Los Alamos NATIONAL LABORATORY EST. 1943

Quantifying Topological Uncertainty in Fractured Systems Using Graph Theory and Machine Learning

Gowri Srinivasan
October 31, 2018

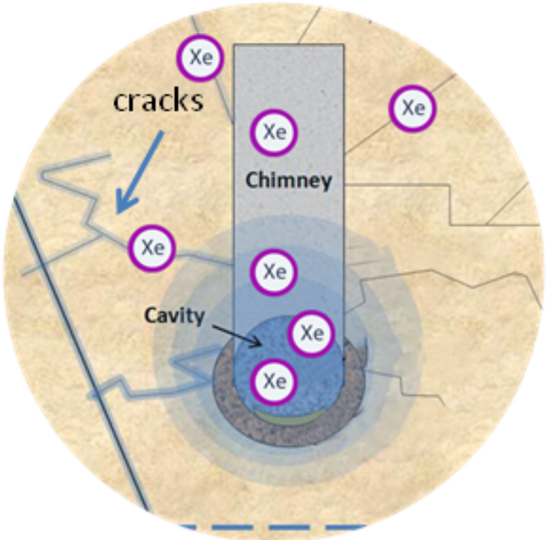
LA-UR-18-30304

Research Team

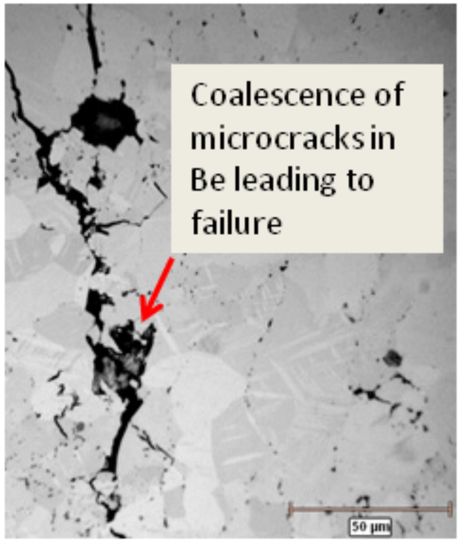
Collaborators

- Aric Hagberg
- Jeffrey Hyman
- Abigail Hunter
- Satish Karra
- Jamal Mohd-Yusof
- Bryan Moore, USC
- Chandramouli Nyshadham, BYU
- Dan O'Malley
- Dave Osthus
- Diane Oyen
- Cory Scott, UC Irvine
- Shriram Srinivasan
- Rosalyne Tchoua, U Chicago
- Hari Viswanathan
- Allon Percus, Claremont Grad U

Timely Topics



Gas Migration from a Nuclear Test
- Static fractures with gas flow



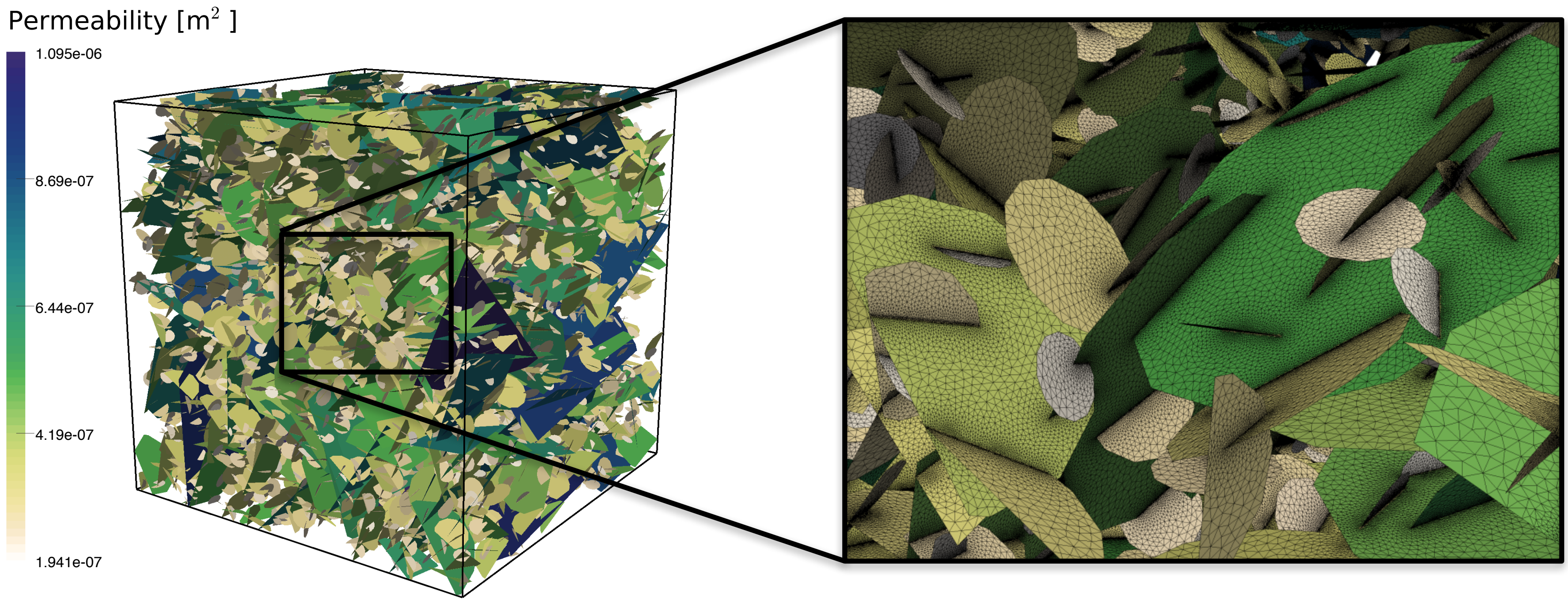
LANL Brittle Fracture Experiment
- Dynamic fracture growth

