Physics-constrained Bayesian inference of an uncertain operator in the sparse-data regime

Teresa Portone (Sandia National Labs),

Bob Moser (UT Austin)

April 12-15, 2022, SIAM UQ 22, Atlanta, GA

Sandia National Laboratories

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2022-4030 C

Bayesian operator inference with limited data

- Common operator inference approaches require copious observations of all state variables as a function of time, for different initial conditions/model parameterizations, etc.

- What can be inferred about an operator for realistic physical applications where limited observational data is available?

- Specifically, about an infinite-dimensional operator appearing in a PDE with only
  - Time-series observations at one location?
  - Spatial-series observations at one time?

- We pose a Bayesian inverse problem for an infinite-dimensional operator.
- Limited observational data augmented by imposing physical constraints on operator and encoding qualitative information through the prior distribution.

Testbed problem – Field-scale contaminant transport

$$\frac{\partial c(\boldsymbol{x},t)}{\partial t} + \nabla \cdot \big(\boldsymbol{u}(\boldsymbol{x})c(\boldsymbol{x},t)\big) = \nu_p \Delta c(\boldsymbol{x},t), \qquad \boldsymbol{x} \equiv (x,y) \in [0,L_x]\times\big[0,L_y\big]$$

$$\boldsymbol{u}(\boldsymbol{x}) = -\kappa(\boldsymbol{x})\nabla p(\boldsymbol{x})$$

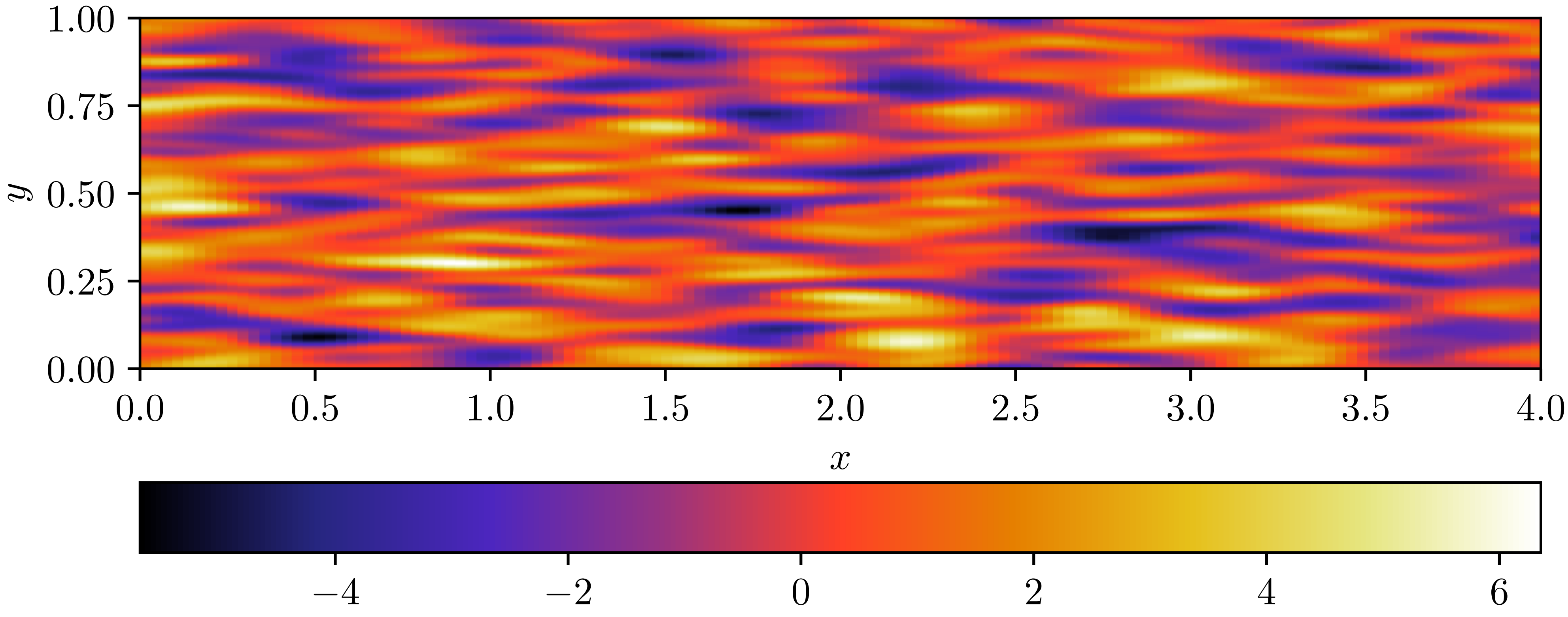
$$\nabla \cdot \big(\boldsymbol{u}(\boldsymbol{x})\big) = 0$$

Periodic in *x*, zero Neumann in *y*

Problem: Can’t observe $\kappa(\boldsymbol{x})$ over entire domain.

Permeability fields highly heterogeneous, observed to vary over several orders of magnitude.

$$\ln(\kappa) \sim \mathcal{N}\left(0, \sigma^2 \exp\left(-\frac{1}{2}\left[\frac{(x-x\prime)^2}{\ell_x^2} + \frac{(y-y\prime)^2}{\ell_y^2}\right]\right)\right), \sigma^2 = 3.04,\ \ell_x = 0.09,\ \ell_y = 0.04$$



However, (it is often assumed) their statistics are homogeneous.

- Could sample from κ distribution and solve the detailed model many times to predict mean transport behavior.

- For field-scale transport in realistic problems this can be computationally challenging.

- Often the equations are statistically averaged to predict mean transport behavior.

