_{\text{critical}})$, the *critical pressure*

Want **online** choice of smallest E which gives high *confidence*

$$P(p^c(E,G) \le \bar{p})$$

in keeping low enough pressure at the critical location

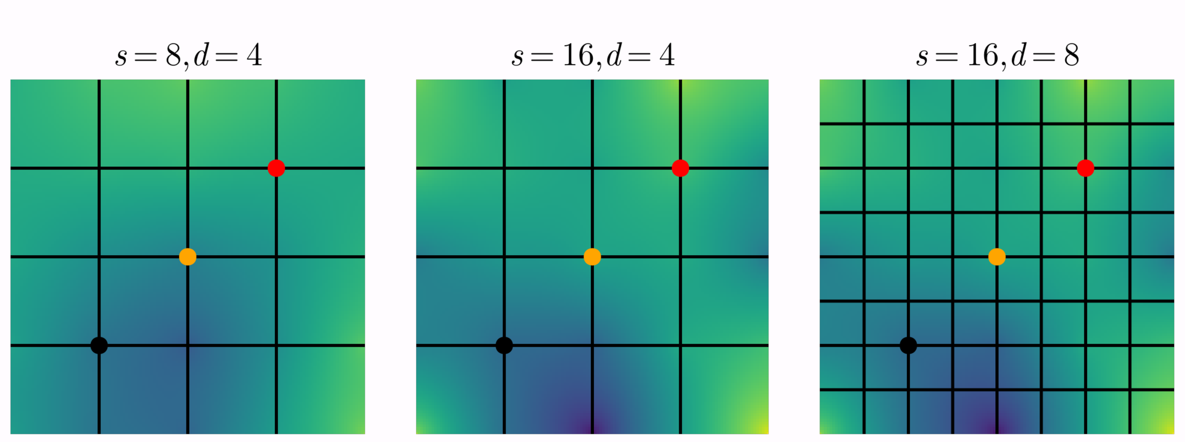
Challenges

- Must approximate $P(p^c(E,G) \le \bar{p})$ for many different E values
- Approximating $P(p^c(E,G) \le \bar{p})$ requires many samples of $G(x)$

Numerical Solution $p^c_{s,d}(E,\boldsymbol{Z})$

s, **KL Dimension:** $G(x) \approx \sum_{j=1}^s \sqrt{\lambda_j} Z_j \varphi_j(x)$ and $\boldsymbol{Z} \sim \mathcal{N}(0, I_s)$

d, **Discretization Dimension:** Mesh width $1/d$ in each physical dimension



Don't Approximate $P(p^c(E,G) \le \bar{p})$ by $P(p^c_{s,d}(E,\boldsymbol{Z}) \le \bar{p})$

- Estimating $P(p^c_{s,d}(E,\boldsymbol{Z}) \le \bar{p})$ for many E requires large number of $\boldsymbol{Z}$ samples
- Numerical solutions of $p^c_{s,d}(E,\boldsymbol{Z})$ are too expensive to compute online
- $P(p^c_{s,d}(E,\boldsymbol{Z}) \le \bar{p}) \neq P(p^c(E,G) \le \bar{p})$ generally

Use Data $\{p^c_{s,d}(E_i,\boldsymbol{Z}_i)\}_{i=1}^n$ to Build a Model for $p^c(E,G)$

