Physics-informed Machine Learning: A very gentle introduction

Ilias Bilionis
Predictive Science Laboratory
School of Mechanical Engineering
Purdue University
www.predictivescience.org

PREDICTIVE
SCIENCE LABORATORY

Predictive Science Laboratory


Murali Rajasekharan Pillai


Vanessa Kwarteng


Salar Safarkhani


Nimish Awalgaonkar

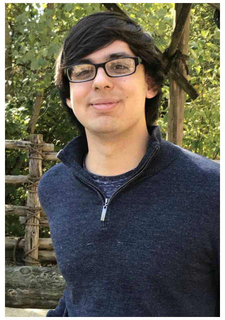

Atharva Hans


Ali Lenjani


Rohit Tripathy


Alana Lund


Sharmila Karamuri


Alex Alberts


Andres Beltran

Physics-informed Machine Learning Subgroup

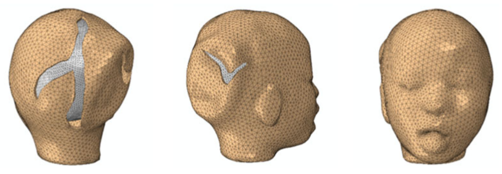
PREDICTIVE
SCIENCE LABORATORY

Three basic problems that we would like to be able to solve.

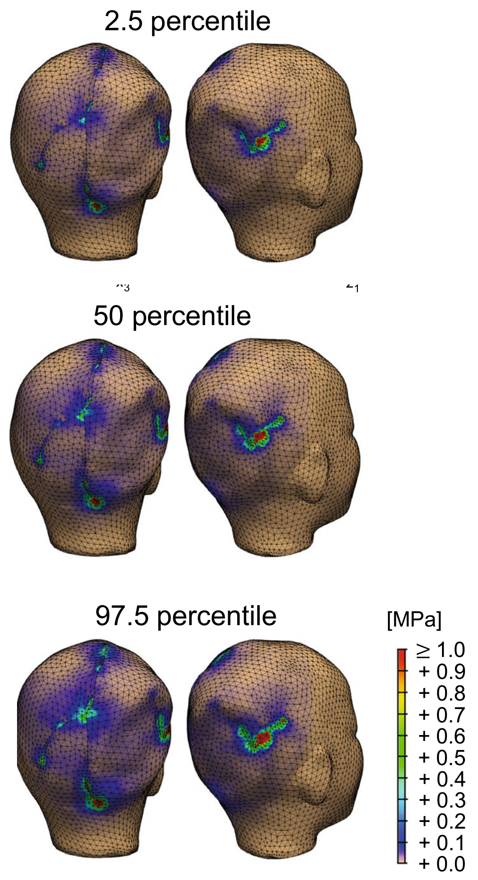
The Uncertainty Propagation Problem (reconstructive surgery)

Table 1 Range of HGO parameters based on Annaidh et al. (2012) and Tonge et al. (2013)

Parameter	Range	Mean
μ (MPa)	[0.004774, 0.2014]	0.04498
k_1 (MPa)	[0.000380, 24.530]	4.9092
k_2 (–)	[0.133, 161.862]	76.64134



Non-linear Elasticity

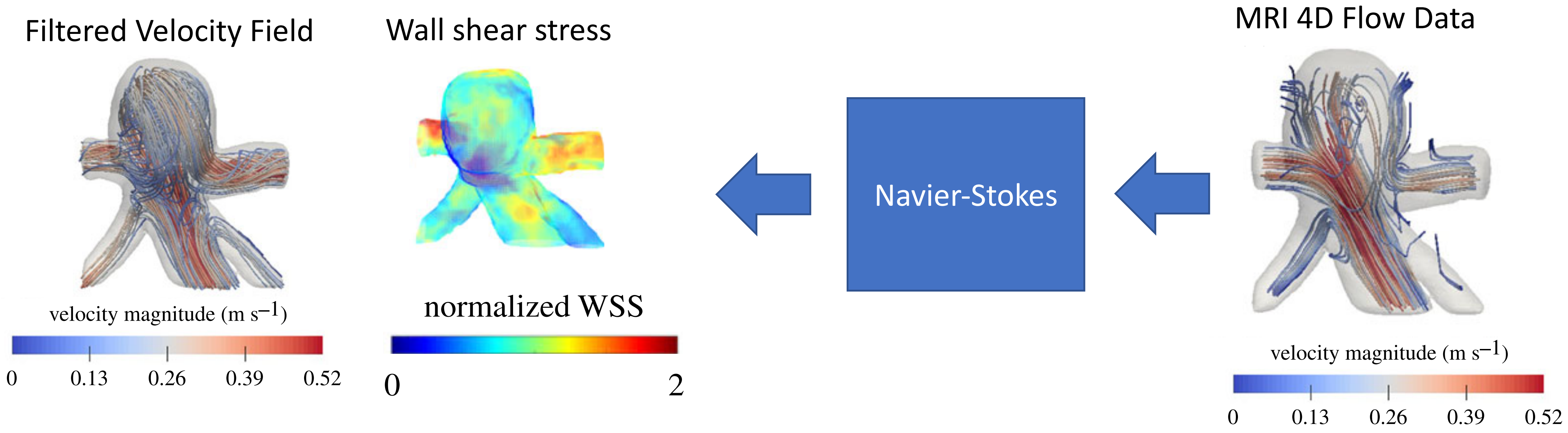


Stress statistics

PREDICTIVE SCIENCE LABORATORY

Lee, T., Turin, S.Y., Gosain, A.K. et al. Biomech Model Mechanobiol (2018) 17: 1857. https://doi-org.ezproxy.lib.purdue.edu/10.1007/s10237-018-1061-

