SAND2023-00292C

Sandia National Laboratories

# Improving Bayesian networks multifidelity surrogate construction with basis adaptation

Xiaoshu Zeng[1], Gianluca Geraci[2], Alex Gorodetsky[3], John Jakeman[2], Mike Eldred[2] and Roger Ghanem[1]

1 University of Southern California
2 Sandia National Laboratories
3 University of Michigan

U.S. Department of Energy NNSA

Supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a

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 & owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Secu DE-NA0003525.

January 24, 2022

## Plan of the talk

- Motivation and background
- MFNets
- Basis Adaptation
- Numerical Results
- Closing Remarks

# Motivation and background

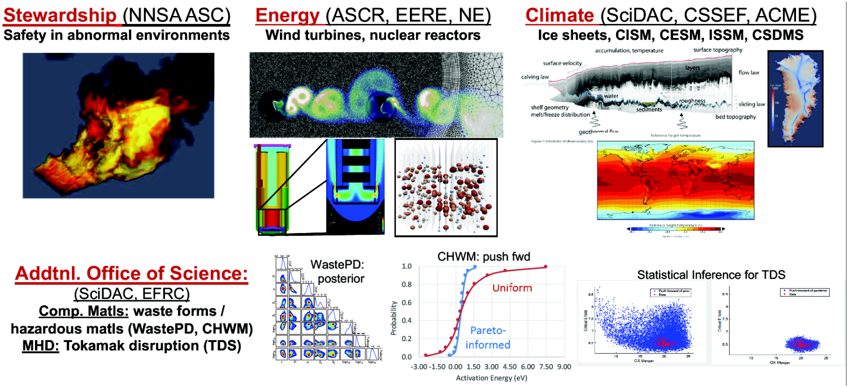


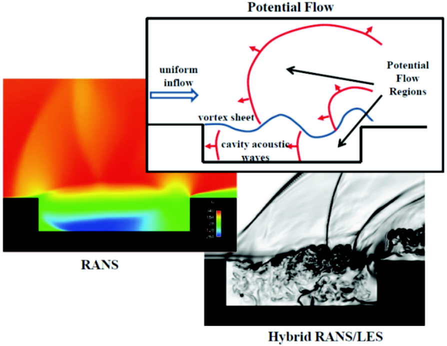

**Figure:** Courtesy of Mike Eldred

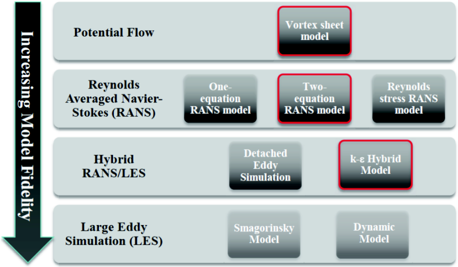
**High-fidelity** state-of-the-art modeling and simulations with HPC

- **Severe** simulations **budget constraints**
- **Significant dimensionality** driven by model complexity

**Multi-fidelity:** several accuracy levels available

- Physical models (Laminar/Turbulent, Reacting/non-reacting, viscous/inviscid...)
- Numerical methods (high/low order, Euler/RANS/LES, *etc.*...)
- Numerical discretization (fine/coarse mesh...)
- Quality of statistics (long/short time history for turbulent flow...)




**Relationships** amongst models can be difficult to anticipate

- **Hierarchical** relationships usually correspond to modeling choices like, *e.g.* discretization
- However, **peer** relationships are often observed in the presence of physical approximations

### MF Sampling methods

- Derive directly from **Monte Carlo**
- Exploit **correlation** among model outputs
- Are build to obtain an **estimator variance reduction**
- Ex: MLMC (Giles,2015), MFMC (Peherstorfer *et al.*, 2016), ACV (Gorodetsky *et al.*, 2020)

### MF Surrogates methods

- Provide an approximation of the *input-output* mapping
- Achieve **rapid error decrease** (as the amount of data increases), provided that the *input-output* mapping is smooth
- Ex: Co-Kriging (Gradiet and Garnier, 2014) and stage-fitting (Liu *et al.* 2018)

### Challenges in MF UQ

1. **Existing strategies assume a prescribed relationship among models**: a general procedure to encode and exploit *a priori* knowledge is not available
2. The presence of **noisy or corrupted data** are not explicitly addressed
3. Heterogeneous sources with **dissimilar uncertainty inputs** are not routinely considered
4. In general, uncertainty estimates are not provided

# MFNets

**Q:** How do we formulate a general approach from which existing strategies (MLMC, MLMF, ACV, co-Kriging *etc.*) can be derived as particular instances?