- We take this approach here, additionally averaging in depthwise direction.
  - Observations limited to depthwise averaged concentrations (mixing when drawing fluid from well)
  - Streamwise variation in contaminant concentration the relevant quantity of interest.

$$\langle f(x,y)\rangle \equiv \frac{1}{L_y}\int_0^{L_y} \mathbb{E}_\kappa[f(x,y)]dy \qquad (1)$$

$$f(x,y) = \langle f(x,y)\rangle + f'(x,y) \qquad (2)$$

Represent $c, \boldsymbol{u}$ using (2), apply (1) to the ADE to get

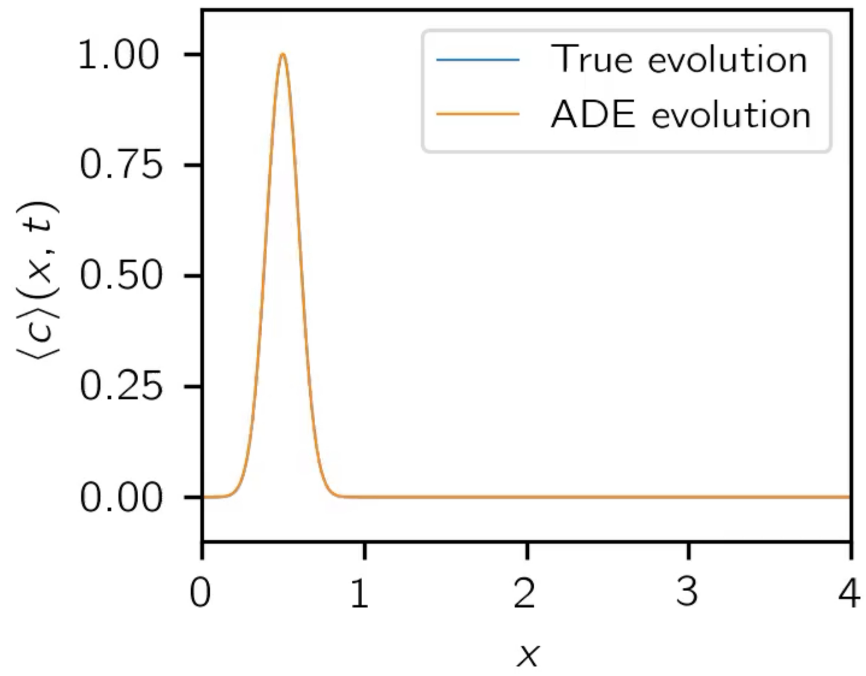
$$\frac{\partial\langle c\rangle(x,t)}{\partial t} + \langle u\rangle \frac{\partial\langle c\rangle(x,t)}{\partial x} + \frac{\partial\langle u'c'\rangle}{\partial x} = \nu_p \frac{\partial^2\langle c\rangle(x,t)}{\partial x^2}$$

Can't observe $\langle u'c'\rangle \Rightarrow$ dependence uncertain.

Typical closure model for $\langle u'c' \rangle$ is gradient-diffusion, but it is known to be inadequate for heterogeneous porous media.

True evolution: mean transport through heterogeneous media.

ADE evolution: predicted mean transport using a gradient-diffusion closure.



Uncertain generalized diffusion operator

Instead of gradient-diffusion, infer an uncertain operator acting on $\langle c \rangle$ to represent dispersion:

$$\mathcal{L}\langle c \rangle = -\frac{\partial \langle u' c' \rangle}{\partial x}$$

- Pose Bayesian inverse problem for $\mathcal{L}$
- Impose physical constraints deterministically and through prior probability distribution.

Physical constraints

$$\frac{\partial \langle c \rangle (x,t)}{\partial t} + \langle u \rangle \frac{\partial \langle c \rangle (x,t)}{\partial x} = \nu_p \frac{\partial^2 \langle c \rangle (x,t)}{\partial x^2} + \mathcal{L}\langle c \rangle (x,t)$$

Linearity in $\langle c \rangle$	$\Rightarrow$	$\mathcal{L}$ linear, $\mathcal{L} f_k = \lambda_k f_k$
Shift invariance	$\Rightarrow$	$f_k = \exp(i a_k x)$, $\mathcal{L} = \sum_k \langle \hat{c}_k \rangle \lambda_k \exp(i a_k x)$, $a_k = \frac{2\pi k}{L_x}$
Conservation of mass	$\Rightarrow$	$\lambda_0 = 0$
$\langle c \rangle (x, t = \infty) = C$	$\Rightarrow$	$-\nu_p a_k^2 + \mathcal{R}[\lambda_k] < 0$
Solution propagates downstream	$\Rightarrow$	$\langle u \rangle a_k - \mathcal{I}[\lambda_k] > 0$

Bayesian inverse problem for $\mathcal{L}$

$$p(\boldsymbol{\lambda}|d) \propto p(d|\boldsymbol{\lambda})p(\boldsymbol{\lambda})$$

Prior specification

$$\lambda_k = R_k + iI_k$$

$$\tilde{R}_k \equiv \nu_p a_k^2 - R_k > 0 \quad \Rightarrow \quad \tilde{R}_k \sim \gamma_k^R \exp\left(-\gamma_k^R(\tilde{R}_k)\right)$$

$$\tilde{I}_k \equiv \langle u \rangle a_k - I_k > 0 \quad \Rightarrow \quad \tilde{I}_k \sim \gamma_k^I \exp\left(-\gamma_k^I(\tilde{I}_k)\right)$$

γ's defined so 95% probability mass falls within plausible bounds on solution behavior:

- Solution doesn't decay too rapidly
- Dispersion's contribution to transport doesn't exceed that of advection

$$p(\boldsymbol{\lambda}) = \prod_k p(R_k)p(I_k)$$

Bayesian inverse problem dependence on data

Want to understand

- How solution to inverse problem depends on amount and type of data
- If $\mathcal{L}$ could be successfully inferred with limited data.

Performed inference with data generated from known $\mathcal{L}$ to understand this.