» Nuclear Nonproliferation: Gas migration through static fractures from an underground nuclear test.

» Brittle Material Failure: Dynamic fracture propagation and failure in brittle materials due to loading, e.g. Be alloys in aircraft wings, ceramics.

High Fidelity Model: Flow and Transport

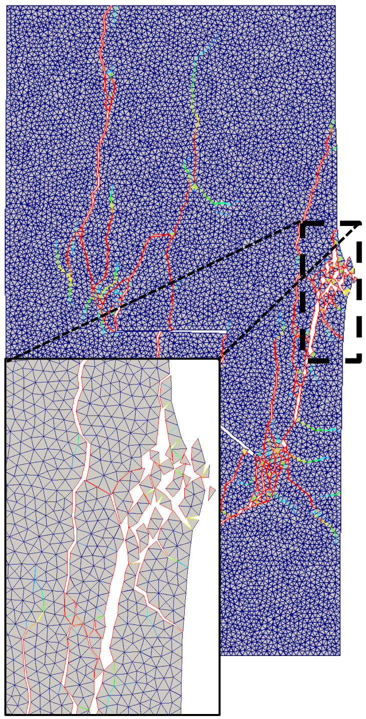
» Discrete Fracture Network (DFN) models are one tool to simulate flow and transport through fracture networks in low permeability rocks.