Inverse Problem Example (Cerebral aneurysm)

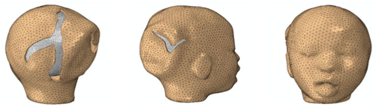


Melissa C. Brindise, Sean Rothenberger, Benjamin Dickerhoff, Susanne Schnell, Michael Markl, David Saloner, Vitaliy L. Rayz, Pavlos P. Vlachos, Multi-modality cerebral aneurysm haemodynamic analysis: *in vivo* 4D flow MRI, *in vitro* volumetric particle velocimetry and *in silico* computational fluid dynamics **16** *J. R. Soc. Interface* http://doi.org/10.1098/rsif.2019.0465

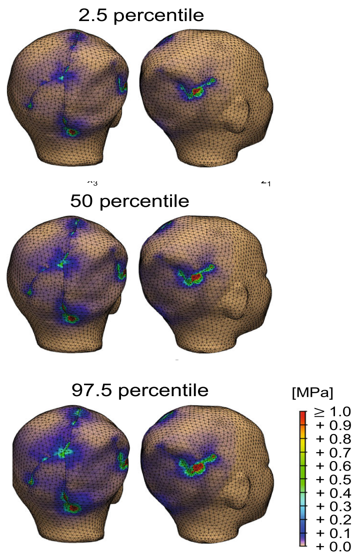
Example of a Design Problem (reconstructive surgery)

Table 1 Range of HGO parameters based on Annaidh et al. (2012) and Tonge et al. (2013)

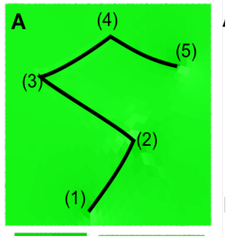
Parameter	Range	Mean
μ (MPa)	[0.004774, 0.2014]	0.04498
k_1 (MPa)	[0.000380, 24.530]	4.9092
k_2 (–)	[0.133, 161.862]	76.64134



Non-linear Elasticity



Stress statistics



Optimal flap design

PREDICTIVE SCIENCE LABORATORY

Figures courtesy of Buganza's group.

We know how to pose these problems mathematically!

We just can’t solve them…

Common Solution Approaches and Their Computational Intractability

- All problems can, in principle, be solved by Monte Carlo sampling.
- Infeasible to do directly with physical simulator.
- Idea -> Replace the simulator with a surrogate model.
- Problem -> ***Curse of dimensionality***.

PREDICTIVE
SCIENCE LABORATORY

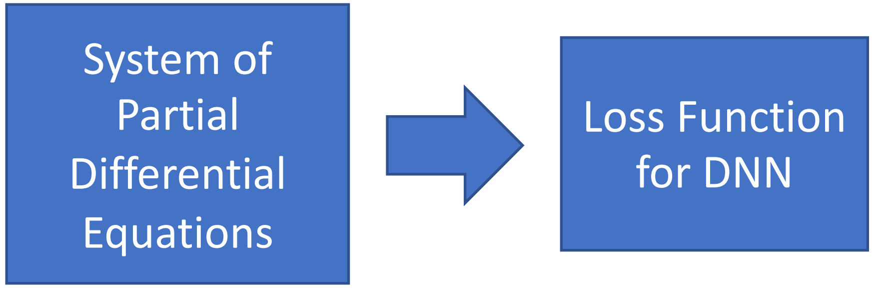
IDEA 1: Use Deep Neural Networks (DNN) to Represent the Response Surface

- Universal function approximators.
- Layered representation of information.
- Tremendous success in high-dimensional applications such as *image classification*, *autonomous driving*.
- Availability of libraries such as *tensorflow*, *keras*, *theano*, *PyTorch*, *caffe* etc.

PREDICTIVE SCIENCE LABORATORY

Tripathy, R. K.; Bilionis, I. Deep UQ: Learning Deep Neural Network Surrogate Models for High Dimensional Uncertainty Quantification. Journal of Computational Physics 2018, 375, 565–588. https://doi.org/10.1016/j.jcp.2018.08.036.

IDEA 2: Get rid of PDE Solver



- Lagaris et al., 1991
- Raisi, Predikaris, Karniadakis, 2019.
- {Raisi, Perdikaris, Karniadakis, Zabaras}* {2018, 2019}.
- Karumuri, Tripathy, Bilionis, Panchal, 2019.
- …

Illustrative Uncertainty Propagation Example With Physics-Informed DNN

Karumuri, S.; Tripathy, R.; Bilionis, I.; Panchal, J. Simulator-free Solution of High-Dimensional Stochastic Elliptic Partial Differential Equations Using Deep Neural Networks. Journal of Computational Physics 2019 (under review). https://arxiv.org/abs/1902.05200.