**MFNets main features:**

- The formulation *unifies* both sampling and surrogate based approaches
- Latent variables (LVs) are used to explain observed relationships among data sources
- LVs allow to leverage **common causes**, not just model outputs (**effects**)

**Conceptual steps:**

- **LVs definition:** LVs can represent both parameters of a simulation or the coefficients of its data-driven representation
- **Dependencies definition:** Bayesian Networks (BNs) provide a mechanism to encode how the data sources are related
- **LVs inference:** conditional independence among LVs is exploited to reduce the computational cost associated to this step
- **UQ analysis:** LVs are used for the propagation of the input variables (and make predictions)

**Linear-Gaussian models**

- **Linear expansion for each model** w.r.t. features (*e.g.* a polynomial chaos basis)

$$Y_i = \sum_{k=0}^{\mathcal{P}_i - 1} \phi_{ik}(X_i)\theta_{ik}$$

- **Gaussian distribution** to encapsulate uncertainty in the LVs

$$\theta = (\theta_1, \theta_2, \ldots, \theta_M) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_M \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{M1} & \cdots & \boldsymbol{\Sigma}_{MM} \end{bmatrix},$$

- **Conditional probability** distributions are restricted to linear-Gaussian:

$$\theta_i | \theta_{\mathrm{pa}(i)} \sim \mathcal{N}\left(\boldsymbol{A}_{i|\mathrm{pa}(i)} \theta_{\mathrm{pa}(i)} + \boldsymbol{b}_{i|\mathrm{pa}(i)}, \boldsymbol{\Gamma}_{i|\mathrm{pa}(i)}\right),$$

- Finally, the **conditional distribution** for $\theta$ given the data $\boldsymbol{y}$ (posterior)

$$\bar{\boldsymbol{\Sigma}} = \left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Phi}^T \Sigma_n^{-1} \boldsymbol{\Phi}\right)^{-1}$$

$$\bar{\boldsymbol{\mu}} = \bar{\boldsymbol{\Sigma}} \boldsymbol{\Phi}^T \Sigma_n^{-1} \boldsymbol{y} + \bar{\boldsymbol{\Sigma}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}.$$

# Basis Adaptation

Let's approximate each model source via a Polynomial Chaos Expansion (PCE)

$$Y_i(X_i) = \sum_{k=0}^{\mathcal{P}_i - 1} \theta_{ik} \psi_{ik}(X_i)$$

Basis Adaptation (BA) is a **dimension reduction** strategy based on the following steps

1 Seek rotated/adapted variables $\boldsymbol{\eta}$

$$\boldsymbol{\eta}_i = \boldsymbol{\eta}_i(X_i) = R_i X_i \, . \tag{1}$$

2 Reduce the dimension, the first $r_i$ important dimensions are adequate to represent $Y_i$

$$\boldsymbol{\eta}_i = \begin{bmatrix} R_{i_r} X_i \\ R_{i_{\neg r}} X_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{\eta}_{i_r} \\ \boldsymbol{\eta}_{i_{\neg r}} \end{bmatrix} \, . \tag{2}$$

3 The original model is approximated with the important dimensions

$$Y_i^{R_{i_r}}(\boldsymbol{\eta}_{i_r}) = \sum_{k=1}^{\mathcal{P}_i^r} \theta_{ik}^{R_{i_r}} \psi_{ik}(\boldsymbol{\eta}_{i_r}) \, , \tag{3}$$

**Why is embedding BA into MFNets a good idea?**

1. It is possible to show that the correlation of two models is higher if *important variables* are used;
2. The number of PCE coefficients, of each model, can be highly reduced

$$\mathcal{P}_i^r = \frac{(r_i + p_i)!}{r_i!\,p_i!} << \frac{(d_i + p_i)!}{d_i!\,p_i!}$$