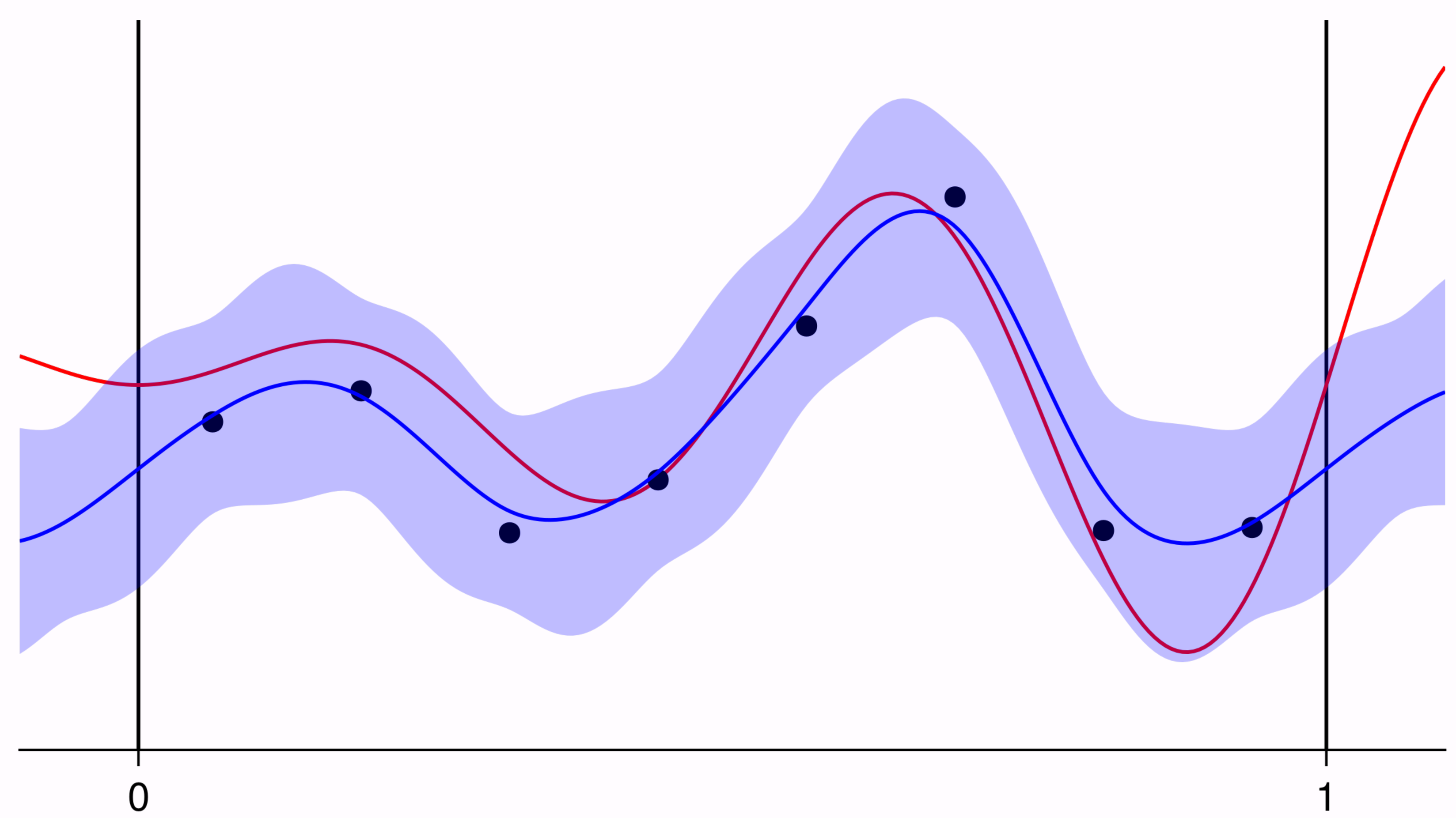
- Can choose sampling nodes $(E_i,\boldsymbol{Z}_i)_{i=1}^n$ to explore the space well
- Numerical solutions $\{p^c_{s,d}(E_i,\boldsymbol{Z}_i)\}_{i=1}^n$ can be computed offline
- Surrogate model can be fit offline and evaluated quickly online

$\therefore$ Inexpensive to estimate confidence $P(p^c(E,G) \le \bar{p})$ online using surrogate