Used Kullback-Leibler divergence to measure information gain from inference:

$$D(p(\boldsymbol{\lambda}|\boldsymbol{d}) \,||\, p(\boldsymbol{\lambda})) = \int p(\boldsymbol{\lambda}|\boldsymbol{d}) \log \frac{p(\boldsymbol{\lambda}|\boldsymbol{d})}{p(\boldsymbol{\lambda})} d\boldsymbol{\lambda}$$

Likelihood specification and data scenarios

Likelihood specification:

$$d_i = \langle c \rangle(x_i, t_i; \boldsymbol{\lambda}) + \epsilon_m, \qquad \epsilon_m \sim \mathcal{N}(0, \Sigma), \qquad i = 1, \dots, N_{obs} = N_x \cdot N_t$$

$$p(\boldsymbol{d}|\boldsymbol{\lambda}) = |2\pi\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|\boldsymbol{d} - \langle \boldsymbol{c} \rangle\|^2_{\Sigma^{-\frac{1}{2}}}\right)$$

Inference with known $\mathcal{L}$ data scenarios:

$\Sigma = (0.005)^2 I$

Spatial-series data: $N_t = 1,\ N_x = 32, 64, 512$

Time-series data: $N_x = 1,\ N_t = 32, 64, 512$

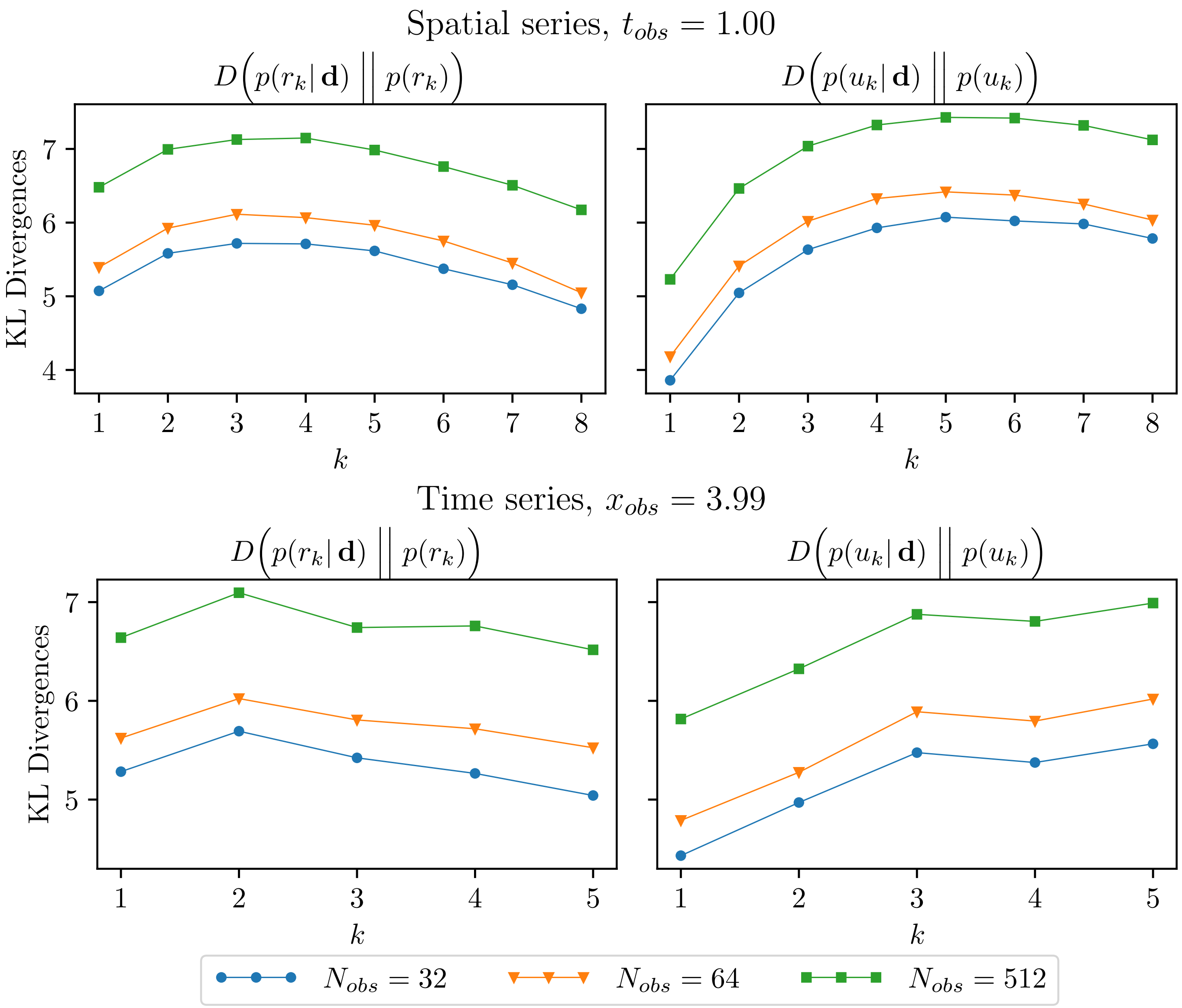
Inverse problem dimensionality

Dimensionality of inverse problem doesn't depend on spatial discretization; instead, it depends on how many eigenvalues $\langle c \rangle$ is sensitive to.