**How do we embed BA into MFNets?**

1. Generate pilot samples for all models ($m_i$ samples per model)
2. Construct rotation matrix (for each model) and generate $\boldsymbol{\eta}_i = R_{i_r} X_i$
3. Assign priors (mean and covariances)
4. Evaluate posterior mean and variances for the high-fidelity model

$$\mu_{\bar{Y}_M} = \boldsymbol{\Phi}_M(\bar{X}_M)\bar{\boldsymbol{\mu}}_M$$
$$\mathbb{V}\mathrm{ar}(\bar{Y}_M) = \mathrm{diag}\left(\boldsymbol{\Phi}_M(\bar{X}_M)\bar{\boldsymbol{\Sigma}}_{MM}\boldsymbol{\Phi}_M(\bar{X}_M)^T\right)$$

# Numerical Results

**Numerical examples**

1. Analytical test problem (verification)
2. Finite element model of nuclear spent fuel

**Two graphs**

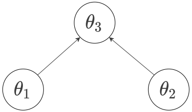

**Figure:** Graph structure with three nodes with (a) a peer and (b) a hierarchical structure. The HF model is $Y_3$ and is described by variable $\theta_3$, while the two low-fidelity models are $Y_2$ and $Y_1$, with variables $\theta_2$ and $\theta_1$, respectively.

**Accuracy/Precision metrics**

- **Mean-Squared Bias** (MSB): average of the squared bias for the test points
- **Mean-VARiance** (MVAR): average variance (of the posterior) over the test points
- **Mean-Mean Squared Error** (MMSE): average MSE over the test points, MMSE = MSB+MVAR

**NOTE:** No hyper-parameter tuning is used for MFNets, few choices for the parameters values are explored

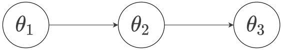
### Peer graph




- $\theta_3 = A_{31}\theta_1 + A_{32}\theta_2 + v_3$
- $\mathrm{Cov}[\theta_\alpha, v_3] = 0\ \forall\alpha$ and $\mathrm{Cov}[\theta_1, \theta_2] = 0$
- $A_{31} = a_{31}I$, $A_{32} = a_{32}I$, $\mathbf{\Sigma}_{11} = s_{11}I$, $\mathbf{\Sigma}_{22} = s_{22}I$, $\mathbf{\Sigma}_{33} = s_{33}I$, and $\Sigma_{v_3} = \mathbf{\Sigma}_{33} - a_{31}^2\mathbf{\Sigma}_{11} - a_{32}^2\mathbf{\Sigma}_{22}$
- Prior covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} s_{11}I & 0 & a_{31}s_{11}I \\ 0 & s_{22}I & a_{32}s_{22}I \\ a_{31}s_{11}I & a_{32}s_{22}I & s_{33}I \end{bmatrix}$$

### Hierarchical graph




- PCE coefficients

$$\theta_2 = A_{21}\theta_1 + v_2$$
$$\theta_3 = A_{32}\theta_2 + v_3\ ,$$

- $\mathrm{Cov}[\theta_\alpha, v_2] = 0$ and $\mathrm{Cov}[\theta_\alpha, v_3] = 0\ \forall\alpha$
- $A_{21} = a_{21}I$, $A_{32} = a_{32}I$, $\Sigma_{v_2} = \mathbf{\Sigma}_{22} - a_{21}^2\mathbf{\Sigma}_{11}$, and $\Sigma_{v_3} = \mathbf{\Sigma}_{33} - a_{32}^2\mathbf{\Sigma}_{22}$
- Prior covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} s_{11}I & a_{21}s_{11}I & a_{21}a_{32}s_{11}I \\ a_{21}s_{11}I & s_{22}I & a_{32}s_{22}I \\ a_{21}a_{32}s_{11}I & a_{32}s_{22}I & s_{33}I \end{bmatrix}$$

