# Sensitivity analysis, parameter inversion, and design of experiments under uncertainty

Alen Alexanderian

Department of Mathematics
North Carolina State University

October 9, 2020

# Global Sensitivity analysis (GSA)

**model:** $y = f(\boldsymbol{\theta})$

- evaluating $f(\boldsymbol{\theta})$ requires evaluating a mathematical model (ODE, PDE, DAE, Stochastic model, ...)
- $\boldsymbol{\theta} \in \mathbb{R}^p$ vector of uncertain model parameters

**global sensitivity analysis (GSA):**

*find entries of* $\boldsymbol{\theta}$ *that are "most important to model output"*

- various ways of defining parameter importance: variance based, derivative based, moment independent, etc.

**why do GSA?**

- understand the model better
- reduce dimensionality of input parameter (accelerate forward/inverse UQ problems)
- simplify the model ...

# Some standard GSA measures

**variance based:**

$$\mathcal{S}_i = \frac{\mathrm{Var}\left\{\mathbb{E}\left\{f(\boldsymbol{\theta})|\theta_i\right)\right\}\}}{\mathrm{Var}\left\{f(\boldsymbol{\theta})\right\}}, \quad i = 1, \ldots, p$$

estimation requires Monte-Carlo sampling

**derivative based:**

$$\nu_i = \mathbb{E}\Big\{ \left(\frac{\partial f}{\partial \theta_i}\right)^2 \Big\}, \quad i = 1, \ldots, p$$

**and others:** moment independent; Morris screening; activity scores; ...

R.C. Smith, Uncertainty quantification: theory, implementation, and applications. 2013.

# Challenges/opportunities

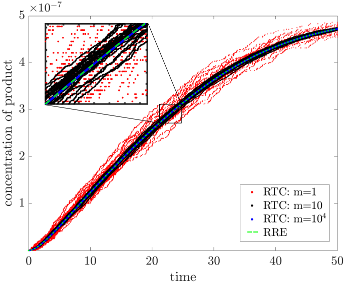
**model:** $y = f(\boldsymbol{\theta})$

- high dimensional $\boldsymbol{\theta}$
- expensive to evaluate $f$: models governed by complex physics systems
- entries of $\boldsymbol{\theta}$ might be correlated
- results depend on how parameter uncertainty is modeled; robustness issues
- $f$ might be time/space dependent $\Rightarrow$ high-dimensional output
- stochasticity in the model, $y = f(\boldsymbol{\theta}, \omega)$; e.g., stochastic compartment models, stochastically forced dynamical systems;
- problems that are defined at multiple scales

**in this talk**:

we will discuss GSA across physical scales

# GSA across scales: application to stochastic chemical systems



# Stochastic chemical kinetics: basic setup

- species $S_1, S_2, \ldots, S_N$
- state vector $\boldsymbol{X}(t) = (X_1(t), \ldots, X_N(t))^T$

  $X_i(t) =$ number of molecules of $i$th species at $t$

- $M$ reactions; each reaction has
  - a stoichiometric (state change) vector $\boldsymbol{\nu}_j$
  - a reaction rate $k_j$ (we model as uncertain)
  - a propensity function $a_j(\boldsymbol{X})$

    $a_j(\boldsymbol{X}(t))dt = \mathrm{Prob}\{\text{reaction } j \text{ occurs in time interval } [t, t+dt)\}$

**example**: two species model

$$S_1 \xrightarrow{k_1} S_2 \qquad S_2 \xrightarrow{k_2} S_1$$

$$a_1(\boldsymbol{X}) = k_1 X_1 \quad \boldsymbol{\nu}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$a_2(\boldsymbol{X}) = k_2 X_2 \quad \boldsymbol{\nu}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

# Multiple scales

- microscale: track number of molecules of each species over time

$$\boldsymbol{X}(t,\omega) \in \mathbb{Z}^N_{\geq 0} \quad \text{(use Stochastic Simulation Algorithm)}$$

- mesoscale: track species evolution via real valued stochastic process (Chemical Langevin Equation)

$$\boldsymbol{Y}(t,\omega) \in \mathbb{R}^N_{\geq 0} \quad \text{(use methods for SDEs)}$$

- macroscale: track species concentration using a system of ODEs (reaction rate equations — RREs)

$$\boldsymbol{y}(t) \in \mathbb{R}^N_{\geq 0} \quad \text{(use methods for ODEs)}$$