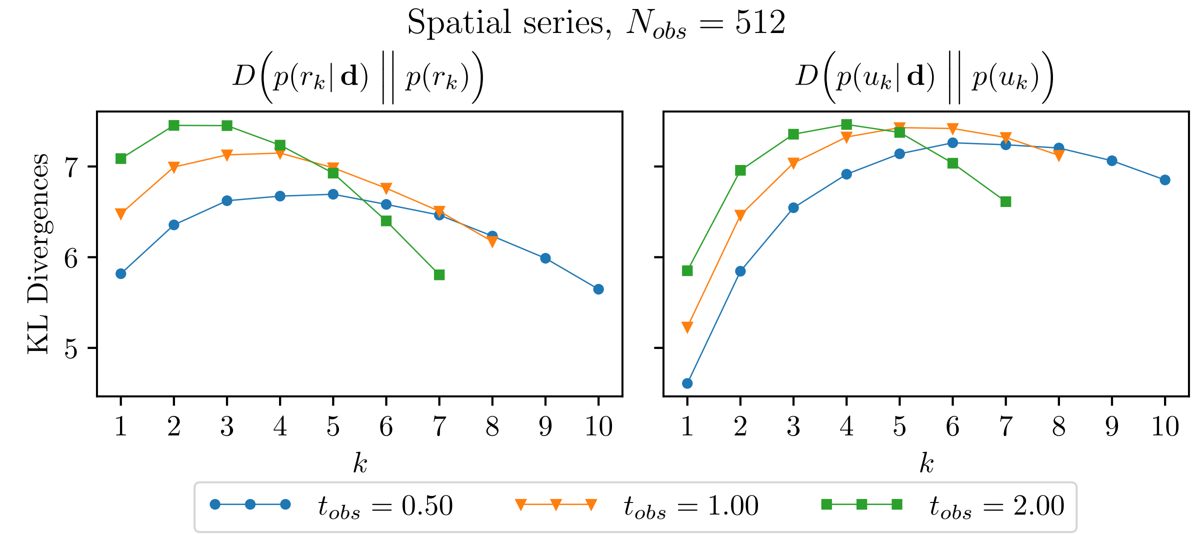
Determined sensitive eigenvalues using global sensitivity analysis for each data scenario.

Spatial series		**Time series**	
Observation time	# sensitive eigenvalues	Observation location	# sensitive eigenvalues
0.5	10	2.0	5
1.0	8	3.0	5
2.0	7	4.0	5

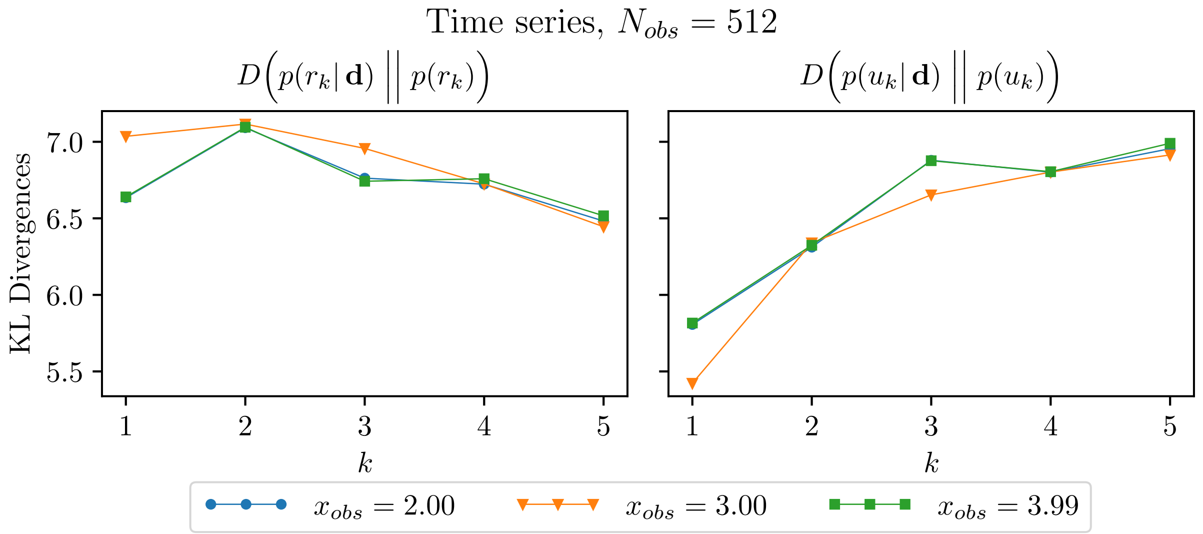
Information gain increases with observation frequency



For spatial-series observations of fixed frequency, information gain and number of eigenvalues informed depended on observation time.