Gaussian Processes as Probabilistic Surrogates for p_c

Assume $p^c_{s,d} = p^c + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0,\zeta_{s,d})$

numerical solution = solution + Gaussian noise



solution p^c, numerical solutions, approximate solution, uncertainty in solution

Structuring Nodes and Kernel Enables $\mathcal{O}(n \log n)$ GP Scaling

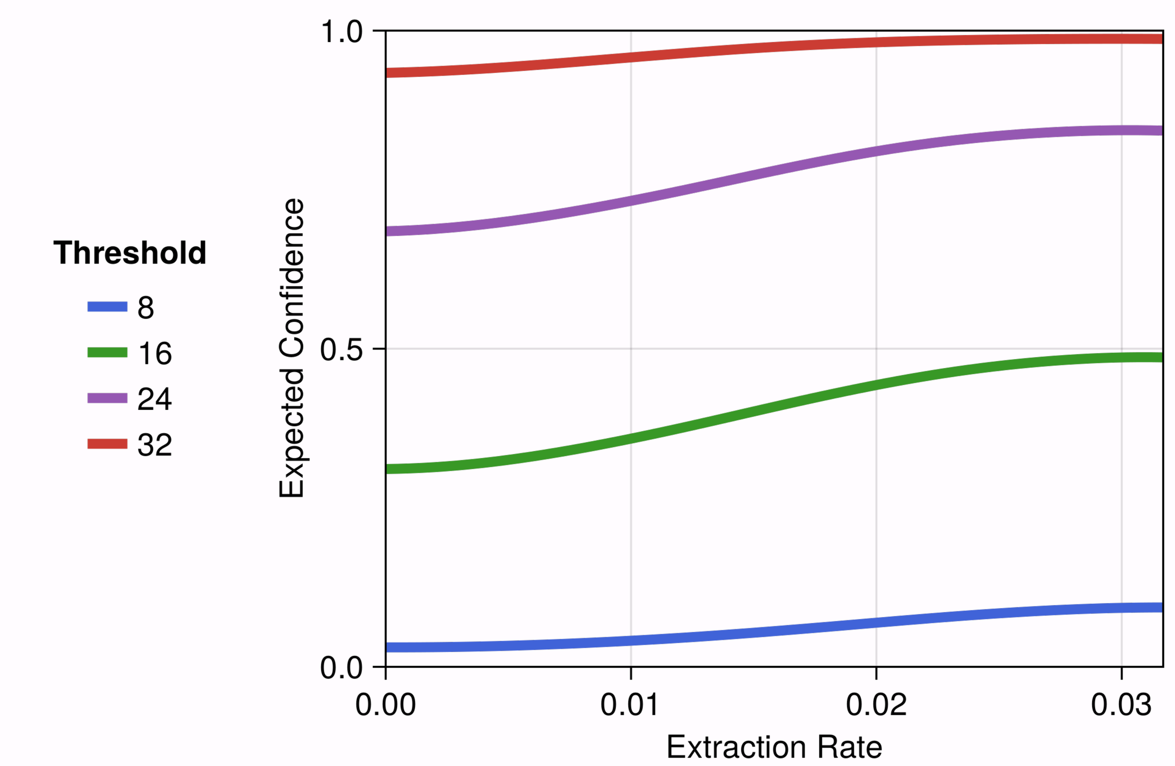
- Want GP fit to large number of samples nodes e.g. $n > 10,000$
- Classic GPs cost $\mathcal{O}(n^3)$ to fit
- Quasi-random nodes with matching kernels produces nice kernel matrices [2, 3]
- Solving systems with these nice kernel matrices costs $\mathcal{O}(n \log n)$