7200 fractures

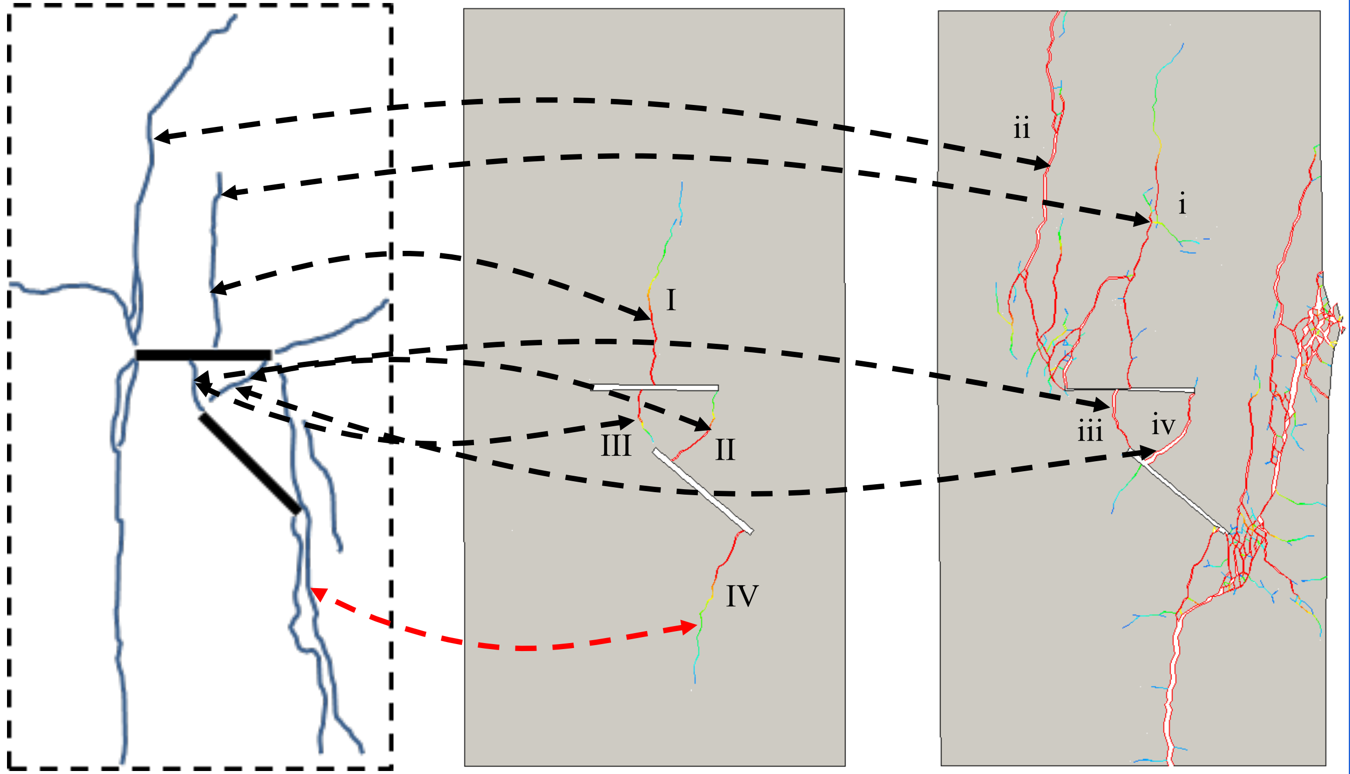
15,871,865 Mesh Nodes

High Fidelity Model: Fracture Propagation



HOSS Meshing

- 6 cm x 12 cm
- 20,000 2-D elements
- 140 fractures
- 4 hours w/ 400 processors
- Petabytes of data in 3D



Experiments

FDEM – Onset of Damage

FDEM – Final Damage

HOSS Validation Example

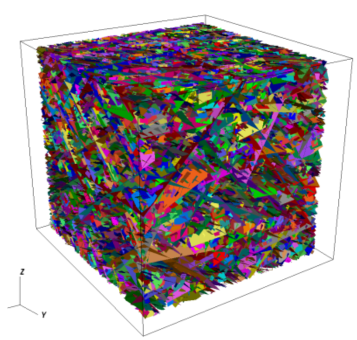
Problem Overview

<u>IS&T Challenge</u>: Structured systems represent a broad class of problems where the geometry and connectivity are critical to behavior.

» Representing structure is computationally intensive (Data Science at Scale)

» 1000s of runs needed to constrain topological uncertainty (UQ & Data Assimilation on reduced order models)

High Fidelity Fracture Models
1 cell has 10^6-10^{10} dof

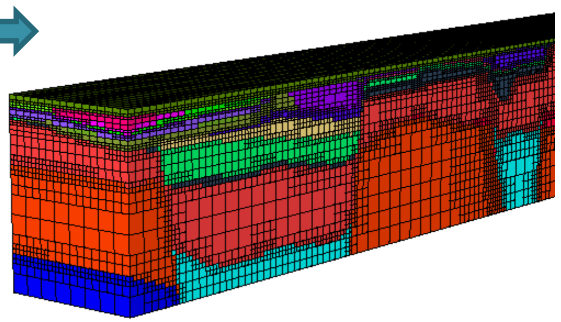


How do we efficiently relay microstructure geometry and topology to macroscale models?

?

Method must scale in a UQ framework to account for uncertainties in geometry / topology

Continuum Model
Averaged microstructure properties ~10^3 – 10^4 cells

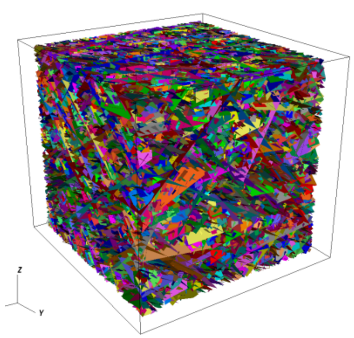


Our Novel Solution

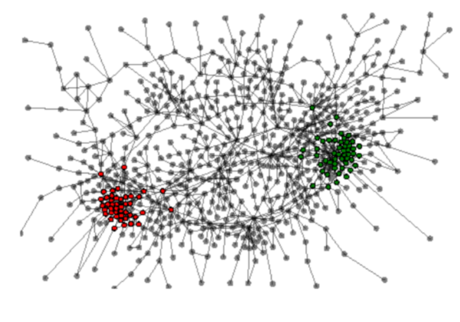
Computationally intensive grids → Efficient graphs → Tractable UQ cycles

(Petabytes of data) **(Terabytes of data)**

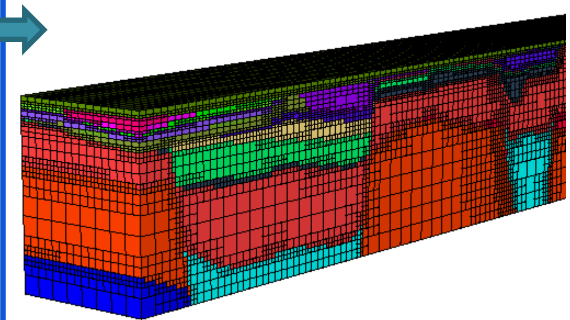
High Fidelity Fracture Models
1 cell has 10^6-10^{10} dof



Equivalent Graph Representation retains Geometry and Connectivity
10^3-10^7 degrees of freedom



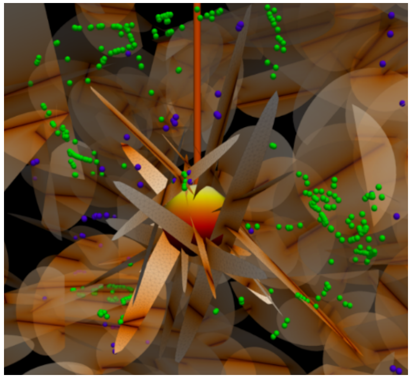
Continuum Model Averaged microstructure properties ~ 10^3 – 10^4 cells