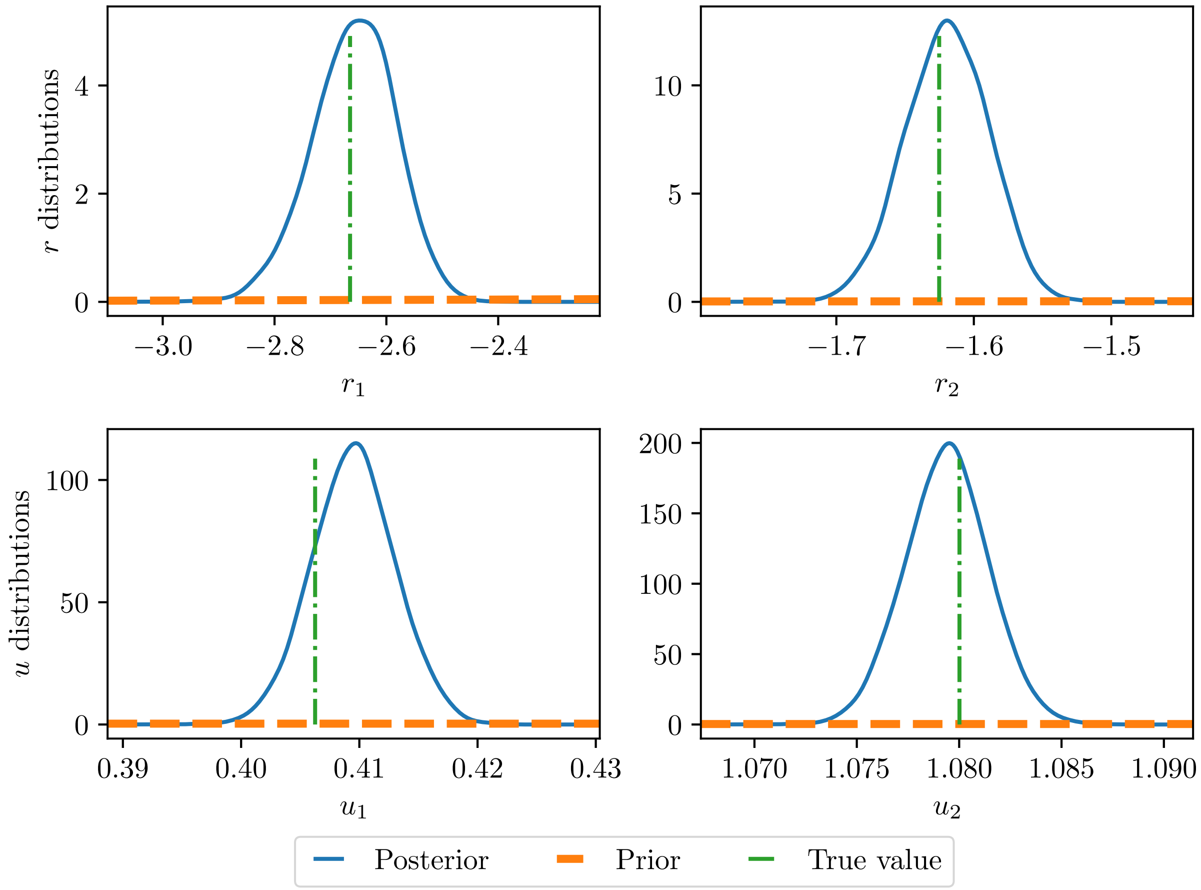


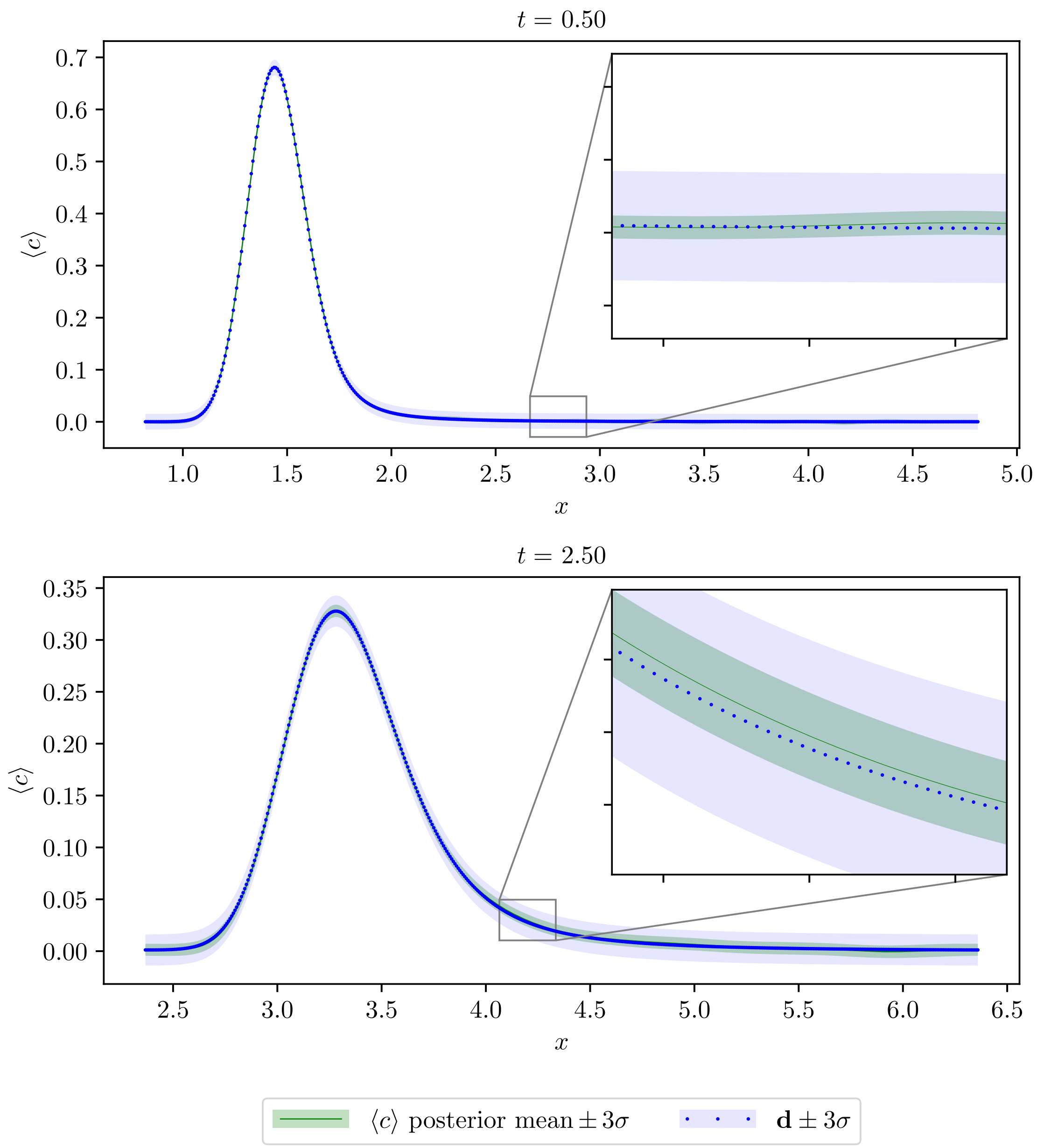
Information gain and number of informed eigenvalues were not similarly dependent on observation location for time-series observations.