# Model problem

**Michaelis-Menten system:**

$$S + E \xrightarrow{k_1} C$$
$$C \xrightarrow{k_2} S + E$$
$$C \xrightarrow{k_3} P + E$$

**species:** $S$ (substrate), $E$ (enzyme), $C$ (complex); $P$ (product).

**uncertain parameters**: $\boldsymbol{\theta} = (k_1, k_2, k_3)^T$

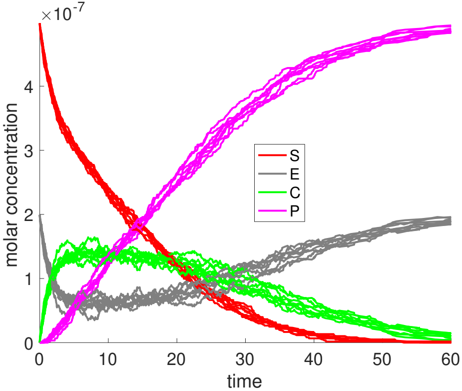
**quantity of interest (QoI)**: average product concentration:

$$f(\boldsymbol{\theta}, \omega) = \frac{1}{T} \int_0^T [P](t, \boldsymbol{\theta}, \omega)\, dt$$

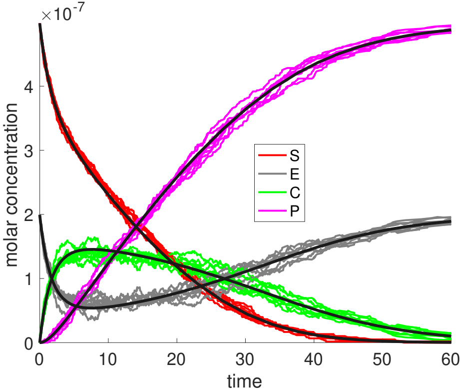
**GSA**: with respect to reaction rates

can we use the macroscale model for efficient GSA of the microscale model?

# Michaelis–Menten simulation



# Michaelis–Menten simulation

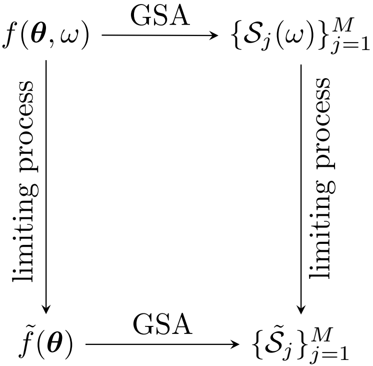



Black lines: solution of RREs

# GSA across scales in a nutshell

**stochastic model:** $y = f(\boldsymbol{\theta}, \omega)$

**for example:** $f(\boldsymbol{\theta}, \omega) = \frac{1}{T}\int_0^T [P](t, \omega, \boldsymbol{\theta})\, dt$




**questions:** does this diagram commute? in what sense? what assumptions are needed?

# Relation between micro and macro scale models

**system size:** $V = \text{Avogadro number} \times \text{system volume}$

**evolution of stochastic species concentrations:**

$$\boldsymbol{Z}^V(t) = \boldsymbol{Z}^V(0) + \sum_{j=1}^{M} \frac{1}{V} \boldsymbol{\nu}_j Y_j \left( V \int_0^t \tilde{a}_j(\boldsymbol{Z}^V(s))\, ds \right)$$

**thermodynamic limit:**

$$\lim_{V\to\infty} \sup_{s\le t} |\boldsymbol{Z}_V(s) - \boldsymbol{y}(s)| = 0 \quad \text{a.s. for all } t > 0$$

where $\boldsymbol{y}(t)$ satisfies the RREs

$$\frac{d\boldsymbol{y}}{dt} = F(\boldsymbol{y}(t)), \quad \boldsymbol{y}(0) = \boldsymbol{y}_0, \quad \text{with } F(\boldsymbol{y}) = \sum_{j=1}^{M} \boldsymbol{\nu}_j \tilde{a}_j(\boldsymbol{y})$$

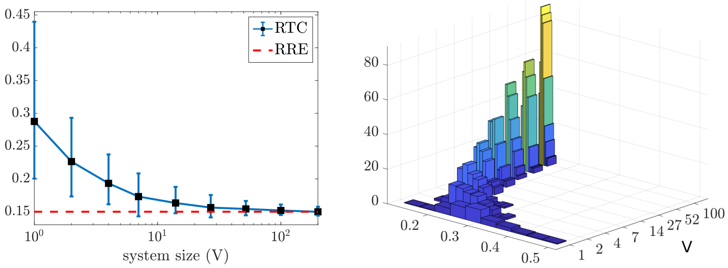
**example of stochastic QoI**: $f(\boldsymbol{\theta}, \omega) = \int_0^T Z_1^V(t, \boldsymbol{\theta}, \omega)\, dt$

