High-dimensional Stochastic Interpolation, Optimization and Inversion via Manifold Learning

CASIS 2018- Annual Signal & Image Sciences Workshop

Charanraj A. Thimmisetty[a], Xiao Chen[a], Roger Ghanem[b], Joshua A. White[a], Charles H. Tong[a] and Panagiotis Tsilifis[b]

[a] Lawrence Livermore National Laboratory, USA

[b] University of Southern California, USA

May 23, 2018

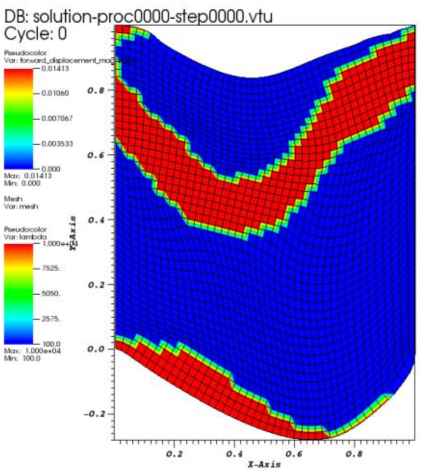
LLNL-PRES-751618-DRAFT

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

High-dimensionality poses challenges in several scientific and engineering applications such as design, model parameter inversion and interpolation.

Elasticity inversion



Target: Invert elasticity field
Data: Displacement measurements

Oil well placement

Goal: Place injection and production wells that maximizes production random permeability field

Subsurface Characterization

Goal: Given proxy measurements, obtain wave velocity characterization of the subsurface

Common challenges:

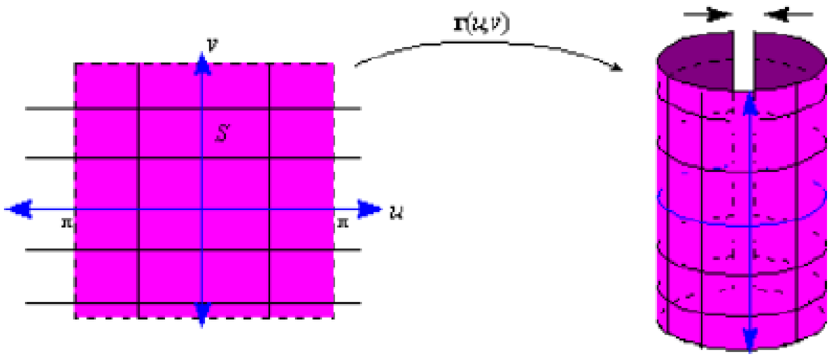
- **Data** (measurements)
  - —High dimensional
  - —Nonlinearly correlated
  - —Non-Gaussian
  - —Noisy and sparse
  - —Expensive to obtain
- **Simulations**
  - —Computationally intensive
  - —Models have uncertainties

Need:

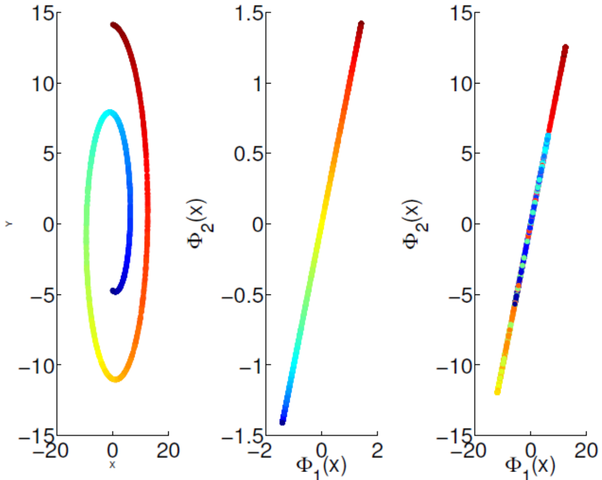
— Understand patterns in the data and build data driven models