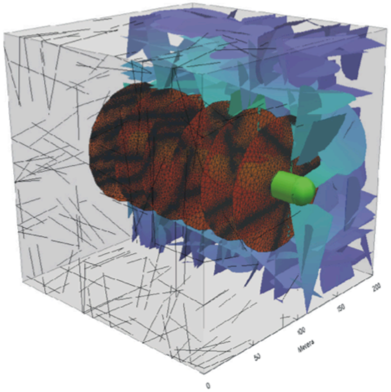
Graph G(size, length, location, orientation, betweenness centrality, eccentricity …)

Subsurface Flow and Transport Modeling

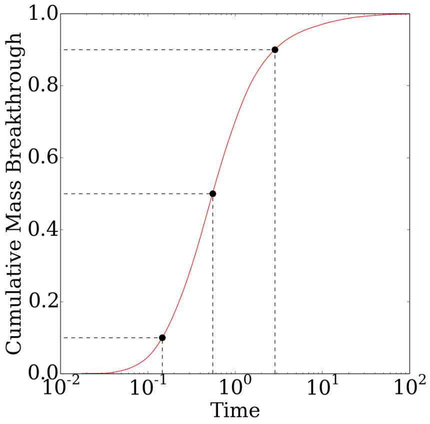
***Goal*: Represent flow and transport in subsurface fracture networks efficiently and integrate into a robust UQ framework.**