Inference successful for copious spatial data

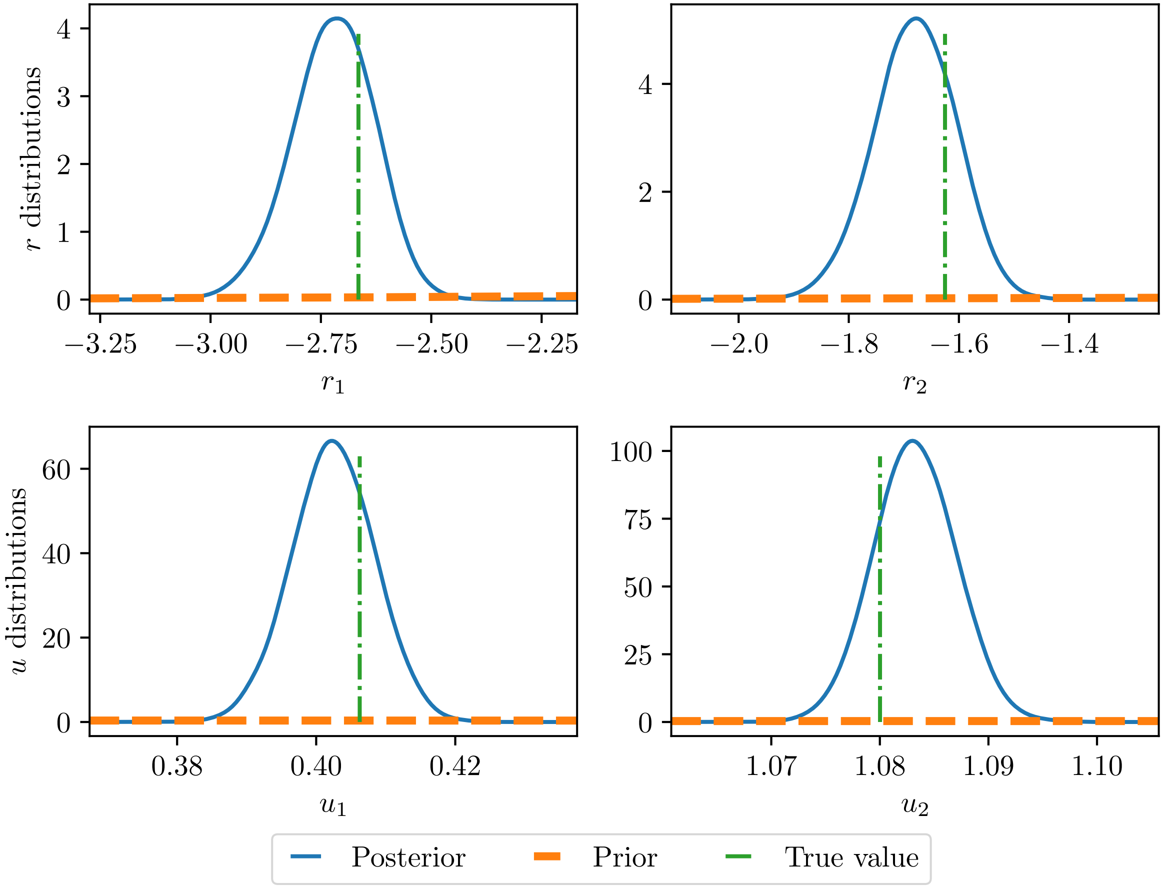
Data: 512 spatial observations taken at $t = 0.5$