— Produce reliable and faster solutions for timely analysis

Manifold learning can provide us easy-to-search space and true metric



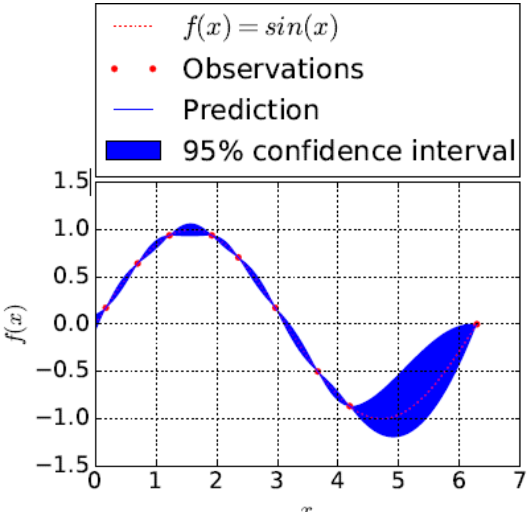
Euclidian distance have different meaning on both geometries even though inhabitants have similar experience

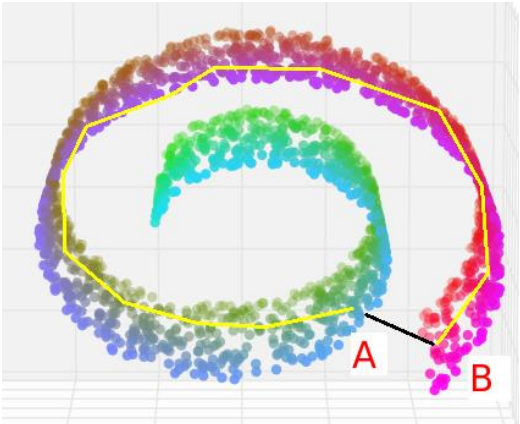


Example: Left original manifold (left), 1D projection using diffusion maps (middle), 1D projection using PCA (right)

- Manifold: A topological space that resembles Euclidean space near each point
- Intrinsic geometry: Geometry experienced by the inhabitants
- Few intrinsic parameters: Length, Area, Gaussian curvatures
- Most of the dataset require non-linear manifold learning techniques to identify underlying manifold
- Non-linear manifold techniques explored in this work include diffusion maps and kernel principle component analysis (KPCA)

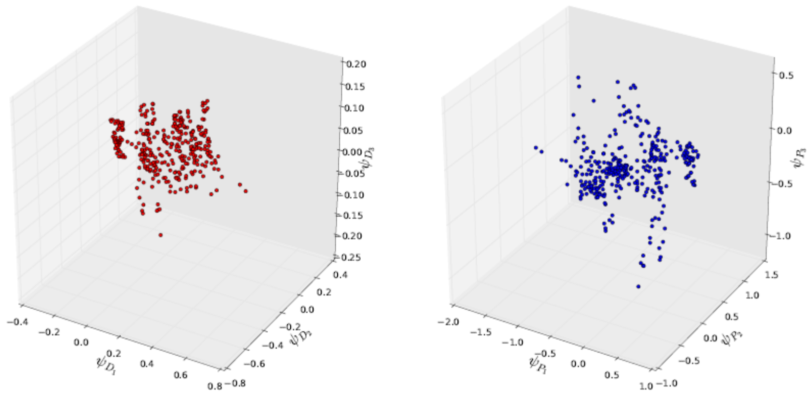
Stochastic intrinsic interpolation: Gaussian process built on the manifold uses true metric and provides probabilistic estimates



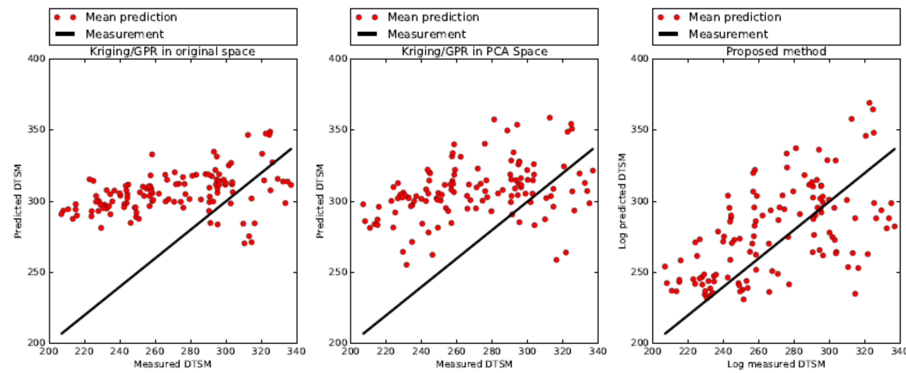


- ❑ Gaussian Process Regression (GPR)/ Kriging acts as interpolation tool for standard kernels
- ❑ Advancements in data acquisition techniques provide high-dimensional proxy datasets
- ❑ GPR produces uninformative predictions as the dimension of predictors increases
- ❑ Soft computing methods such as neural network and support vector machines by default will not provide a full probabilistic predictions
- ❑ We propose a novel intrinsic interpolation method
- ❑ We obtain a low dimensional embedding of the data using diffusion maps
- ❑ GPR is built on the manifold that takes into account of the distance on manifold (diffusion distance) instead of the Euclidean distance