Stochastic Elliptic Partial Differential Equation

PDE:

$$\nabla(a(\mathbf{x})\nabla u(\mathbf{x})) = 0,$$

$$\mathbf{x} = (x_1, x_2) \in \Omega = [0,1]^2,$$

Boundary conditions:

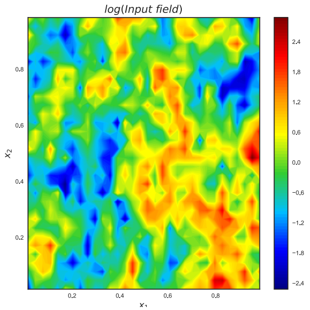
$$u = 0, \forall x_1 = 1,$$

$$u = 1, \forall x_1 = 0,$$

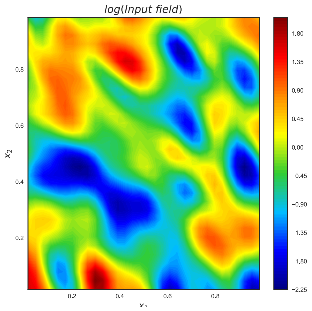
$$\frac{\partial u}{\partial n} = 0, \forall x_2 = 1.$$

Uncertain conductivity:

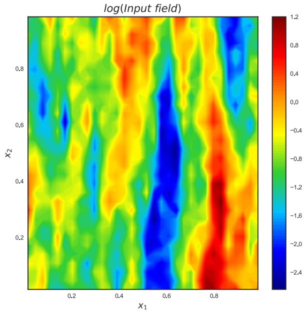
$\log a(x) =$



or



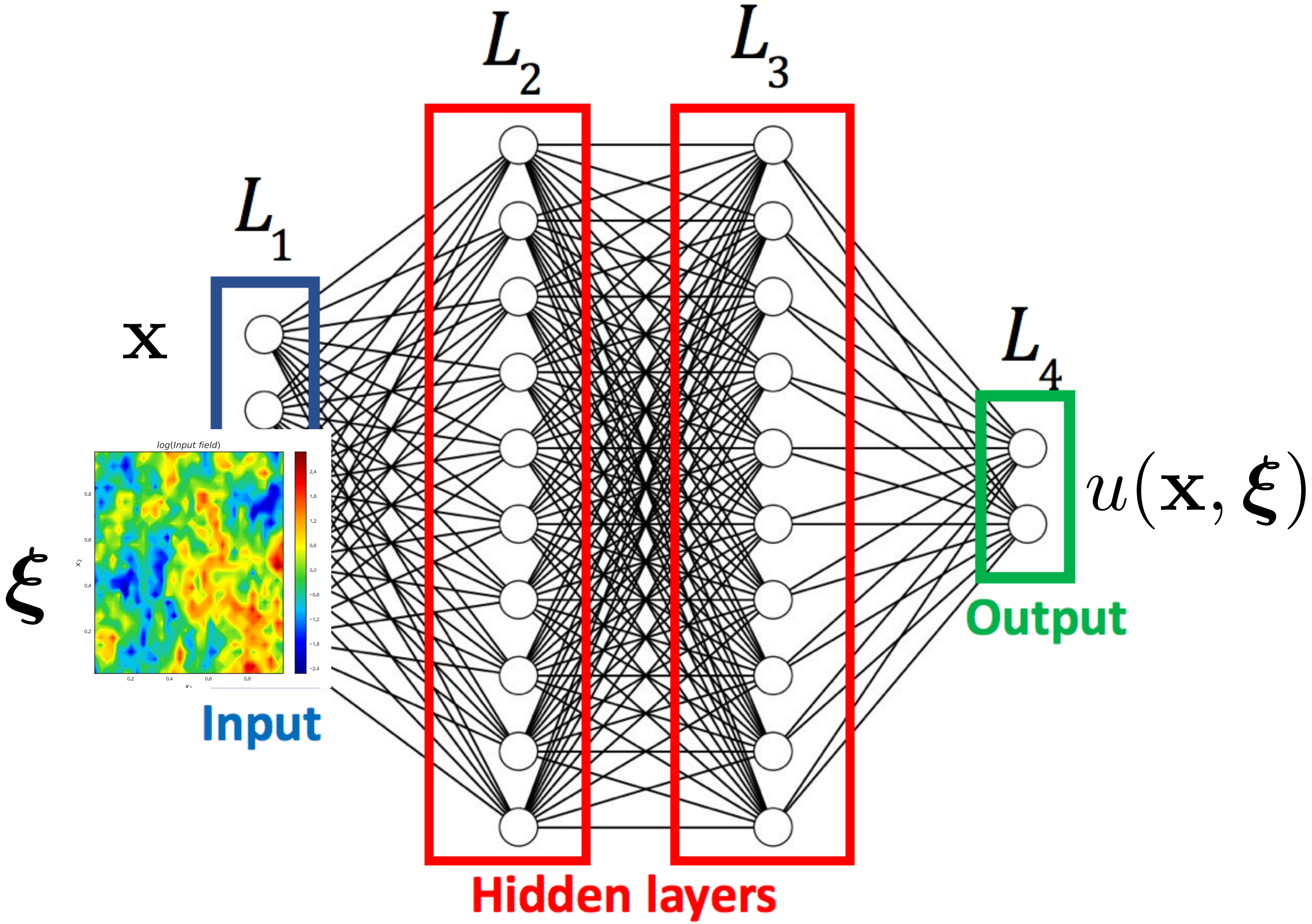
or



...

PREDICTIVE
SCIENCE LABORATORY

Representing the Solution of the Stochastic PDE as a DNN



PREDICTIVE
SCIENCE LABORATORY

How to turn the PDE into a loss function? Integrated Squared Residual

- Move all PDE terms to the left hand side.
- Square and integrate over space/time.
- Take expectation over random parameters.
- Minimize what you get over the space of DNNs subject to any boundary conditions.

$$J[u] = \mathbb{E}_{\xi}\left[\int_{[0,1]^2} \left(\nabla \cdot (a(x,\xi)\nabla \mathrm{u})\right)^2 dx\right].$$

Works, but may have lots of local minima…

Can we do better?