**the corresponding "deterministic" QoI:** $\tilde{f}(\boldsymbol{\theta}) = \int_0^T y_1(t, \boldsymbol{\theta})\, dt$

# Numerical results: back to the Michaelis-Menten system

$$S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_3} P + E$$

$$f(\boldsymbol{\theta}, \omega) = \int_0^T [P](\boldsymbol{\theta}, \omega, t)\, dt \quad \boldsymbol{\theta} = (k_1, k_2, k_3)^T$$

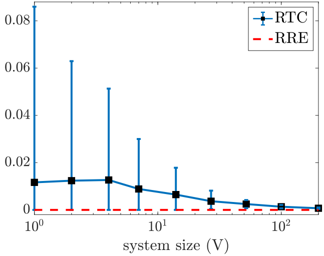


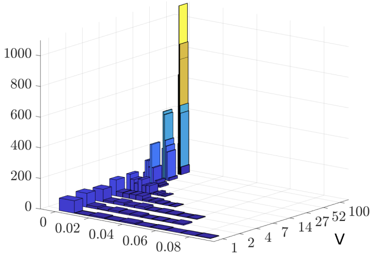

convergence of $T_1(\omega)$ in distribution for increasing $V$
($T_1$: total Sobol' index corresponding to $k_1$)

# Numerical results: back to the Michaelis-Menten system

$$S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_3} P + E$$

$$f(\boldsymbol{\theta}, \omega) = \int_0^T [P](\boldsymbol{\theta}, \omega, t)\, dt \quad \boldsymbol{\theta} = (k_1, k_2, k_3)^T$$



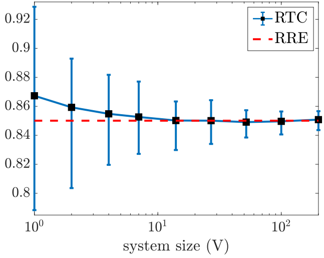



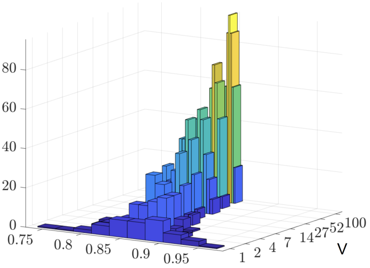

convergence of $T_2(\omega)$ in distribution for increasing $V$

# Numerical results: back to the Michaelis-Menten system

$$S + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C \xrightarrow{k_3} P + E$$

$$f(\boldsymbol{\theta}, \omega) = \int_0^T [P](\boldsymbol{\theta}, \omega, t)\, dt \quad \boldsymbol{\theta} = (k_1, k_2, k_3)^T$$




convergence of $T_3(\omega)$ in distribution for increasing $V$

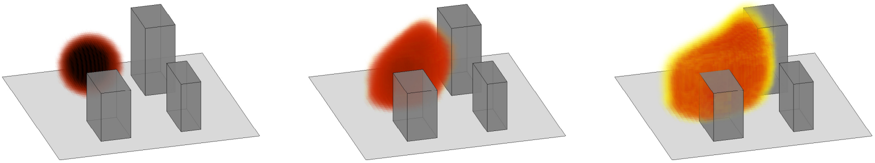
# GSA across scales: summary

- can use the (deterministic) RREs for efficient GSA of stochastic chemical system
- numerical illustrations on biochemical reaction networks
- analysis of convergence
- many interesting remaining questions ...

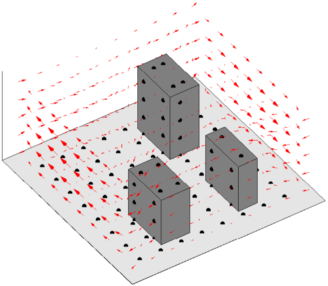
M. Merritt, A. Alexanderian, and P.A. Gremaud, Multiscale global sensitivity analysis for stochastic chemical systems. SIAM Journal on Multiscale Modeling and Simulation, under revision. 2020.