Example: Characterization of geo-physical parameters using a suite of well-log measurements



Scatter plot of first three diffusion & principle components

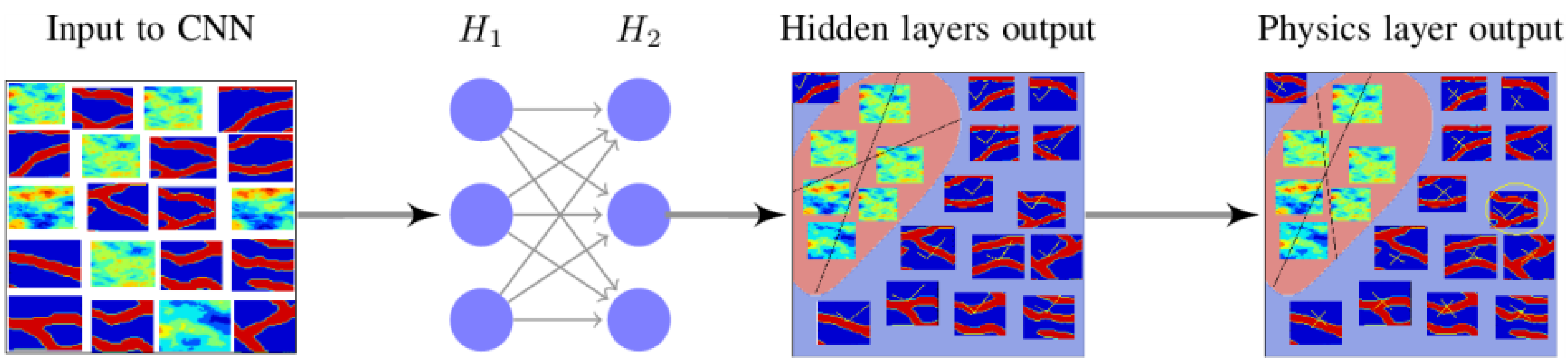


DTSM predictions in Original space (left), Principle component space (middle) and diffusion space (right)

- Goal: To build a regression model for the wave velocities
- Direct measurement of wave velocities can be difficult and/or expensive.
- Proxy quantities are available at dense locations while wave velocities are measured at sparse locations
- Out method gives better metric and more data points per dimension while training thus improved interpolation accuracy

Thimmisetty, C. A., Ghanem, R. G., White, J. A., & Chen, X. (2017). High-Dimensional Intrinsic Interpolation Using Gaussian Process Regression and Diffusion Maps. *Mathematical Geosciences*, 1-20.

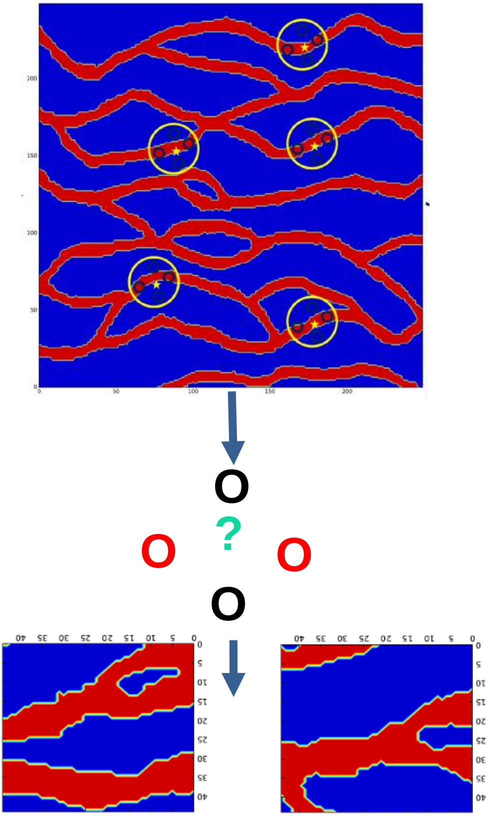
Inversion on the manifolds: Manifolds can provide us easy-to-search space where the inference is computationally cheaper