### Common assumptions

- Five scalars for parametrizations ($s_{11} = s, s_{22} = s, s_{33} = s, a_{31}, a_{32}$), with $s = 1$
- Two cases
  - Uncorrelated coefficients $a_{31} = a_{32} = 0$
  - Correlated coefficient $a_{31} = a_{32} \approx \rho$

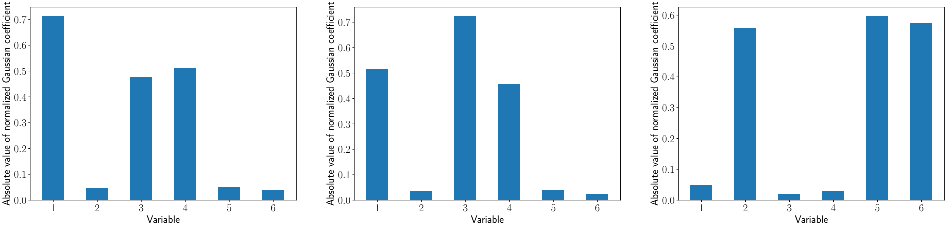
## A three-model analytical example

**Definitions**

$$
\begin{aligned}
f_1(\boldsymbol{x}) &= \exp\left(x_1 + 0.05x_2\right) + \exp 0.8x_3 + \exp\left(0.8x_4 + 0.05x_5 + 0.05x_6\right), \\
f_2(\boldsymbol{x}) &= \log\left(0.75x_1 + 0.05x_2 + 1\right) + \log\left(x_3 + 0.05x_5 + 1\right) + \log\left(0.5x_4 + 0.05x_6 + 1\right), \\
f_3(\boldsymbol{x}) &= \exp\left(0.1x_1 + 1.2x_2\right) + \exp 0.05x_3 + \exp\left(0.05x_4 + x_5 + x_6\right) \\
&\quad + \log(0.05x_1 + 0.8x_2 + 1) + \log(0.75x_5 + x_6 + 1),
\end{aligned}
$$

**Training Dataset:** $m_1 = m_2 = 100$ and $m_3 = 20$
**Testing Dataset:** 200 points

**Basis Adaptation**

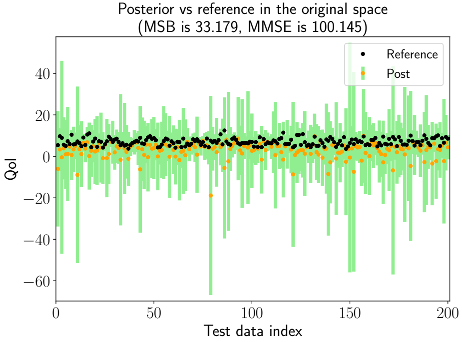



(a) Low-fidelity ($f_1$) (b) Low-fidelity ($f_2$) (c) High-fidelity ($f_3$)

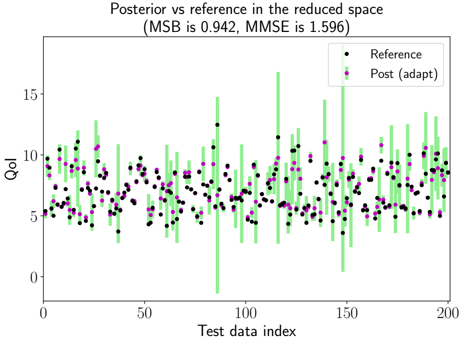
**Figure:** Normalized first-order PCE coefficients in the original space.