**Inverse problems and optimal design of experiments**

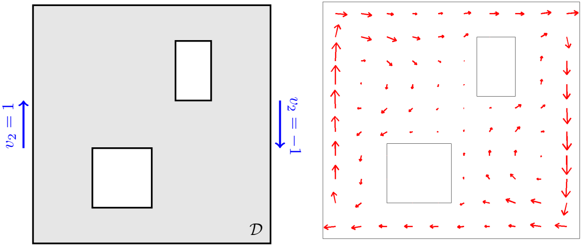
# Example: Diffusive transport of a contaminant with uncertain initial condition



- **Governing PDE** (forward model): advection-diffusion equation
- **Unknown/uncertain parameter**: initial concentration field
- **Inverse problem**: Use a vector $\boldsymbol{d}$ of space/time sensor measurements of concentration to reconstruct the initial state



# 2D Model problem

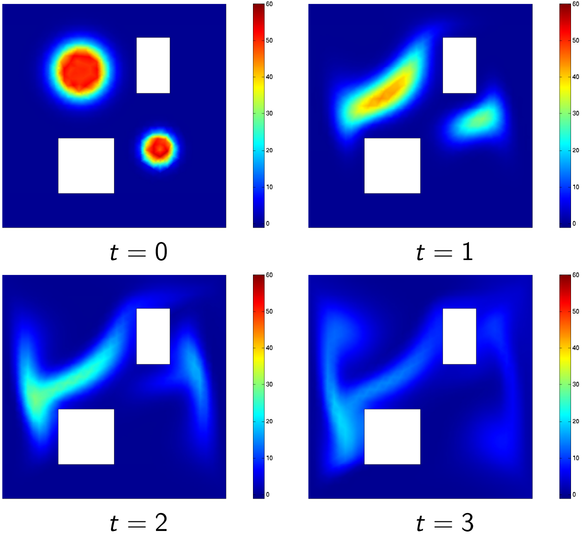



- Forward problem: time dependent advection-diffusion

$$
\begin{aligned}
u_t - \kappa \Delta u + \boldsymbol{v} \cdot \nabla u &= 0 && \text{in } \mathcal{D} \times [0, T] \\
u(0, \boldsymbol{x}) &= m && \text{in } \mathcal{D} \\
\kappa \nabla u \cdot \boldsymbol{n} &= 0 && \text{on } \partial\mathcal{D} \times [0, T]
\end{aligned}
$$

- $m$: *unknown* initial condition
- $\boldsymbol{v}$: velocity field

# Solution of the forward problem



# The inverse problem: reconstruct initial condition

The inverse problem of finding the unknown initial state based on sensor data

$$\min_m \frac{1}{2}\|\mathcal{B}u(m) - \boldsymbol{d}\|^2 + \frac{\alpha}{2}\langle \mathcal{A}m, m\rangle$$

where

$$\begin{aligned} u_t - \kappa\Delta u + \mathrm{v}\cdot\nabla u &= 0 \qquad \text{in } \mathcal{D}\times[0,T] \\ u(0,\boldsymbol{x}) &= m \qquad \text{in } \mathcal{D} \\ \kappa\nabla u\cdot\boldsymbol{n} &= 0 \qquad \text{on } \partial\mathcal{D}\times[0,T] \end{aligned}$$

- $\mathcal{B}$: observation operator
- $\boldsymbol{d} = [\boldsymbol{d}_1^T \boldsymbol{d}_2^T \cdots \boldsymbol{d}_{n_t}^T]^T$, $\quad \boldsymbol{d}_i \in \mathbb{R}^{n_s}$, $n_s =$ number of sensors
- $u$ linear in $m$, $u = \mathcal{S}m \Longrightarrow$ linear parameter-to-observable map: $\mathcal{F} = \mathcal{B}\mathcal{S}$
- Can rewrite the optimization problem as

$$\min_m \mathcal{J}(m) := \frac{1}{2}\|\mathcal{F}m - \boldsymbol{d}\|^2 + \frac{\alpha}{2}\langle \mathcal{A}m, m\rangle$$

# Solving the inverse problem

- Derivative of $\mathcal{J}$

$$
\begin{aligned}
D\mathcal{J}(m)(\tilde{m}) &= \frac{d}{d\varepsilon}\mathcal{J}(m+\varepsilon\tilde{m})\,|_{\varepsilon=0} \\
&= \langle \mathcal{F}^*(\mathcal{F}m - \boldsymbol{d}) + \alpha\mathcal{A}m, \tilde{m}\rangle
\end{aligned}
$$