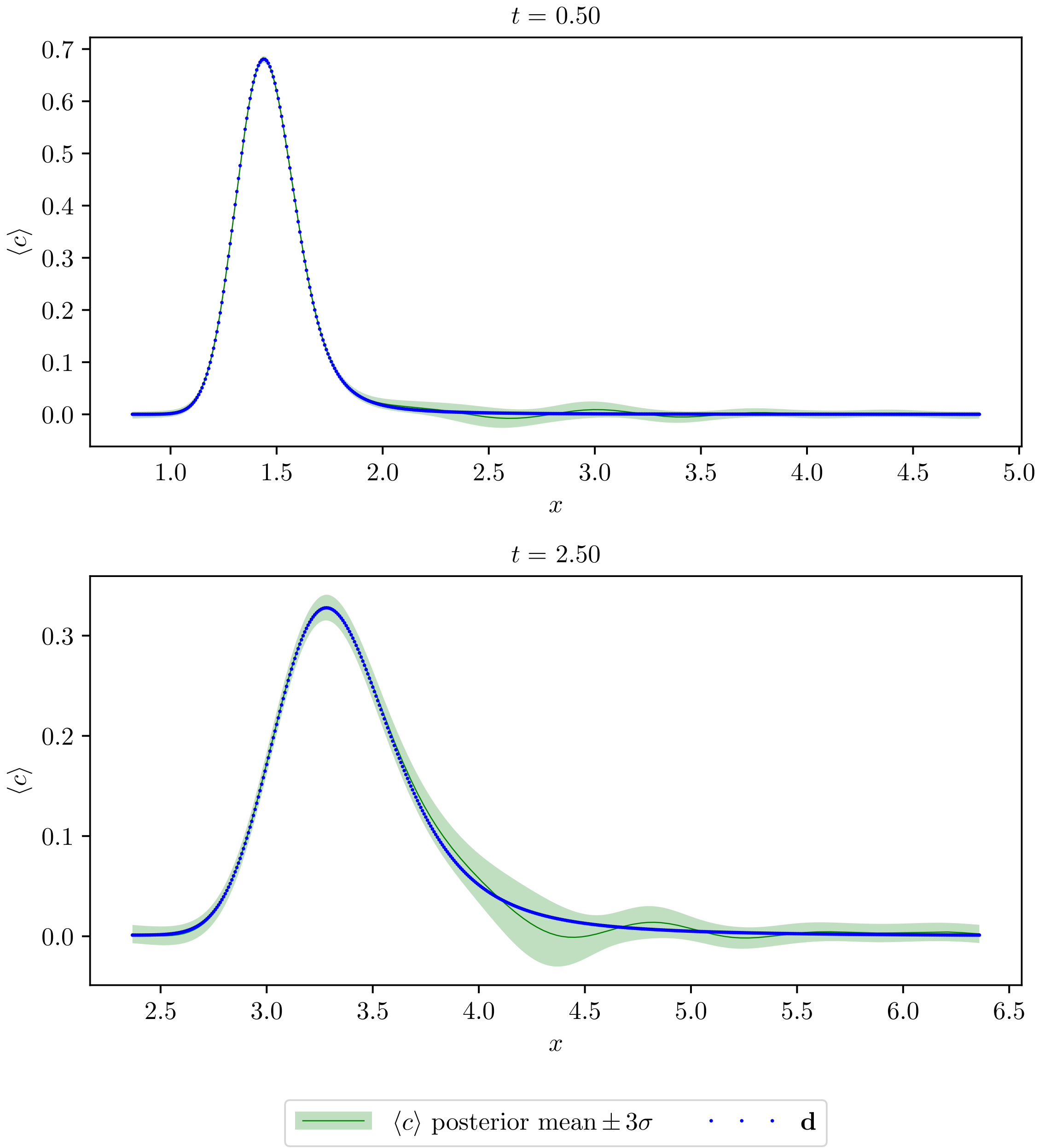




Inverse problem solution with sparse time-series data can be nonphysical

Data: 32 time observations uniformly spaced over [0,4], taken at $x = \frac{L_x}{2} = 2$.





Bayesian inference with unknown $\mathcal{L}$

Can we infer the eigenvalues of $\mathcal{L}$ using observations of $\langle c \rangle$ from a direct numerical simulation of 2D ADE?

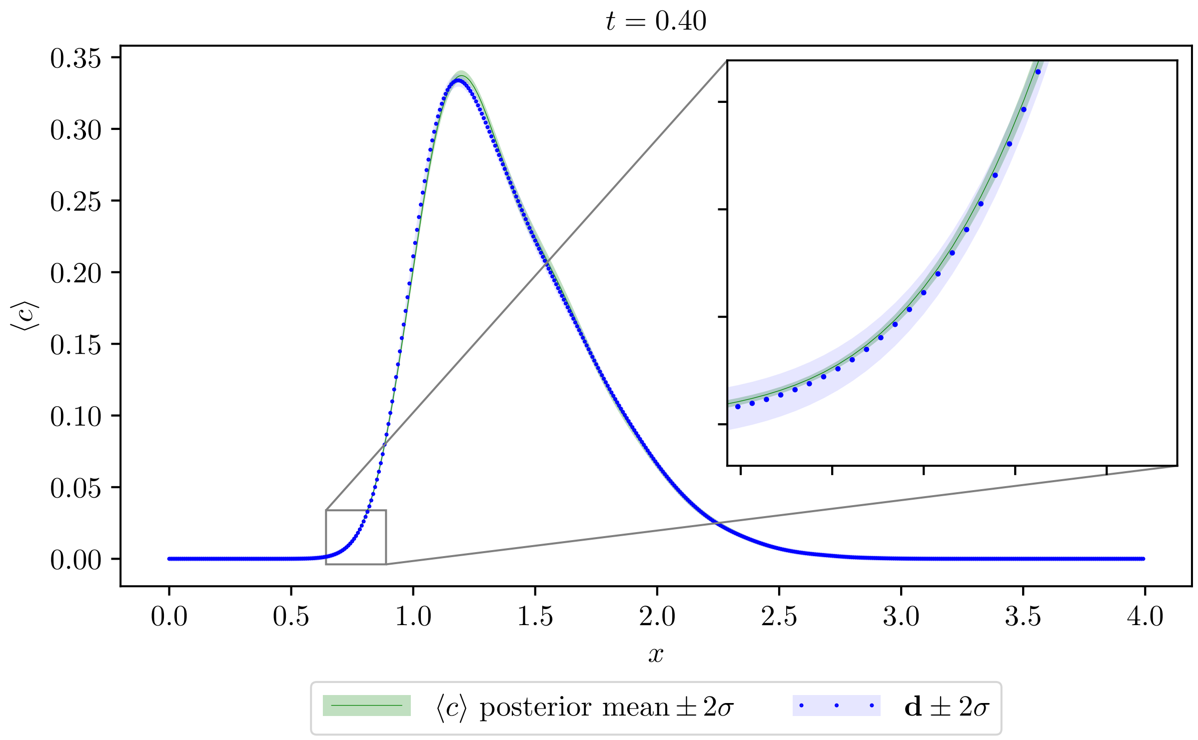
Data: ensemble average from 2D ADE, randomly sampling permeability fields.

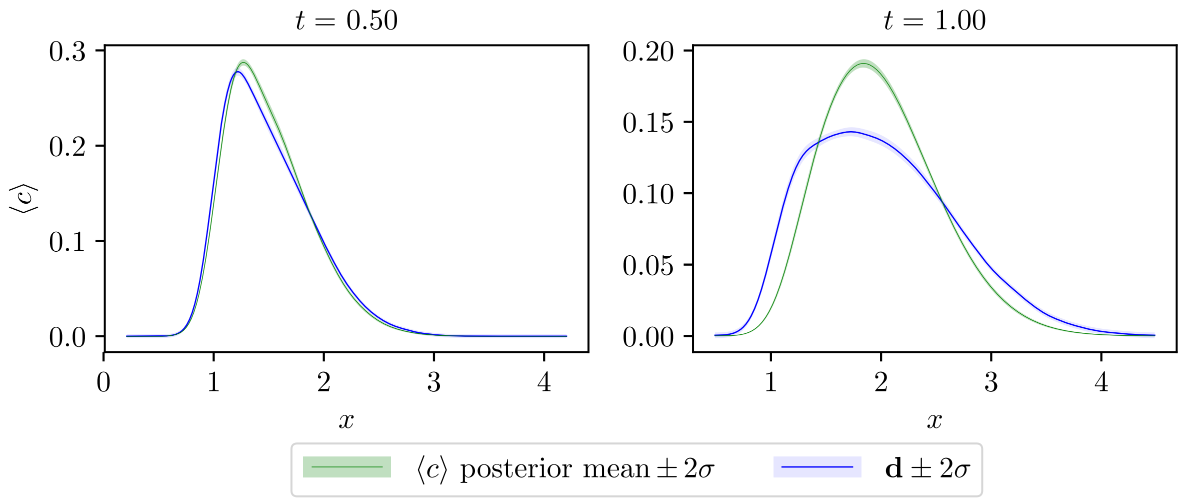
512 spatial observations at early time ($t = 0.4$, 1/10$^{\text{th}}$ of a flowthrough time) to maximize information gain.

$$\boldsymbol{d} = \langle \boldsymbol{c} \rangle_N + \epsilon, \qquad \epsilon \sim \mathcal{N}\left(0, \frac{S_N}{N}\right)$$

Inference can match observations but can't extrapolate in time

$\mathcal{L}$ was assumed to be time independent, but it is known that anomalous diffusion exhibits history dependence.