# Analytical test problem

## CASE 1: Peer Graph and uncorrelated coefficients ($a_{31} = a_{32} = 0$) – third order polynomial




**(a)** Posterior (Without BA)

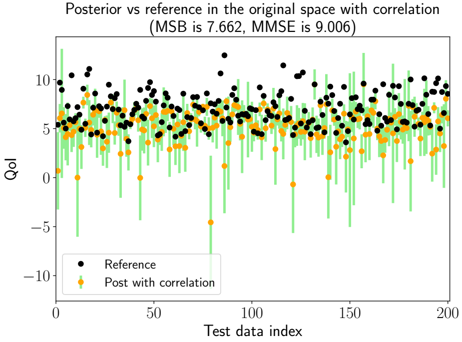



**(b)** Posterior (With BA)

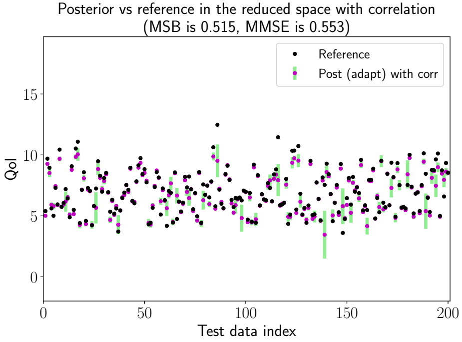
**Figure:** CASE 1 - Analytical example with peer graph and uncorrelated coefficients

# Analytical test problem

## CASE 2: Peer Graph and correlated coefficients ($a_{31} = a_{32} = 0.7$) – third order polynomial




**(a)** Posterior (Without BA)




**(b)** Posterior (With BA)

**Figure:** CASE 2 - Analytical example with peer graph and correlated coefficients

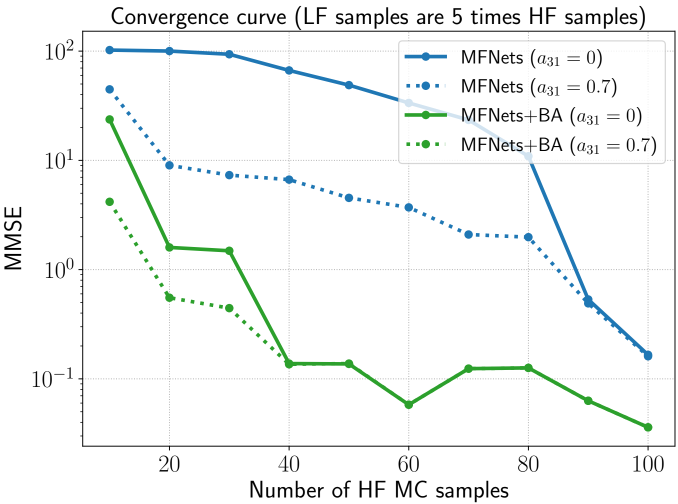
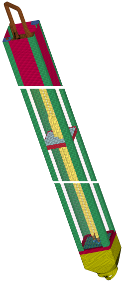
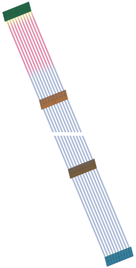



**Figure:** MMSE for different surrogate constructions with MFNets with and without BA – Analytical test problem.

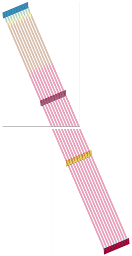
## FEM model for nuclear fuel assembly



**(a)** Detailed fuel assembly (HF)



**(b)** First simplified fuel assembly (LF2)



**(c)** Second simplified fuel assembly (LF1).

**Models' description**

- FEM of a spent nuclear fuel assembly with 92 fuel rods, 2 water rods, and 8 spacers
- Squared channel structure (welded to the upper handle and lower tie plate)
- HF model's DoFs $N = 1,912,506$
- LF2: $10 \times 10$ fuel rods, one upper-tie plate, and one low-tie plate $\rightarrow N = 368,892$ DOFs
- LF1: coarse version of LF2 $\rightarrow N = 64,392$ DOFs

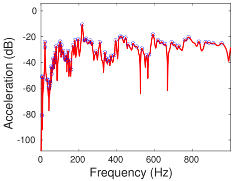
**Training dataset:** $m_1 = m_2 = 1000$, $m_3 = 150$
**Testing dataset:** 100 points

# FEM exemplar
## Problem definition

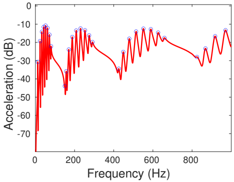
**Uncertainties ($d = 8$)**

- Connections of different structural levels: rod-to-grid and grid-channel $\rightarrow$ beta distributions

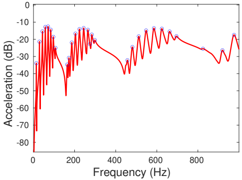
**QoI – Frequency Response Function (Nominal response in the figure)**