- Action of $\mathcal{F}^*$
  $\mathcal{F}^*\boldsymbol{y} = p(\cdot, 0)$, where $p$ is solution of the adjoint equation

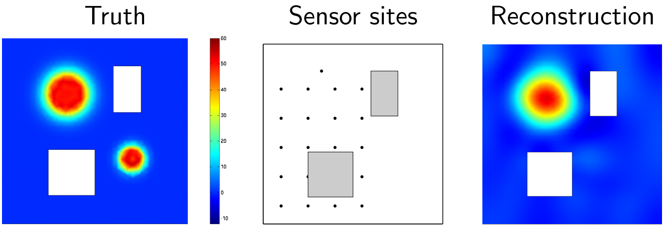
$$
\begin{aligned}
-p_t - \nabla\cdot(p\boldsymbol{v}) - \kappa\Delta p &= -\mathcal{B}^*\boldsymbol{y} \\
p(T) &= 0 \\
(\boldsymbol{v}p + \kappa\nabla p)\cdot\boldsymbol{n} &= 0
\end{aligned}
$$

- Optimality condition

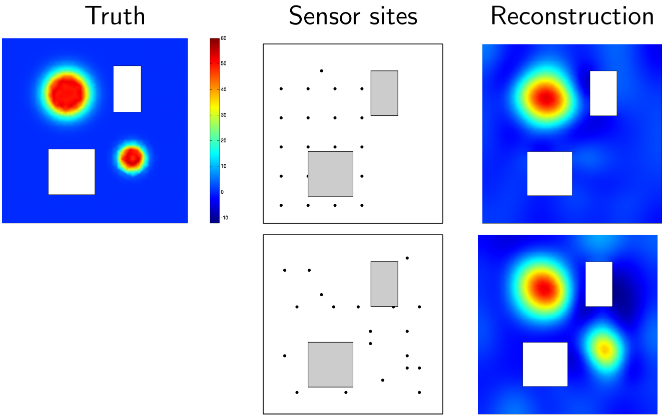
$$
(\mathcal{F}^*\mathcal{F} + \alpha\mathcal{A})m = \mathcal{F}^*\boldsymbol{d} \quad \overset{\textit{discretize}}{\Longrightarrow} \quad (\mathsf{F}^*\mathsf{F} + \alpha\mathsf{A})\boldsymbol{m} = \mathsf{F}^*\boldsymbol{d}
$$

  Solve the linear system using an iterative method, e.g. conjugate gradient

# Solving the inverse problem: numerical results



# Solving the inverse problem: numerical results



# Optimal sensor placement as an optimal design problem

How to place sensors in an "optimal" way?

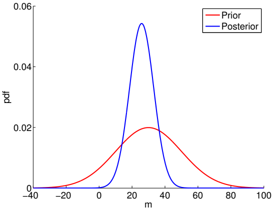
- Can formulate the optimal sensor placement problem as an optimal experimental design (OED) problem
- Can consider a statistical formulation of the inverse problem
- In addition to a reconstruction, we can also compute a statistical distribution of the parameters, conditioned on experimental data
- Find sensor locations so as to optimize the statistical quality of the reconstructed/inferred parameter
- In context of inverse problems a Bayesian formulation is natural

# Bayesian inference: Bayes' formula

$$\pi_{\text{post}}(m|\boldsymbol{d}) \propto \pi_{\text{like}}(\boldsymbol{d}|m)\pi_{\text{prior}}(m)$$

| | |
|---|---|
| $\pi_{\text{post}}(m|\boldsymbol{d})$ | posterior pdf of $m$ |
| $\pi_{\text{like}}(\boldsymbol{d}|m)$ | pdf of $\boldsymbol{d}$ given $m$ (data likelihood) |
| $\pi_{\text{prior}}(m)$ | prior pdf of $m$ |

pdf = probability density function




*Rev. Thomas Bayes*

*Pierre-Simon Laplace*

Bayes, T., An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, FRS Communicated by Mr. Price, in a Letter to John Canton, AMFRS. Philosophical Transactions, 1763.

Laplace, P.S., Théorie analytique des probabilités. 1820.