PREDICTIVE
SCIENCE LABORATORY

How to turn the PDE into a loss function? Energy-based Residual

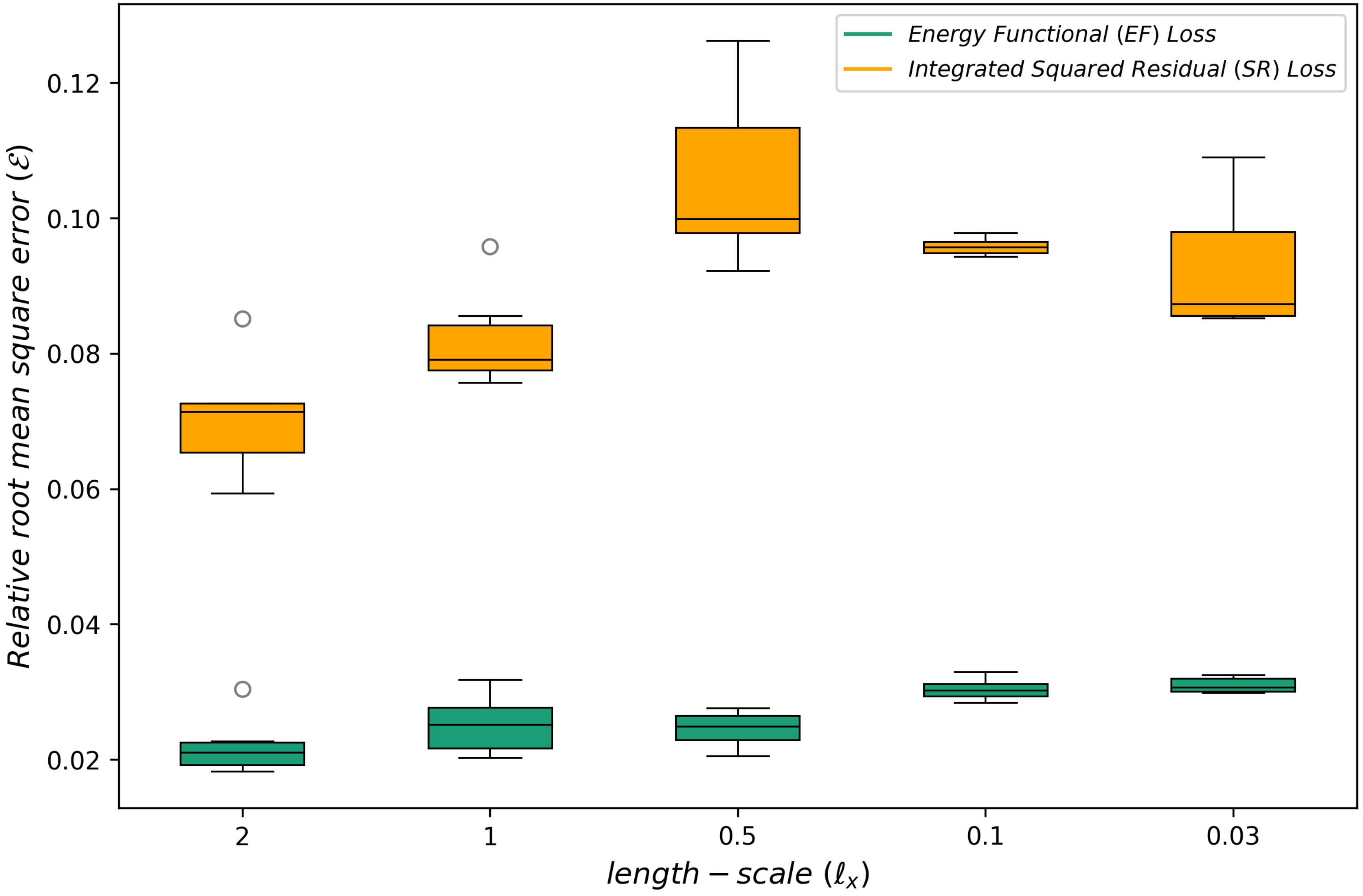
- Write down energy functional for system.
- Take expectation over random parameters.
- Minimize what you get over the space of DNNs subject to any boundary conditions.

$$J[u] = \mathbb{E}_{\xi}\left[\int_{[0,1]}^{2} a(x,\xi) \parallel \nabla u \parallel_2^2 \, dx\right].$$

Energy-based loss is better because you can often prove uniqueness of solution!