**(a)** FRF (detailed model - HF)

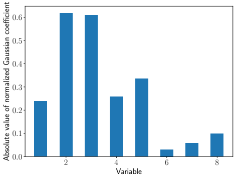



**(b)** FRF (first LF model - LF2)

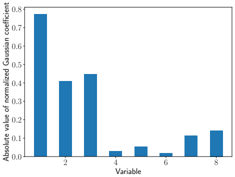



**(c)** FRF (second LF model - LF1)

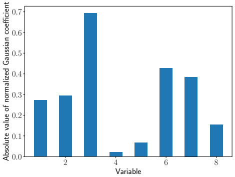
**Basis Adaptation**




**(a)** LF1



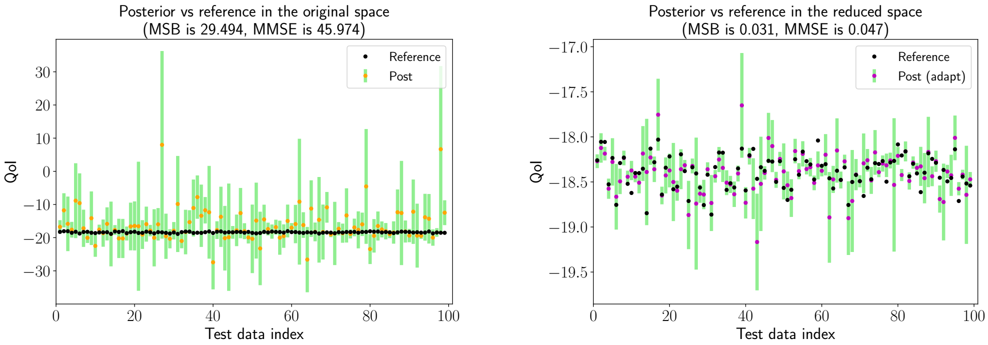

**(b)** LF2




**(c)** HF

# FEM exemplar

## CASE 1: Peer Graph and uncorrelated coefficients ($a_{31} = a_{32} = 0$) – third order polynomial



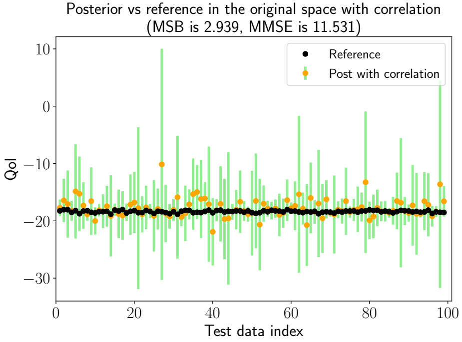

**(a)** Posterior (Without BA)

**(b)** Posterior (With BA)

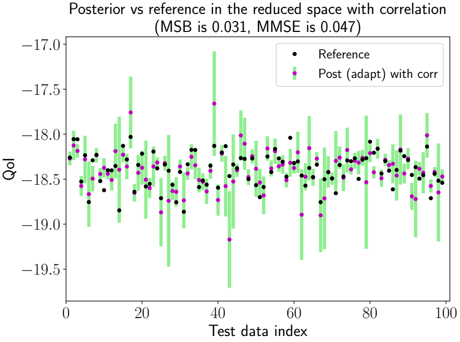
**Figure:** CASE 1 - FEM exemplar with peer graph and uncorrelated coefficients

# FEM exemplar

## CASE 1: Peer Graph and correlated coefficients ($a_{31} = a_{32} = 0.5$) – third order polynomial




**(a)** Posterior (Without BA)




**(b)** Posterior (With BA)

**Figure:** CASE 1 - FEM exemplar with peer graph and correlated coefficients

# Closing remarks

**Summary**

- We explored the embedding of the BA dimension reduction strategy into MFnets
- Several numerical examples were considered
- Results are encouraging and motivate additional tests, e.g., MFNets+BA seems to be insensitive w.r.t. the graph

**Next steps**

- The next step is to include hyper-parameter tuning in both strategies
- Extend the numerical examples to larger ensemble of models

# THANKS!

**Acknowledgements**

- Cosmin Safta (SNL)

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.