# Bayesian linear inverse problems

Assume linear parameter-to-observable map:

$$\boldsymbol{d} = \mathsf{F}\boldsymbol{m} + \boldsymbol{\eta}$$

and assume prior is Gaussian

$$\pi_0(\boldsymbol{m}) \propto \exp(-\tfrac{1}{2}\boldsymbol{m}^T \Gamma_{\text{prior}}^{-1} \boldsymbol{m})$$

Then, the posterior pdf is

$$\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}) \propto \exp\left\{-\tfrac{1}{2}(\boldsymbol{m} - \boldsymbol{m}_{\text{MAP}})^T (\mathsf{F}^T \Gamma_{\text{noise}}^{-1} \mathsf{F} + \Gamma_{\text{prior}}^{-1})(\boldsymbol{m} - \boldsymbol{m}_{\text{MAP}})\right\}$$

$$\Rightarrow \mu_{\text{post}} = \mathcal{N}\left(\boldsymbol{m}_{\text{MAP}}, \Gamma_{\text{post}}\right)$$

$$\Gamma_{\text{post}}^{-1} = \underbrace{\mathsf{F}^T \Gamma_{\text{noise}}^{-1} \mathsf{F}}_{\mathsf{H}_{\text{misfit}}} + \Gamma_{\text{prior}}^{-1} \quad (= D^2_{\boldsymbol{m}}(-\log \pi_{\text{post}}))$$

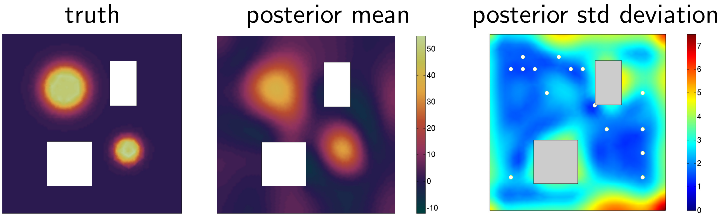
$$\boldsymbol{m}_{\text{MAP}} = \arg\min_{\boldsymbol{m}} \frac{1}{2}\|\mathsf{F}\boldsymbol{m} - \boldsymbol{d}\|^2_{\Gamma_{\text{noise}}^{-1}} + \frac{1}{2}\left\langle \Gamma_{\text{prior}}^{-1}\boldsymbol{m}, \boldsymbol{m}\right\rangle$$

An important problem structure:

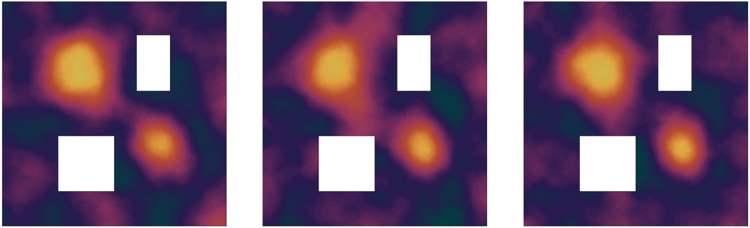
$\mathsf{H}_{\text{misfit}}$ is low rank, and $\Gamma_{\text{prior}}^{1/2}\mathsf{H}_{\text{misfit}}\Gamma_{\text{prior}}^{1/2}$ is even more so ...

# Bayesian inversion of the initial condition for 2D advection-diffusion equation

- Posterior mean, and posterior variance

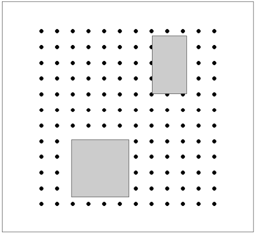



- Posterior samples: $\boldsymbol{\nu} = \boldsymbol{m}_{\mathrm{MAP}} + \Gamma_{\mathrm{post}}^{1/2} \mathsf{M}^{-1/2} \boldsymbol{n}, \qquad \boldsymbol{n} \sim \mathcal{N}(0, \mathsf{I})$



# The optimal experimental design problem

A grid of candidate locations for observation points



- **Experimental design**: locations of observation points / sensors

$$\text{design} := \left\{ \begin{array}{c} \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_s} \\ w_1, \ldots, w_{N_s} \end{array} \right\}$$

- **Bayesian inversion**:

  data + likelihood, prior $\Longrightarrow$ posterior distribution of inversion parameter

- **Optimal experimental design (OED)**:

  Find sensor locations that result in minimized posterior uncertainty

# Commonly used OED criteria

Bayesian A-optimal experimental design:

$$\underset{\boldsymbol{w}\in S}{\text{minimize}}\, \mathrm{tr}\big[\Gamma_{\mathrm{post}}(\boldsymbol{w})\big] + \gamma P(\boldsymbol{w}) \tag{1}$$