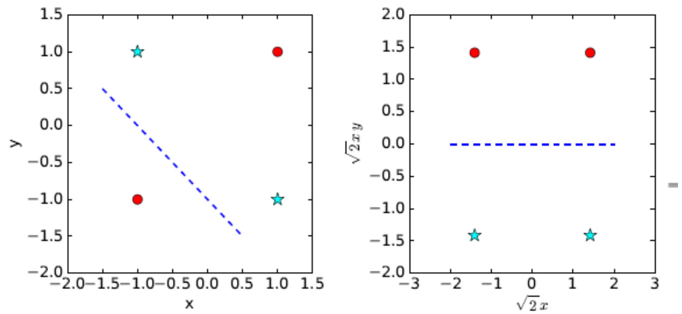
ML algorithms can distinguish channelized and unchannelized sub-surfaces, but they do not account for physical law as a constraint for credible data analysis. Our work will couple the nonlinear manifold deep learning with physics models.

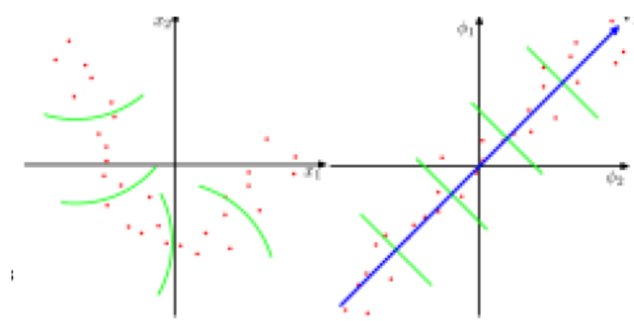
- ❑ Physics constrained data analysis enforces the physics law on the ML algorithm to improve the extracted information and provides valuable insight into distinct datasets
- ❑ Data integration on an easy-to-search feature space extracted by deep learning algorithms reduces the computational complexity of the simulation-driven data integration
- ❑ Seamless data analysis and data integration allow us to identify the relevant features for the quantities of interest in distinct datasets

Machine learning can help us to transform the proxy data into prior knowledge and feature space identification. Adjoint models facilitates higher acceptance rate of MCMC chains.



Snapshots generated from proxy data using multipoint statics





KPCA motivating examples; XOR data (top) and arbitrary non-linear data (bottom)

Ku=F

↓ Adjoint model

$Ku^a=(u-u_m)$

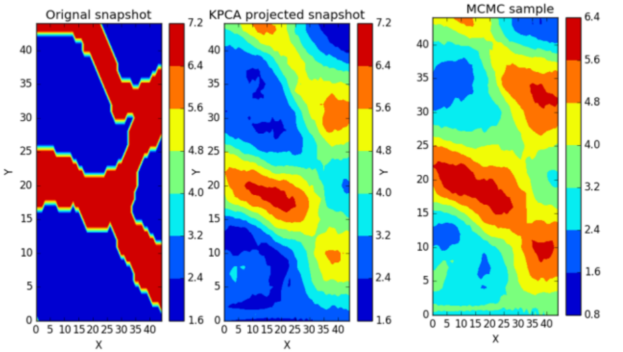
↓ Automatic differentiation

Gradients in feature space

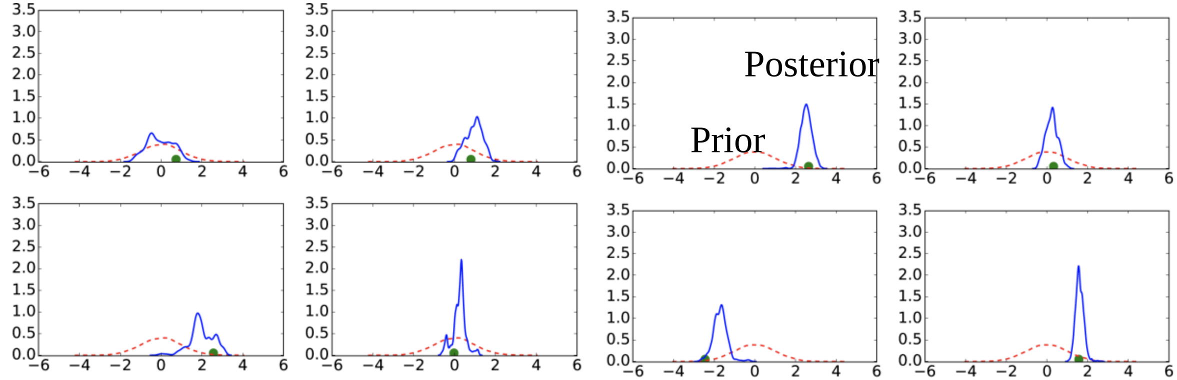
We solve distinct challenges in inverse problems via manifold learning and constructing adjoint PDEs