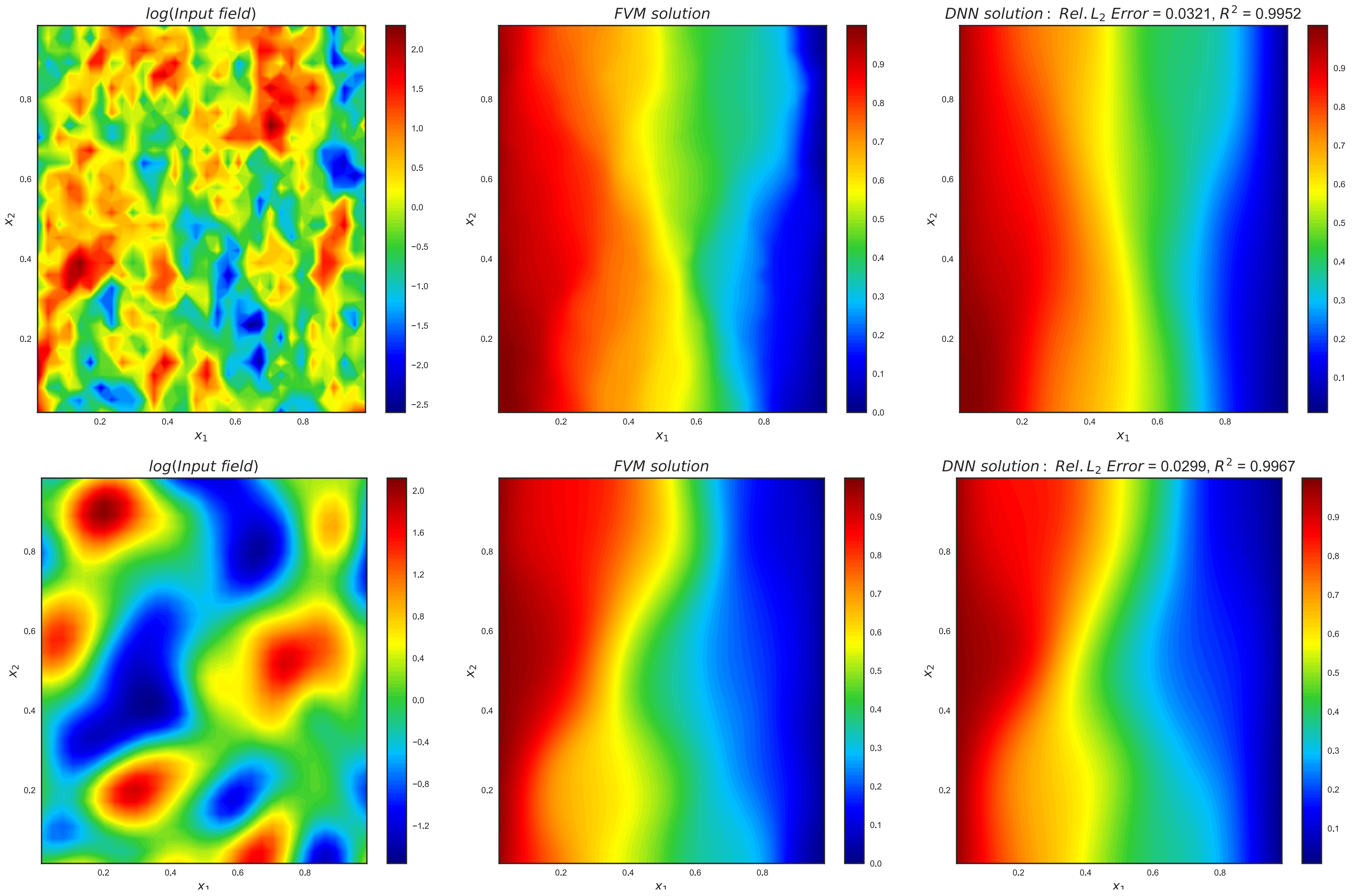
PREDICTIVE
SCIENCE LABORATORY

Integrated Square Residual vs Energy Loss



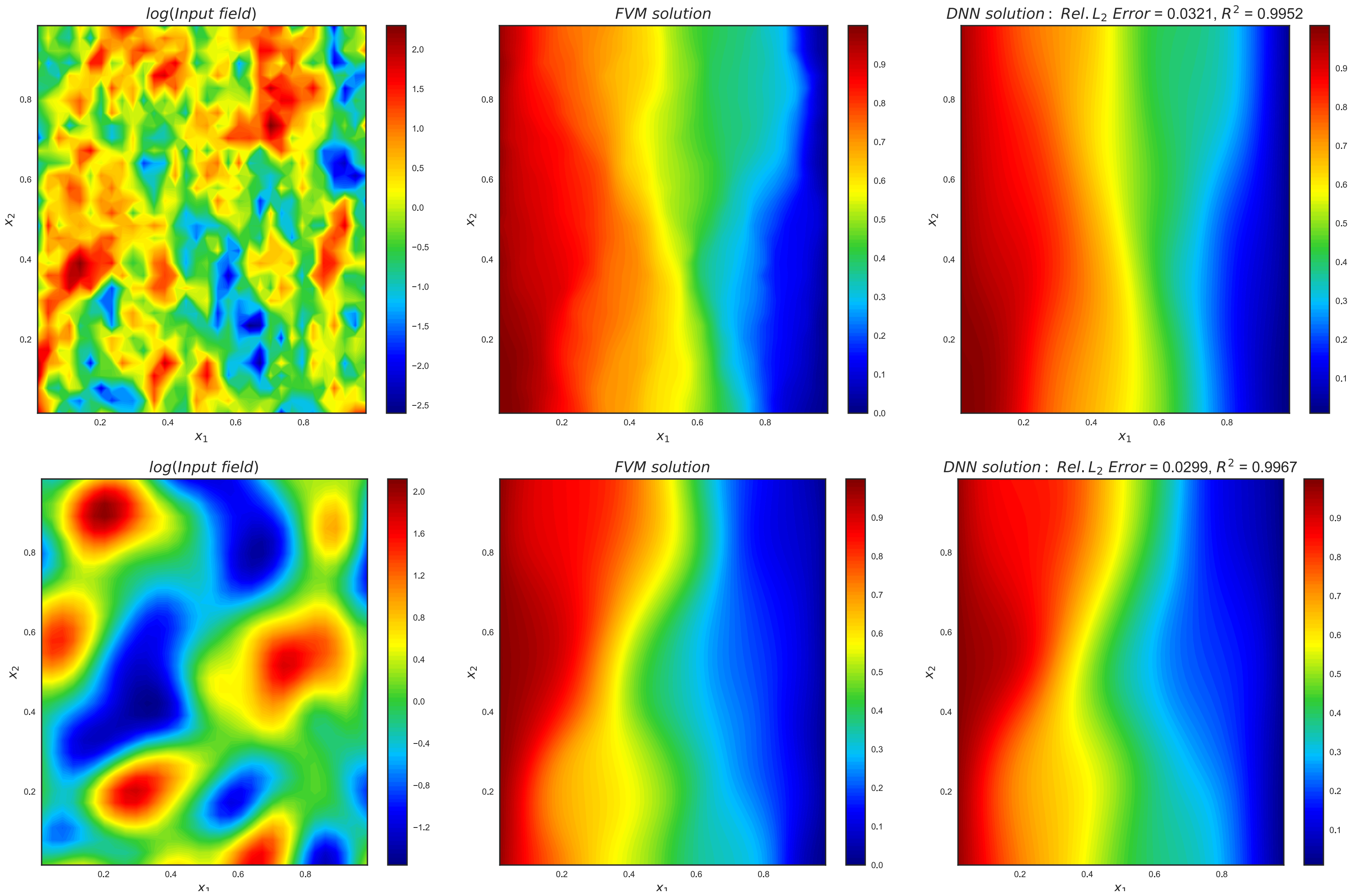
PREDICTIVE
SCIENCE LABORATORY

Numerical Examples: Point-wise Predictions



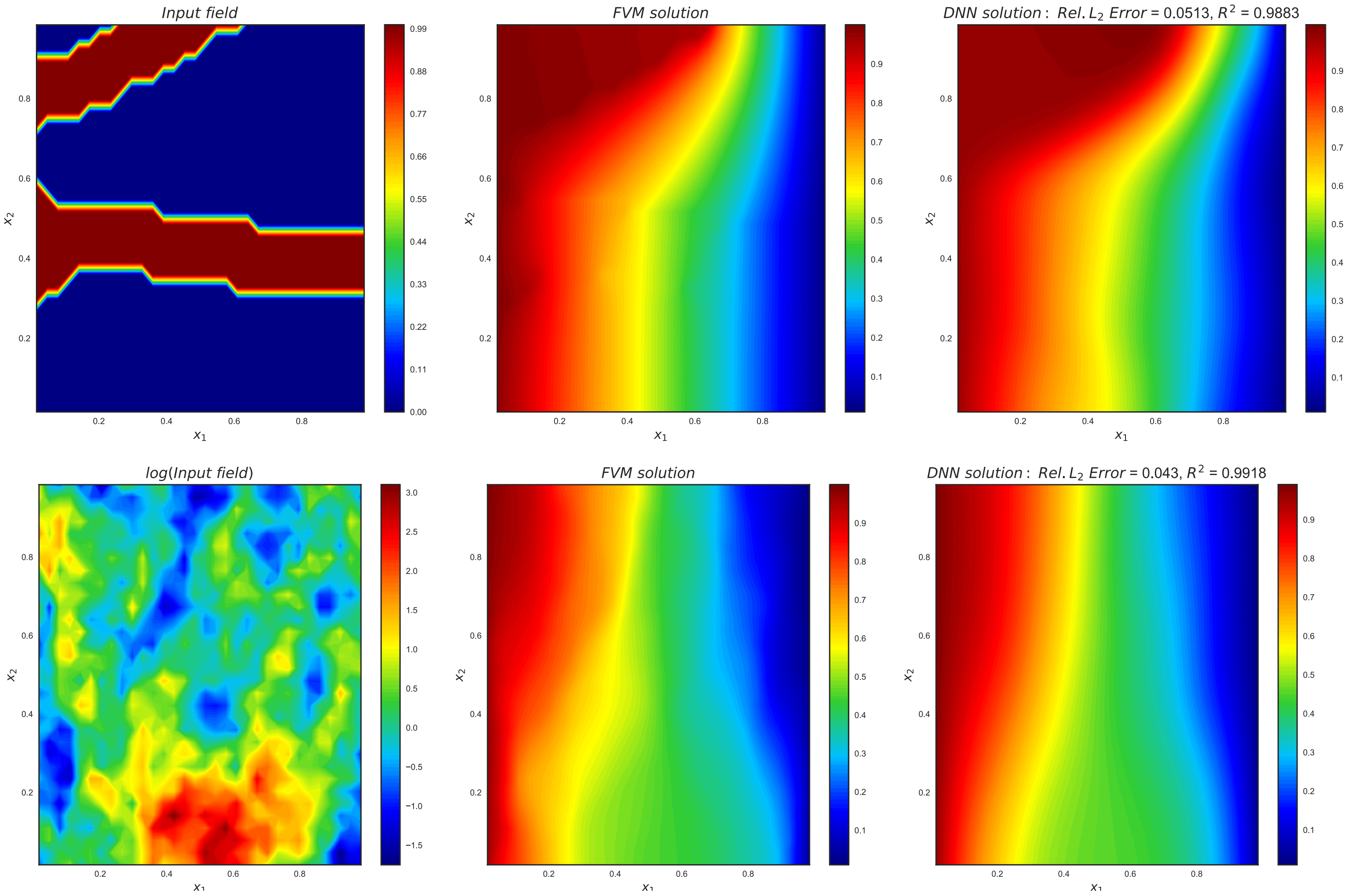
PREDICTIVE
SCIENCE LABORATORY

Numerical Examples: Point-wise Predictions



PREDICTIVE