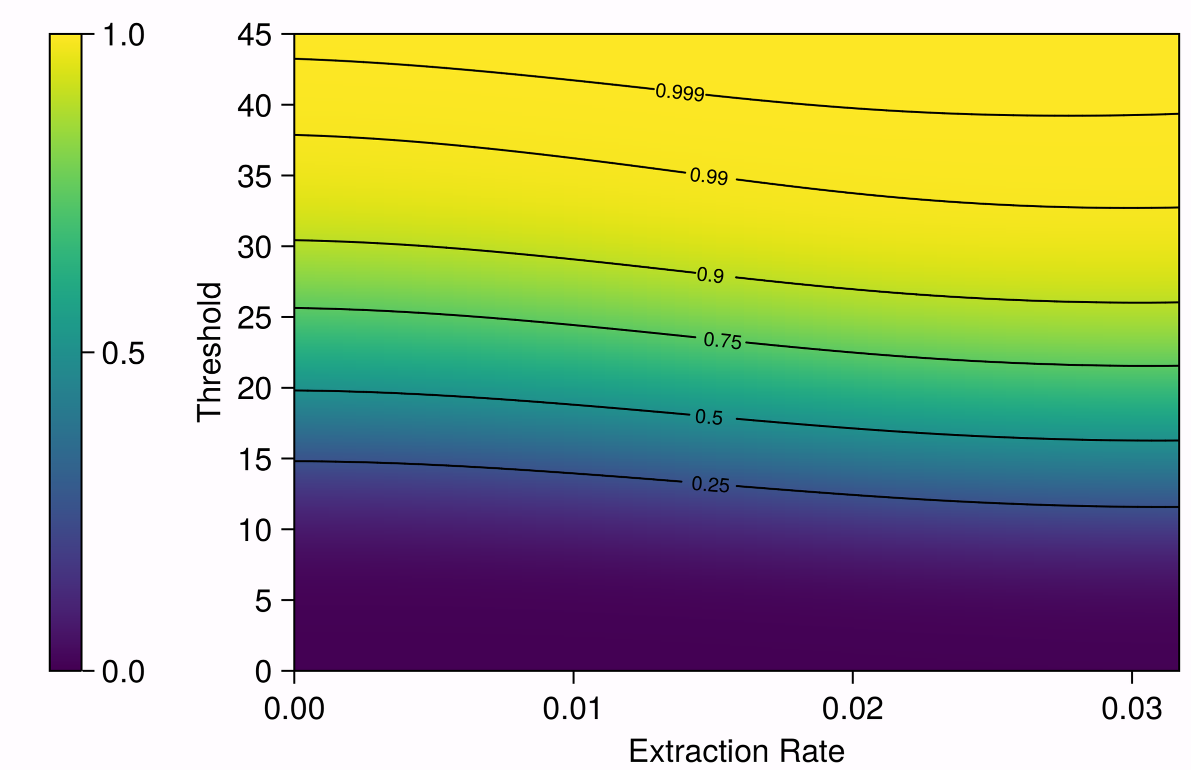
$\therefore$ Can construct GPs with $\mathcal{O}(n \log n)$ cost when we have control over nodes

Calibrate GP Noise Variance $\zeta_{s,d}$ to Match Numerical Error

- Approximate upper bound $\bar{\zeta}_{s,d}$ by tracking convergence as fidelity increases
- Decrease $\zeta_{s,d}$ starting at $\bar{\zeta}_{s,d}$ to optimize Gaussian process likelihood
- Hyperparameter optimization also costs $\mathcal{O}(n \log n)$ for specially structured GPs

Estimate $P(p^c(E,G) \le \bar{p})$ Online with Gaussian Process Surrogate





Takeaways

We fit a Probabilistic Model to a PDE with Random Coefficients

- **Efficient** GP construction in $\mathcal{O}(n \log n)$ instead of $\mathcal{O}(n^3)$ cost
- **Error Aware** GP noise calibrated to numerical error
- **Transferable** to other PDEs with random coefficients

References

[1] *DPFEHM: A Differentiable Subsurface Physics Simulator.* URL: https://github.com/OrchardLANL/DPFEHM.jl.

[2] R. Jagadeeswaran and Fred J. Hickernell. "Fast automatic Bayesian cubature using lattice sampling". In: *Statistics and Computing* 29.6 (2019), pp. 1215–1229. ISSN: 1573-1375. DOI: 10.1007/s11222-019-09895-9. URL: http://dx.doi.org/10.1007/s11222-019-09895-9.

[3] Rathinavel Jagadeeswaran and Fred J Hickernell. "Fast Automatic Bayesian Cubature Using Sobol Sampling". In: *Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer.* Springer, 2022, pp. 301–318.

[4] Aleksandra Pachalieva et al. "Physics-informed machine learning with differentiable programming for heterogeneous underground reservoir pressure management". In: *Scientific Reports* 12.1 (2022), p. 18734.