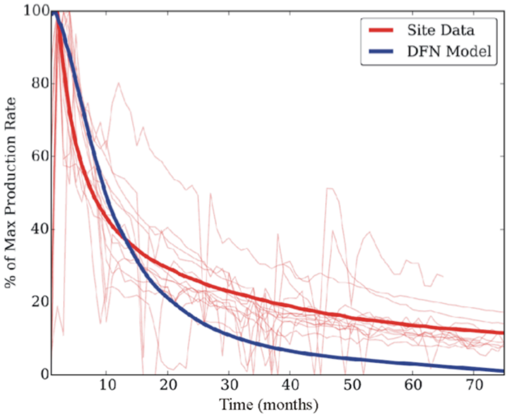
Nuclear Nonproliferation



Hydraulic Fracturing in Shale

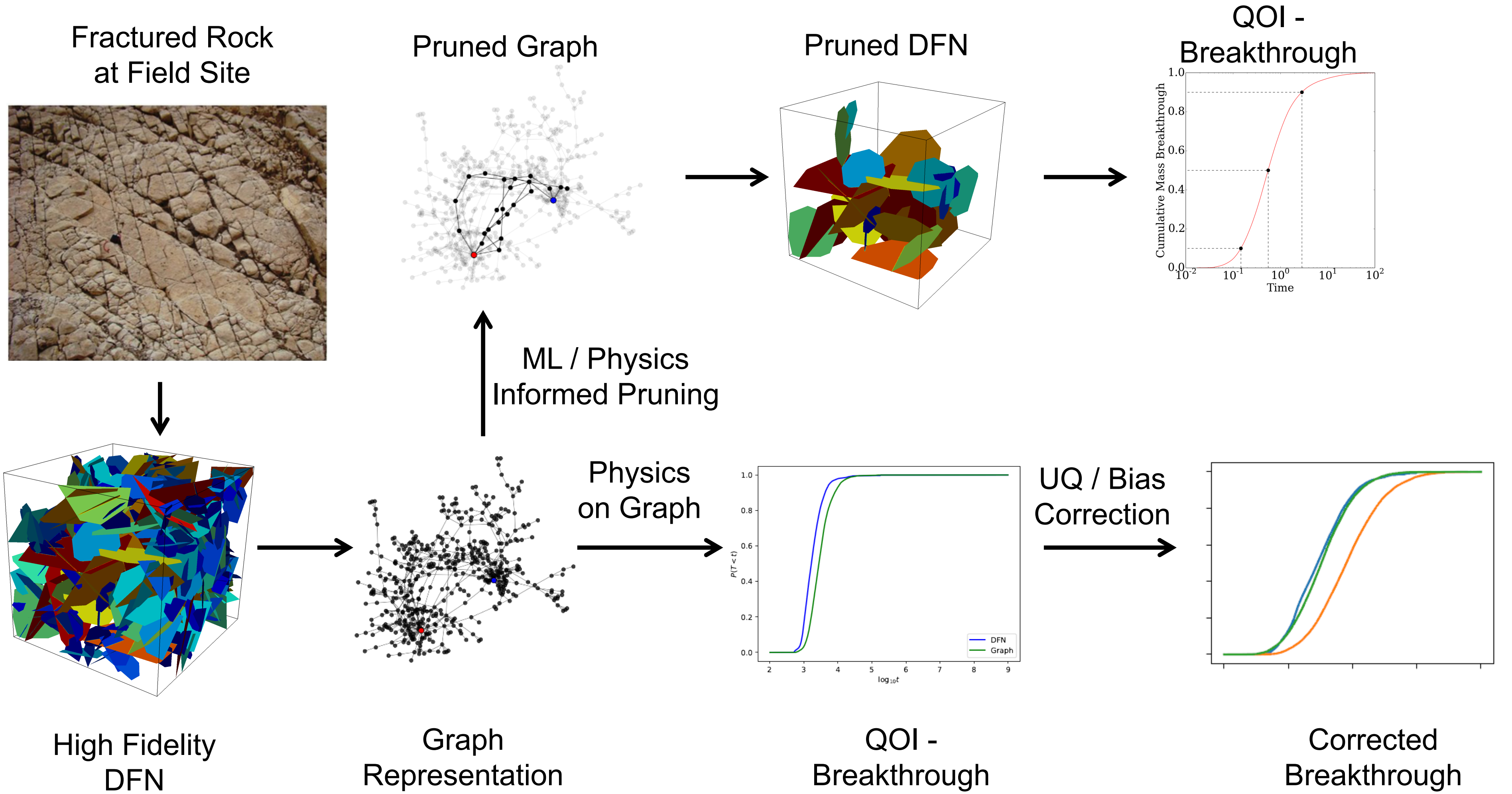


Quantity of Interest: First Arrival Times in the Breakthrough curve



Quantity of Interest: Entire Breakthrough curve

Subsurface Modeling Workflow

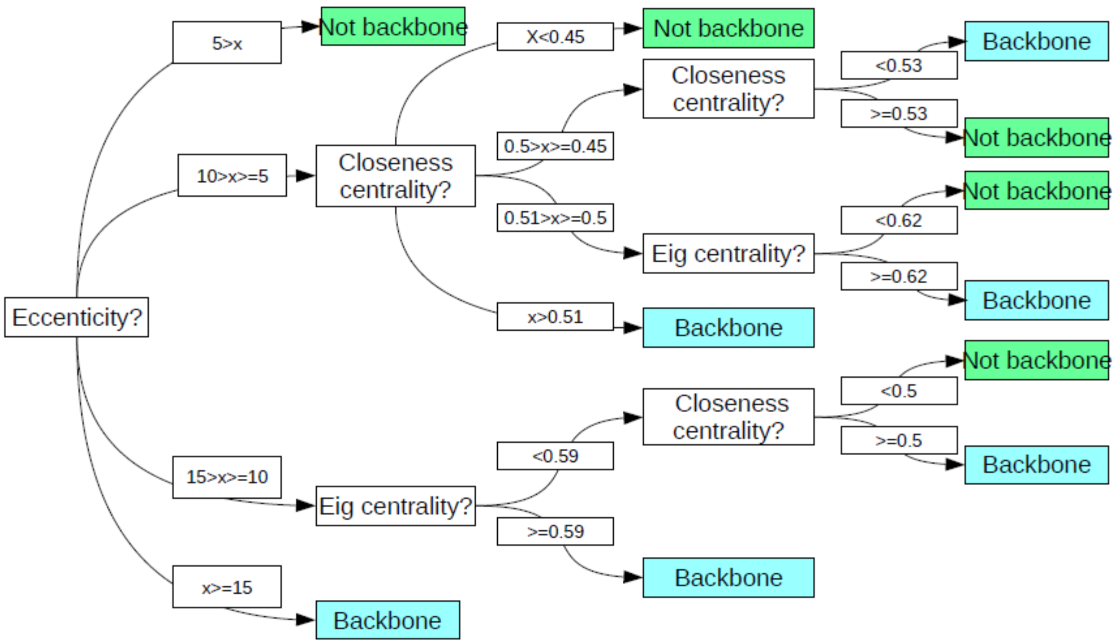


G. Srinivasan, J.D.Hyman, D.Osthus, B.Moore, D.O'Malley, S.Karra, E.Rougier, A.Hagberg, A.Hunter, and H. Viswanathan. Quantifying topological uncertainty in fractured systems using graph theory and machine learning. Nature Scientific Reports, 2018.