SCIENCE LABORATORY

Numerical Examples: Point-wise Predictions



PREDICTIVE
SCIENCE LABORATORY

Ending Remarks

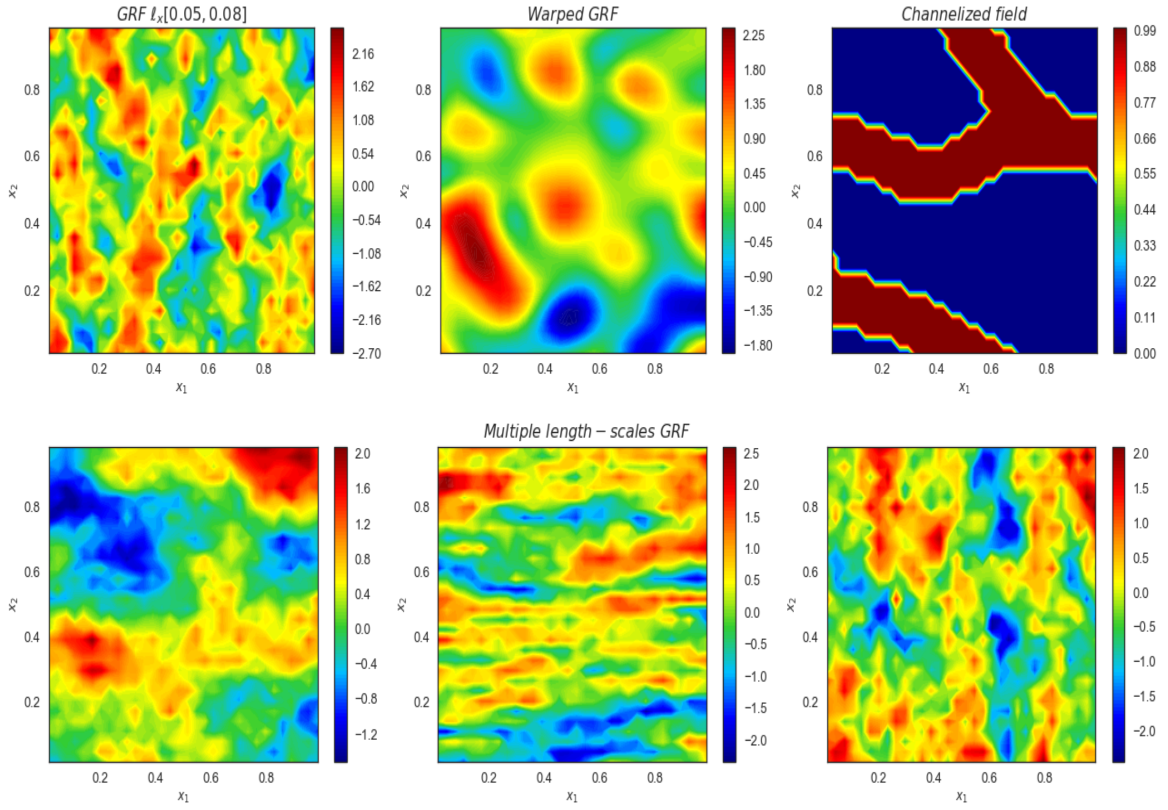
- Lot's of nuances that did not talk about (see paper).
- Can we ditch traditional solvers completely?
- How to pose inverse problems?
- How to pose design problems?
- Best DNN structures?
- Best optimization algorithms?
- Bayesian formulation?

Thank you
ibilion@purdue.edu

But how do I do the integrals?

- You don't have to do the integrals.
- All you need is the ability to sample:
  - uniformly in spatial domain
  - random parameters
- This is sufficient to construct stochastic algorithms that provably converge to a local minimum of the loss (Robbins-Monro, 1956).