Bayesian D-optimal experimental design:

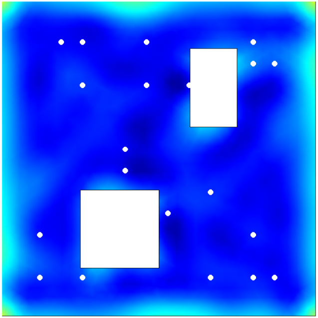
$$\underset{\boldsymbol{w}\in S}{\text{minimize}} -\frac{1}{2}\log\det(\mathsf{I} + \Gamma_{\mathrm{prior}}^{1/2}\mathsf{H}_{\mathsf{misfit}}(\boldsymbol{w})\Gamma_{\mathrm{prior}}^{1/2}) + \gamma P(\boldsymbol{w}) \tag{2}$$

- Need trace/log-determinant of high-dimensional operators
- Need many applications of the forward operator $\implies$ many PDE solves
- OED much harder for nonlinear inverse problems
- Ingredients of efficient OED methods: randomized matrix methods, use of problem structure, low-rank approximations, iterative solvers, gradient based optimization algorithms, adjoint based gradient/Hessian computation, sparsifying penalty methods, ...

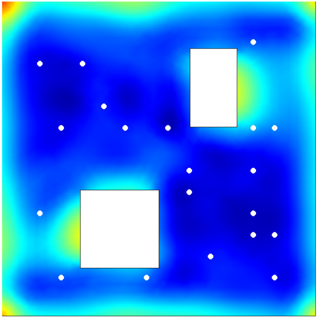
*A.K. Saibaba, A. Alexanderian, and I.C. Ipsen. Randomized matrix-free trace and log- determinant estimators. Numerische Mathematik, 2017.*

*A. Alexanderian and A. Saibaba, Efficient D-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems. SIAM Journal on Scientific Computing. 2018.*

# A-optimal design: the variance field

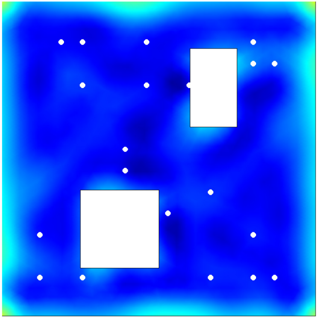


Optimal

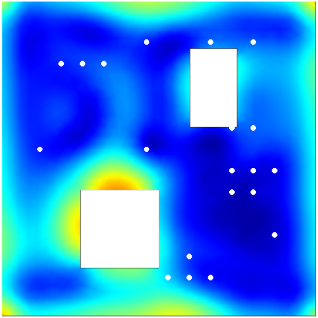


Sub-optimal

# A-optimal design: the variance field

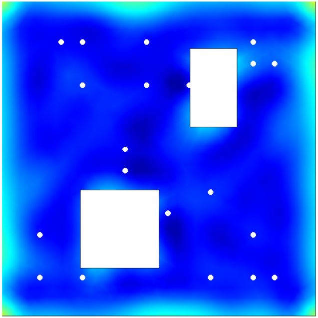


Optimal

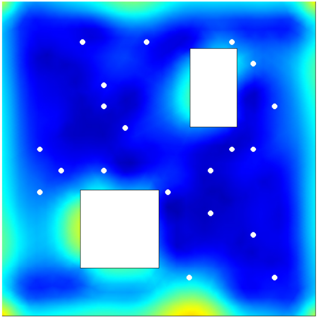


Sub-optimal

# A-optimal design: the variance field

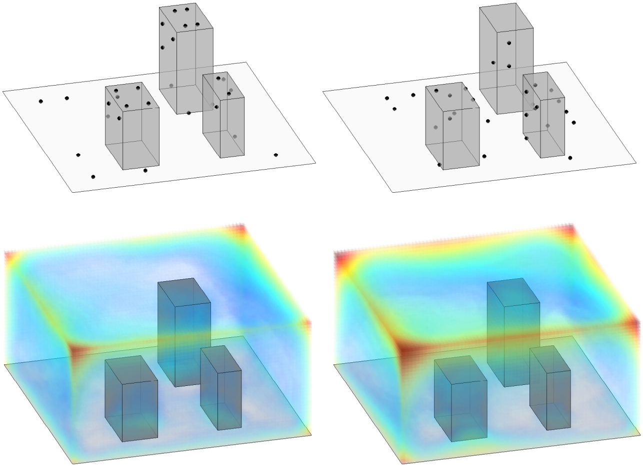


Optimal



Sub-optimal

# OED for 3D model (parameter dim $\sim 10^4$)



# Summary and outlook: Global sensitivity analysis

**Global sensitivity analysis**: apportion uncertainty in the model output to different sources of uncertainty in the model input parameters