Data Driven Pruning with Machine Learning

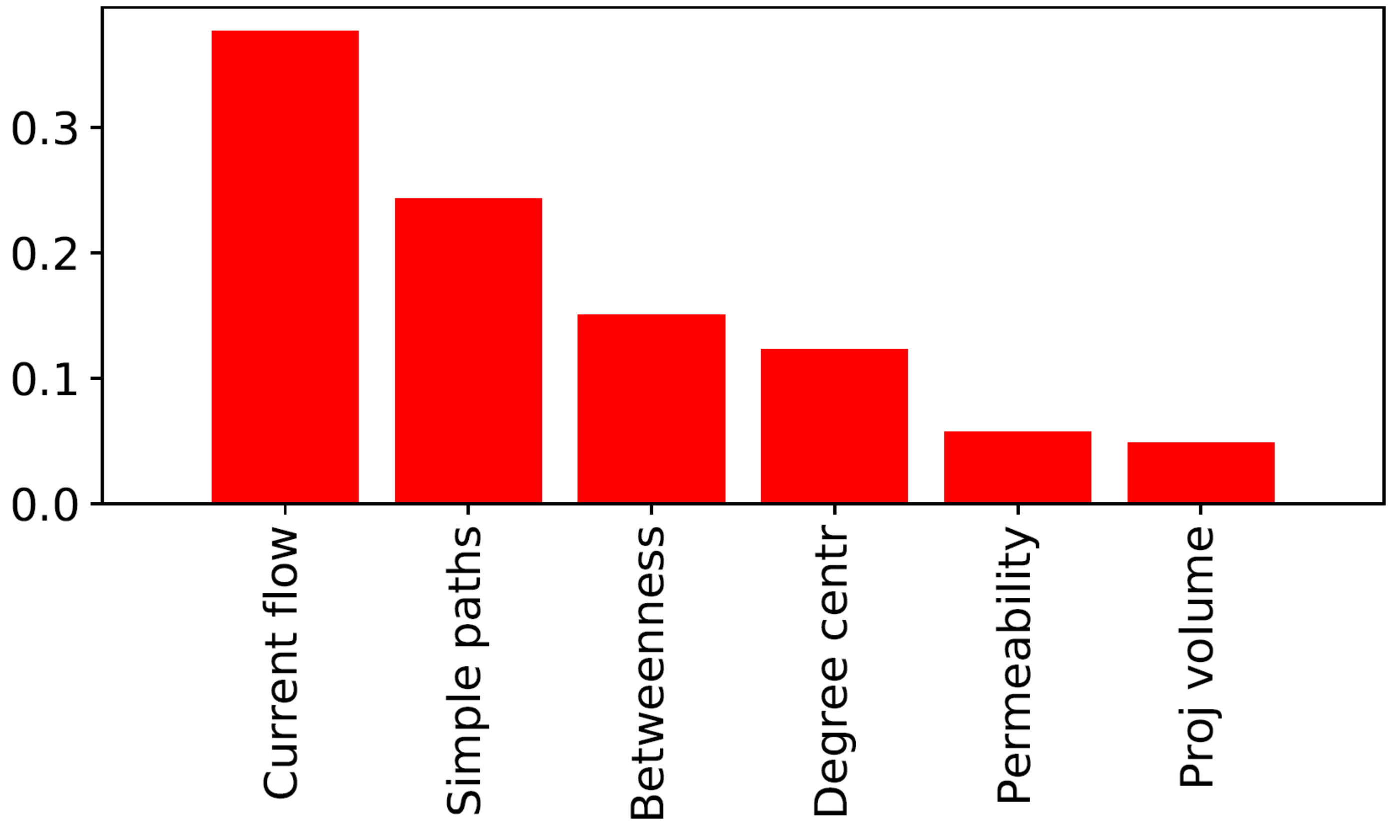
Use Random Forest and SVM to identify flowing backbone

- » Classification problem on whether a vertex is on the flowing backbone
- » 1st mapping where network topology properties are inherite by fractures (vertices)
- » Leave out cross validation: train on 80, test on 20
- » Compare transport breakthrough curves
- » Ensure a connected path
- » False positives better than false negatives