Conclusions

- Explored the feasibility of inferring uncertain infinite-dimensional operator using limited data.
- Augmented data with physical information through deterministic constraints and prior distribution.
- Eigendecomposition formulation exposed inherent dimensionality of the problem.
- Nonphysical oscillations in solution for few time observations highlights importance of imposing physical constraints in sparse data regimes.
- Success with high-frequency data shows promise of the approach.
- Compared inferred eigenvalues from DNS data to spectrum of common closure models in paper:

Portone, Teresa, and Robert D. Moser. "Bayesian Inference of an Uncertain Generalized Diffusion Operator." *SIAM/ASA Journal on Uncertainty Quantification*, February 7, 2022, 151–78. https://doi.org/10.1137/21M141659X.

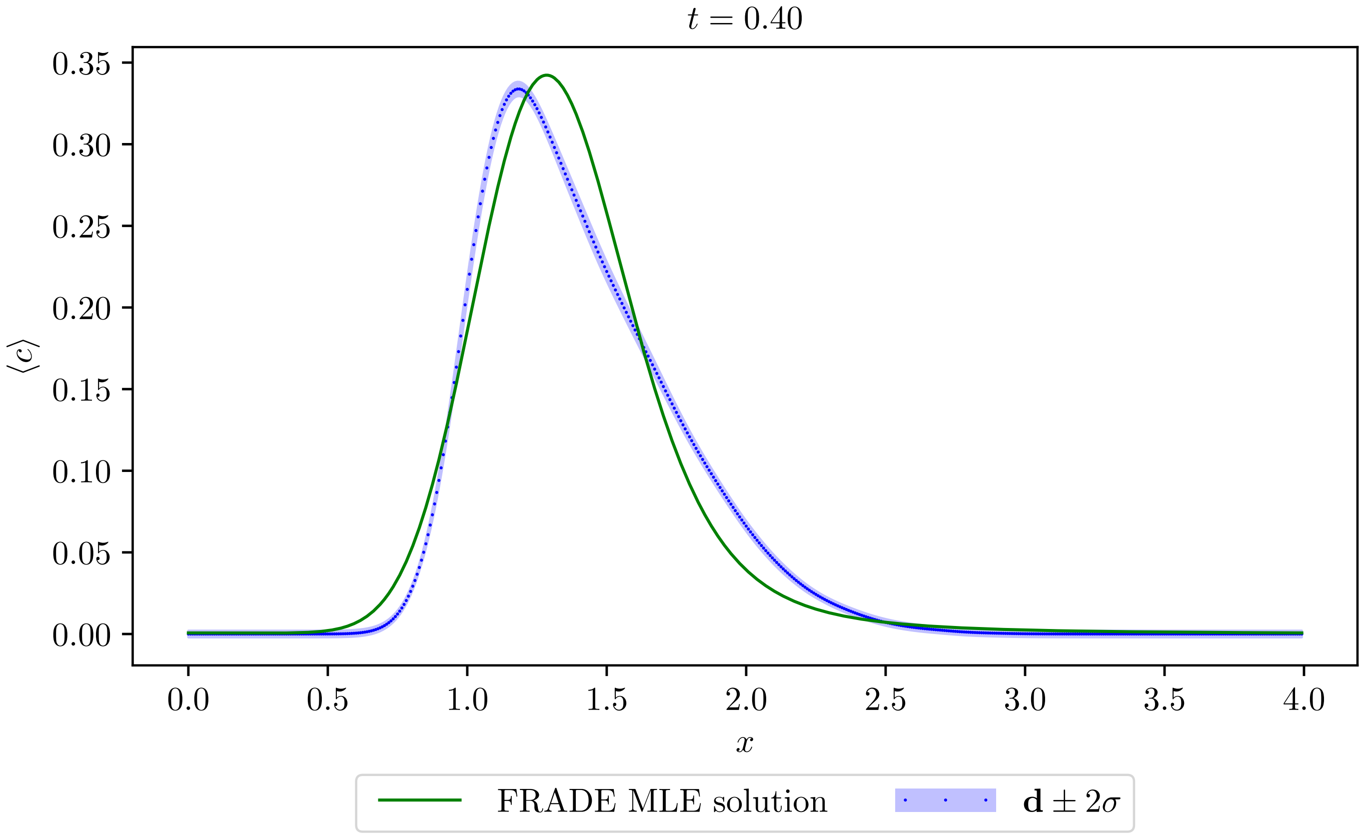
teresaportone.com

Backup slides

How do the inferred eigenvalues of $\mathcal{L}$ compare to the eigenvalues of common closure models for anomalous diffusion?

Spatial fractional advection-diffusion equation (FRADE) can't capture data used in inference

$$\frac{\partial \langle c \rangle}{\partial t} + \langle u \rangle \frac{\partial \langle c \rangle}{\partial x} = \nu \frac{\partial^{\alpha} \langle c \rangle}{\partial x^{\alpha}}$$



Inferred eigenvalues have more complex wavenumber dependence than is assumed for FRADE ($\lambda_k = \nu(ia_k)^\alpha$) or gradient diffusion $\lambda_k = \nu(ia_k)^2$.