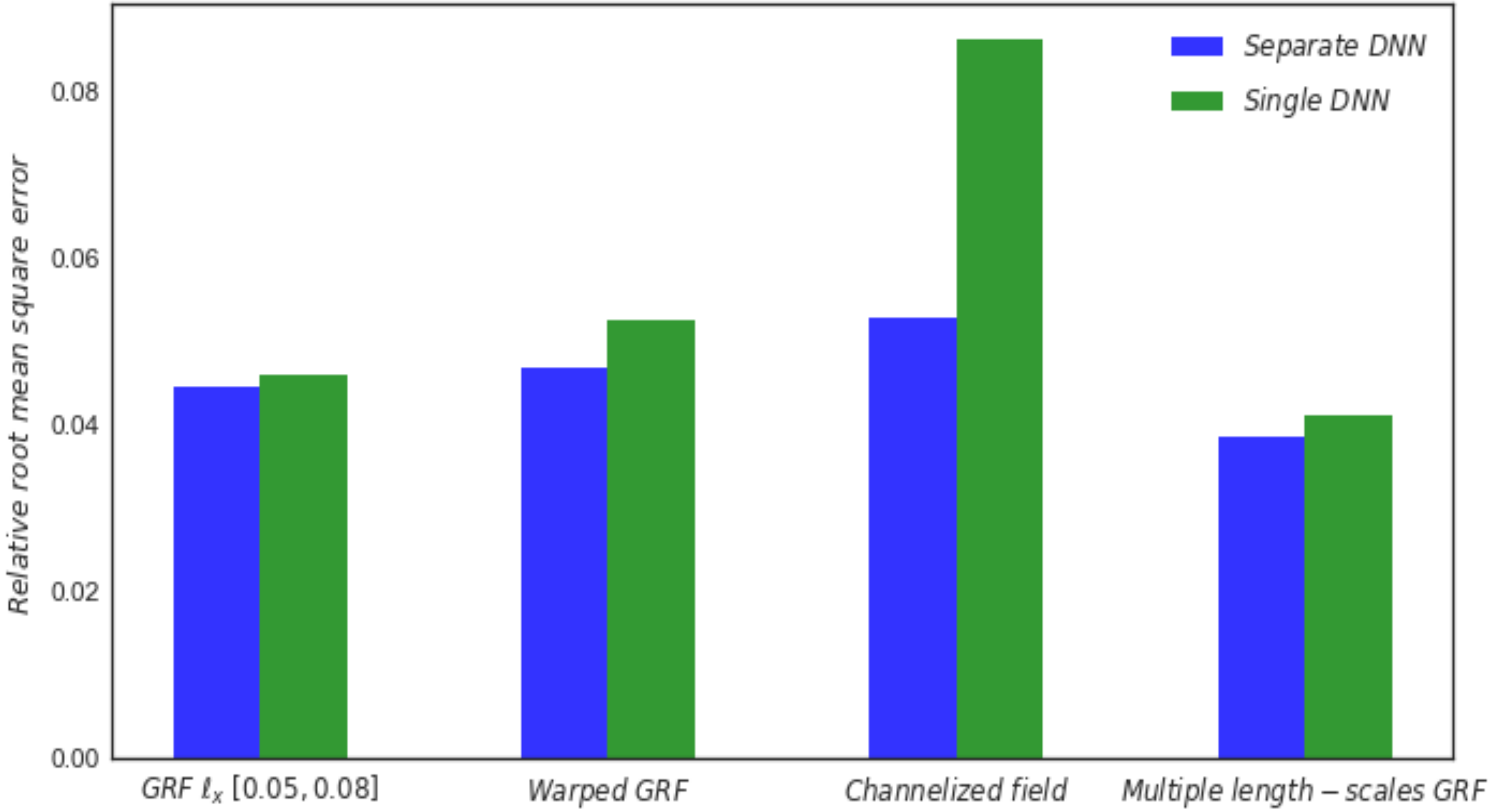
Numerical Examples: Results Summary



Datasets	K	L	n	Number of test samples	$\mathcal{E}$	Number of trainable parameters θ
GRF ℓ_x $[0.05, 0.08]$	3	2	350	2,000	4.45%	1,096,901
Warped GRF	5	2	300	1,000	4.68%	1,211,401
Channelized field	3	2	300	512	5.30%	850,201
Multiple length-scales GRF	3	2	500	9,000	3.86%	2,017,001

PREDICTIVE
SCIENCE LABORATORY

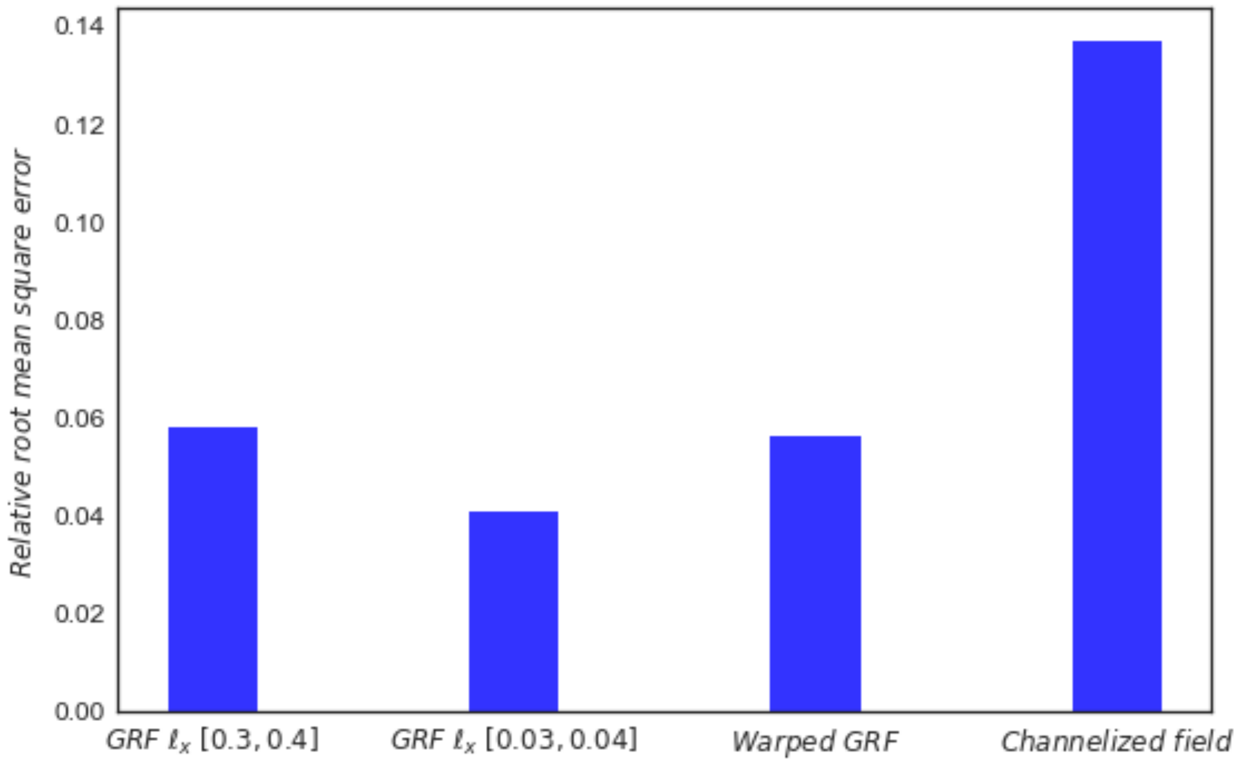
Numerical Examples: One DNN for all fields?



PREDICTIVE
SCIENCE LABORATORY

Numerical Results: Transfer Learning

- Trained on GRF with multiple length scales predicting on other:



PREDICTIVE
SCIENCE LABORATORY