Example Decision Tree

Feature Importance

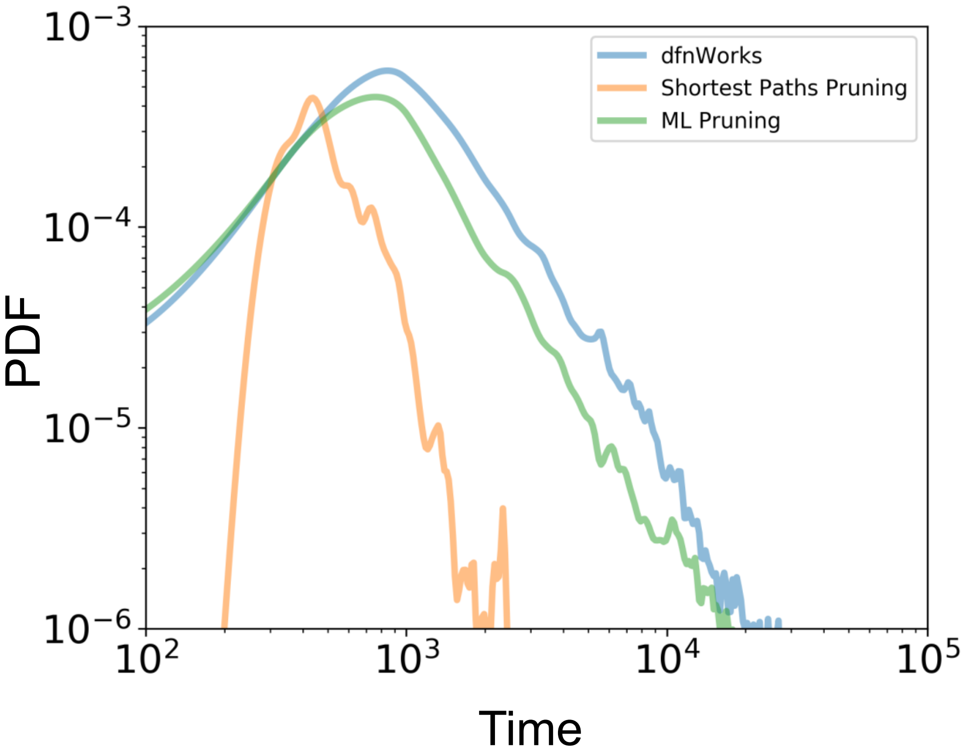


Combination of topological and flow features

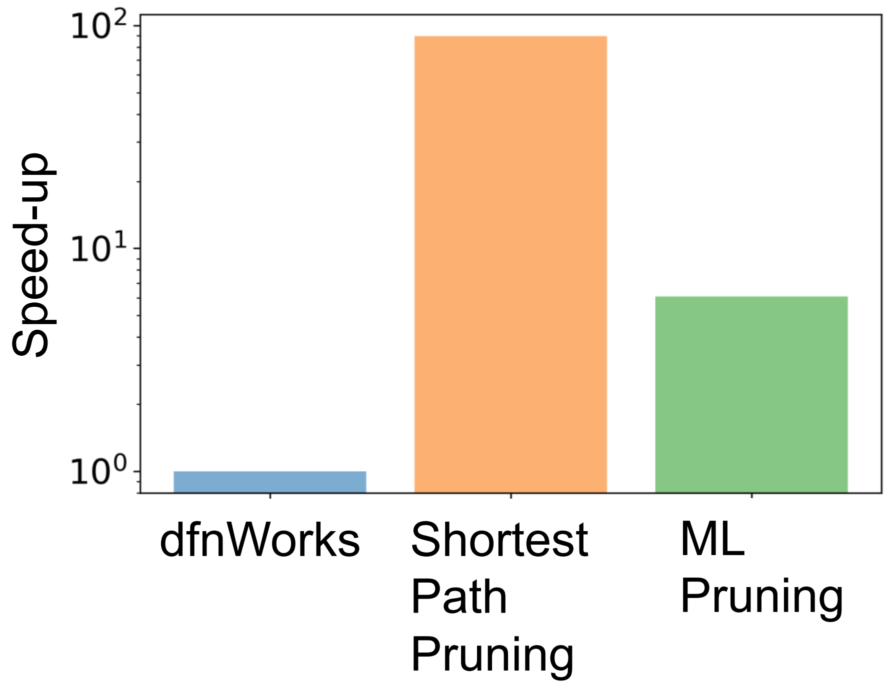
M. Valera, Z. Guo, P. Kelly, S. Matz, A. Cantu, A.G. Percus, J. D. Hyman, G. Srinivasan, and H.S. Viswanathan. Machine learning for graph-based representations of three-dimensional discrete fracture networks. *Computational Geosciences*, 2018