Challenge	Approach	Explanation
High-fidelity gradient computation	Adjoint gradient	Adjoint PDE allows us to compute gradients in the parameter space with two model runs (a forward and adjoint simulation)
High dimensionality of the parameters	Manifold learning via KPCA/Machine learning	KPCA is used to find a low-dimensional feature space where the solution is not an outlier in the prior probability space
Sampling non-Gaussian feature random variables	PCE	PCE is used to sample KPCA feature random variables that are uncorrelated but dependent non-Gaussians
Ill-posedness of the inverse problem	Bayesian inference	Provides a systematic way to address noisy and measurements and provides a probabilistic inverse solution
Computational intractability of the MCMC	Langevin MCMC	MCMC require O(n) computational capacity while LMCMC need O($n^{1/3}$) and inference is done in feature space

Example: Elasticity inversion with KPCA and adjoint models



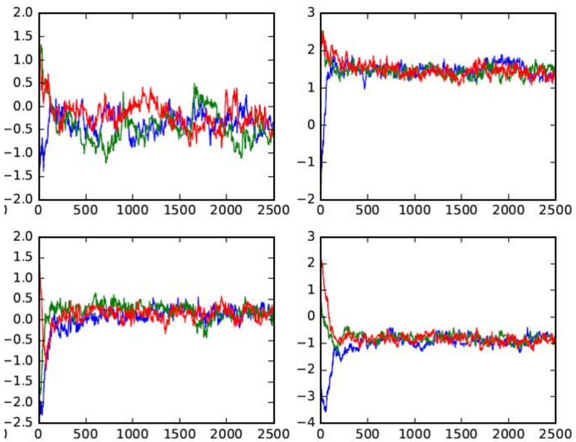
True (left), KPCA projected (middle) and Posterior sample of parameter space.



PCA

KPCA

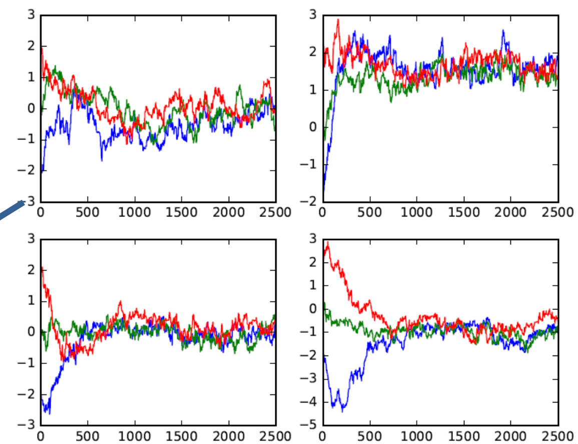
Prior posterior and true densities of feature random variables in PCA space and KPCA space



MCMC Chains

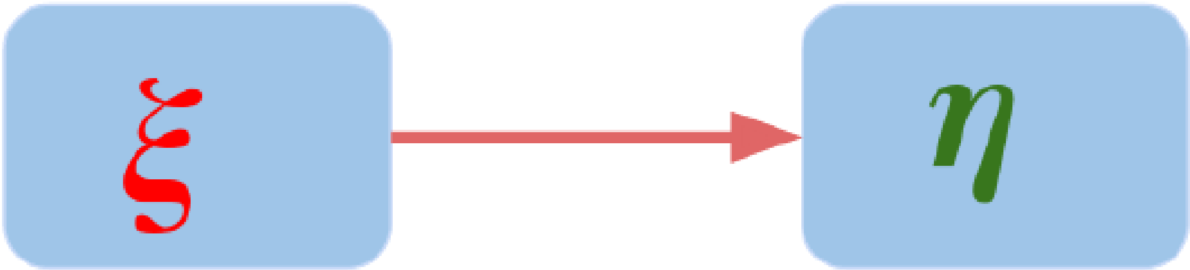
Langevin MCMC

MCMC



Thimmisetty, Charanraj A., et al. "High-dimensional Stochastic Inversion via Adjoint Models and Machine Learning." *arXiv preprint arXiv:1803.06295* (2018).