- Interesting directions: theory and computational methods for global sensitivity analysis with high-dimensional inputs/outputs; GSA for stochastic models; GSA across scales; **sensitivity analysis of inverse problems**
- Applications: flow through porous media, contaminant transport, radioactive waste storage, biotransport in cancerous tumors, chemical kinetics, biochemistry, epidemiology, pharmacokinetics ...
- NCSU faculty collaborators: Pierre Gremaud, Ralph Smith
- Students: Helen Cleaves, Mike Merritt, Isaac Sunseri

# Summary and outlook: Optimal design of inverse problems

Optimal placement for Bayesian inverse problems: find optimal placements of measurement point to minimize the uncertainty in reconstructed parameters

- Interesting research directions: scalable algorithms for sensor placement for linear inverse problems governed by PDEs (randomized methods in numerical linear algebra, low-rank approximations, optimization with exact penalty method, ...); Optimal sensor placement for nonlinear inverse problems, (bi-level PDE-constrained optimization, adjoint based derivative computation, ... ), optimal experimental design under model uncertainty (marginalization, stochastic optimization, ...); sequential design of experiments
- Applications: porous medium flow, contaminant source identification, radiation detection in urban environments

- NCSU faculty collaborators: Arvind Saibaba, Ralph Smith
- Students: Isaac Sunseri, Bekah White

# Relevant courses

| | |
|---|---|
| MA 580 - 780 | Numerical Analysis |
| MA 515 - 715 | Analysis |
| MA 573 - 574 | Mathematical Modeling |
| MA 534 - 734 | Partial Differential Equations |
| MA 546 - 747 | Applied Probability and Stochastic Processes |
| MA 540 | Uncertainty Quantification |
| MA 587 | Numerical Solution of PDEs – Finite Element Method |
| MA 798 | Inverse Problems (offered this spring) |

# Some references: sensitivity analysis


- M. Merritt, A. Alexanderian, and P.A. Gremaud, Multiscale global sensitivity analysis for stochastic chemical systems. SIAM Journal on Multiscale Modeling and Simulation, under revision. 2020.
- A. Alexanderian, P. Gremaud, and R.C. Smith. Variance-based sensitivity analysis for time-dependent processes. Reliability Engineering & System Safety. 2020.
- H. Cleaves, A. Alexanderian, H. Guy, R.C. Smith, and M. Yu. Derivative based global sensitivity analysis for models with high-dimensional inputs and functional outputs. SIAM Journal on Scientific Computing. 41(6):A3524–A3551, 2019.
- Manav Vohra, Alen Alexanderian, Hayley Guy, and Sankaran Mahadevan. Active Subspace-based dimension reduction for chemical kinetics applications with epistemic uncertainty. Combustion and Flame, 204:152–161, 2019.
- J. Hart, A. Alexanderian, and P. Gremaud. Efficient computation of Sobol indices for stochastic models. SIAM Jounral on Scientific Computing. 39(4). 2017.

# Some references: OED and inverse problems


- Alen Alexanderian Optimal Experimental Design for Bayesian Inverse Problems Governed by PDEs: A Review. In revision, 2020. https://arxiv.org/abs/2005.12998
- A. Alexanderian, N. Petra, G. Stadler, and I. Sunseri. Optimal design of large-scale Bayesian linear inverse problems under reducible model uncertainty: good to know what you don't know. In review, 2020. https://arxiv.org/abs/2006.11939
- I. Sunseri, J. Hart, B. van Bloemen Waanders, A. Alexanderian. Hyper-Differential Sensitivity Analysis for Inverse Problems Constrained by Partial Differential Equations. Inverse Problems, Accepted, 2020. https://arxiv.org/abs/2003.00978
- A. Alexanderian and A. Saibaba, Efficient D-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems. SIAM Journal on Scientific Computing. 2018.
- A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. A fast and scalable method for A-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems. SISC, 2016.
- A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. A-optimal design of experiments for infinite-dimensional Bayesian linear inverse problems with regularized $\ell^0$-sparsification. SISC. 2014.