Optimization on the manifolds: Manifolds are adopted to the quantity of interest

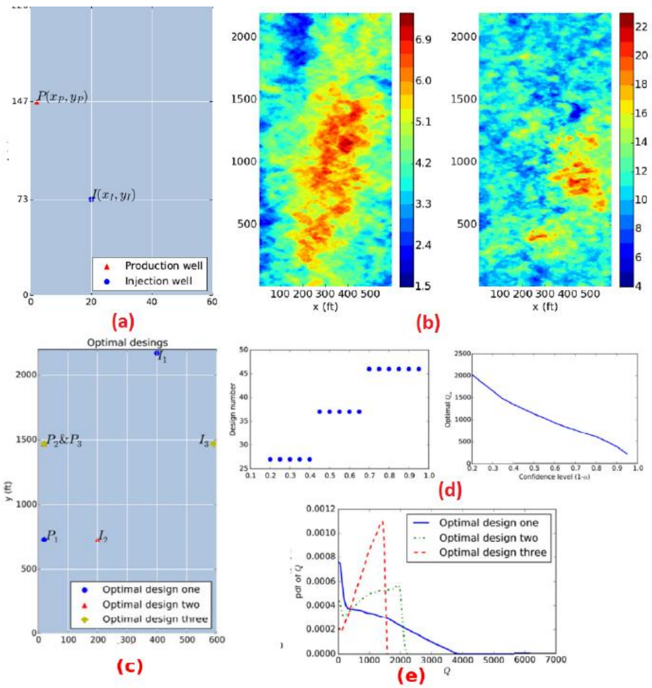


Original random variables | QoI adapted random variables

- In general, the physical system acts a filter (removes the noise in the input) and Quantity of interest (QoI) often lies in a subspace of the input parameter space
- We obtain QoI adapted random variables (η) by rotation of the original random variables (ξ)
- Rotation does not affect accuracy of the solution if the random variables are Gaussian
- Rotation can be found via linear or quadratic adaptation
- We obtain a manifold (subspace) in the rotated space (η) and build surrogate in the subspace

Tipireddy et al. "Basis adaptation in homogeneous chaos spaces." Journal of Computational Physics 259 (2014): 304-317.

Example: Oil well placement problem



(a) Well placement problem (b) mean and variance of permeability (c) three optimal designs (d) (1-α) vs design choice and optimal QoI vs (1 -α) (e) pdf of QoI for the optimal designs

- **Goal**: To maximize the oil production rate (with some confidence, α) by placing injection and production wells at optimal locations
- Uncertainty in permeability **(dim=20)** implies we need to solve optimization problem under uncertainty
- For each design point we have to construct a surrogate model to compute statistics
- Surrogates are built using a manifold adapted to the QOI via basis adaptation **(about 100 model runs for each design point)**
- Computationally efficient compared to MCMC, variants of MCMC and traditional Polynomial Chaos Expansion (PCE)

Thimmisetty et al. "Homogeneous chaos basis adaptation for design optimization under uncertainty: Application to the oil well placement problem." *AI EDAM* 31.3 (2017): 265-276.

Conclusions and future research

- Manifold learning is used improve high-dimensional
  - Stochastic interpolation : Via providing a better metric and more data points per dimension while training thus improved the interpolation accuracy
  - Stochastic inversion : Via doing inversion on easy-to-search feature space extracted by manifold learning reduces the computational complexity of the simulation-driven data integration
  - Optimization under uncertainty : By building quantity of interest aware manifold thus constructing a surrogate with a few model evolutions
- Open Source software (Oct 2018): Data Assimilation for Stochastic Source Inversion (**DASSI**)
- Developing a novel machine learning (ML) framework for seamless data analysis and data integration to identify the relevant features for the quantities of interest in distinct datasets and apply physics constraints on ML models

Thank You

Lawrence Livermore National Laboratory