Backbone Identification Through Machine Learning

QOI: Breakthrough Curves

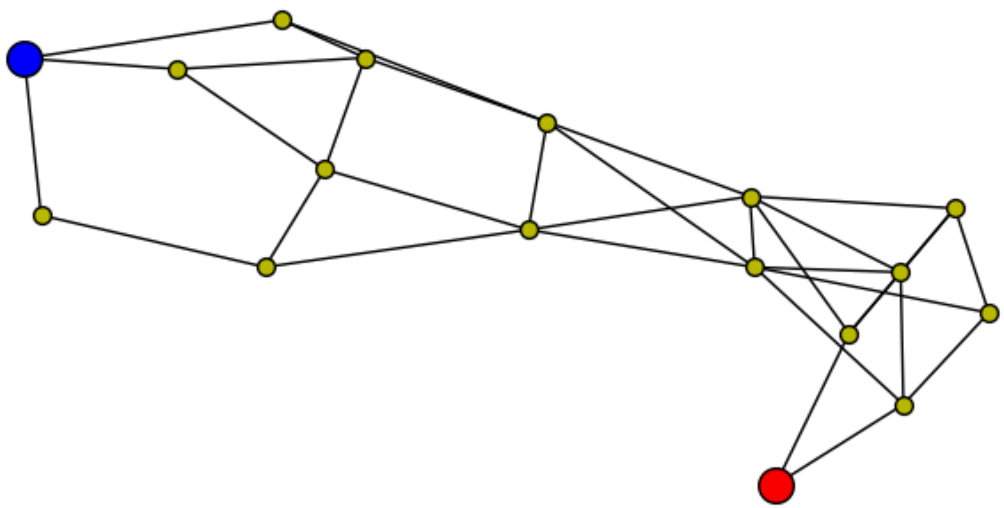


Computational Performance



Sufficiently accurate pruned network with 33% of the original fractures

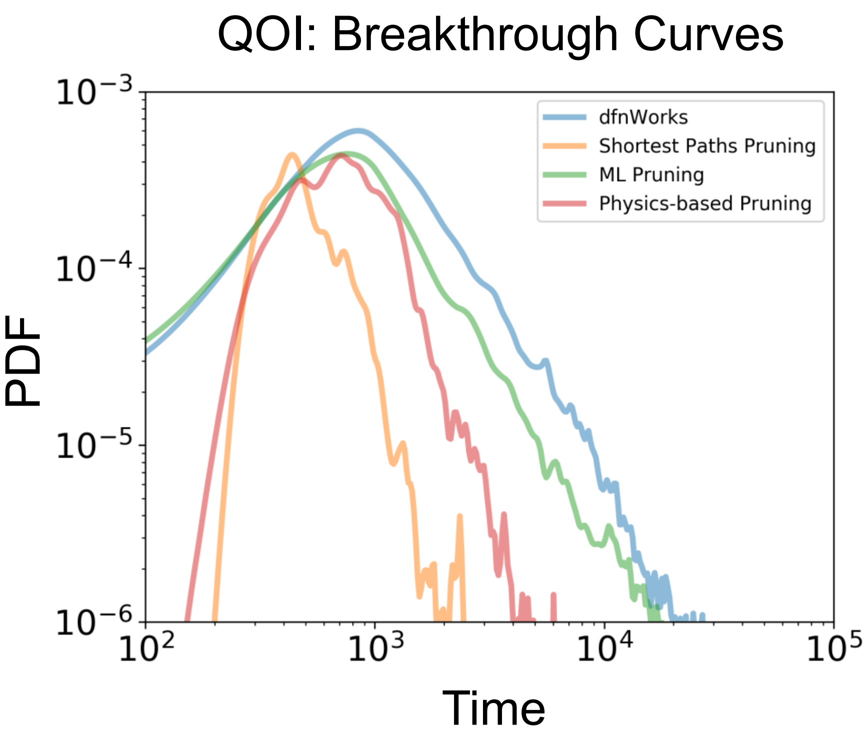
Physics-based Pruning

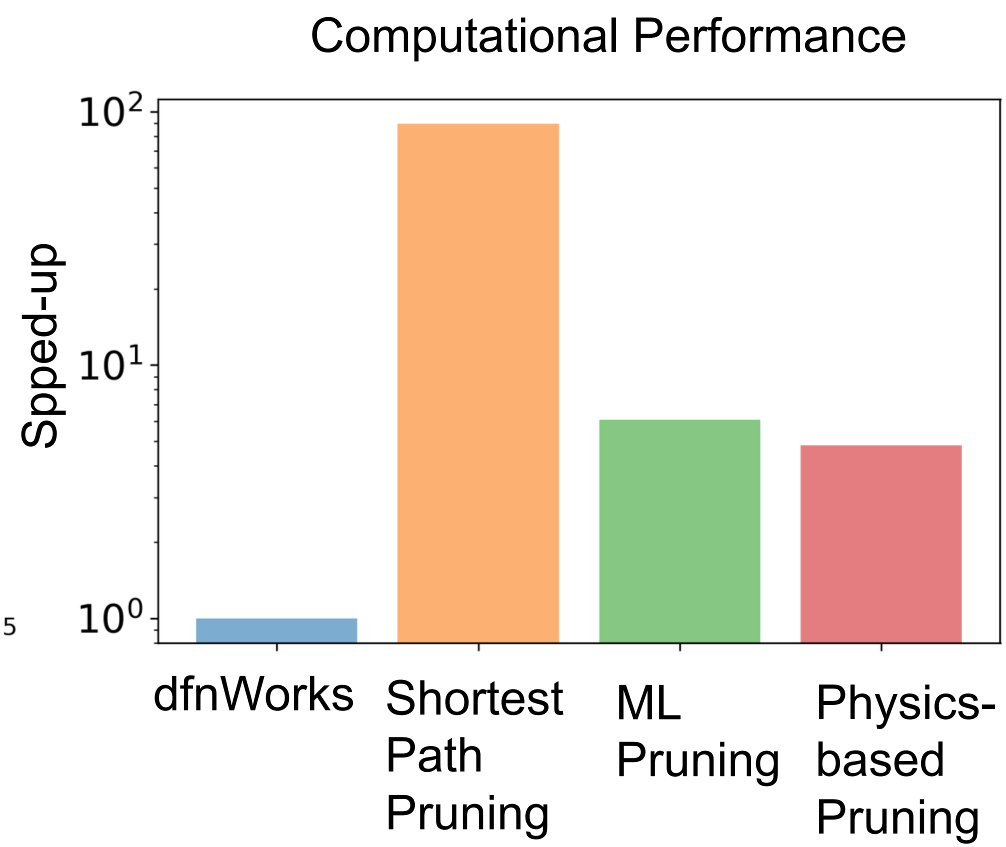


» Threshold on flux on edges ensuring connected path
» Hydrologic properties like permeability can be inherited by edges

Sufficiently accurate pruned network with only 35% of original network

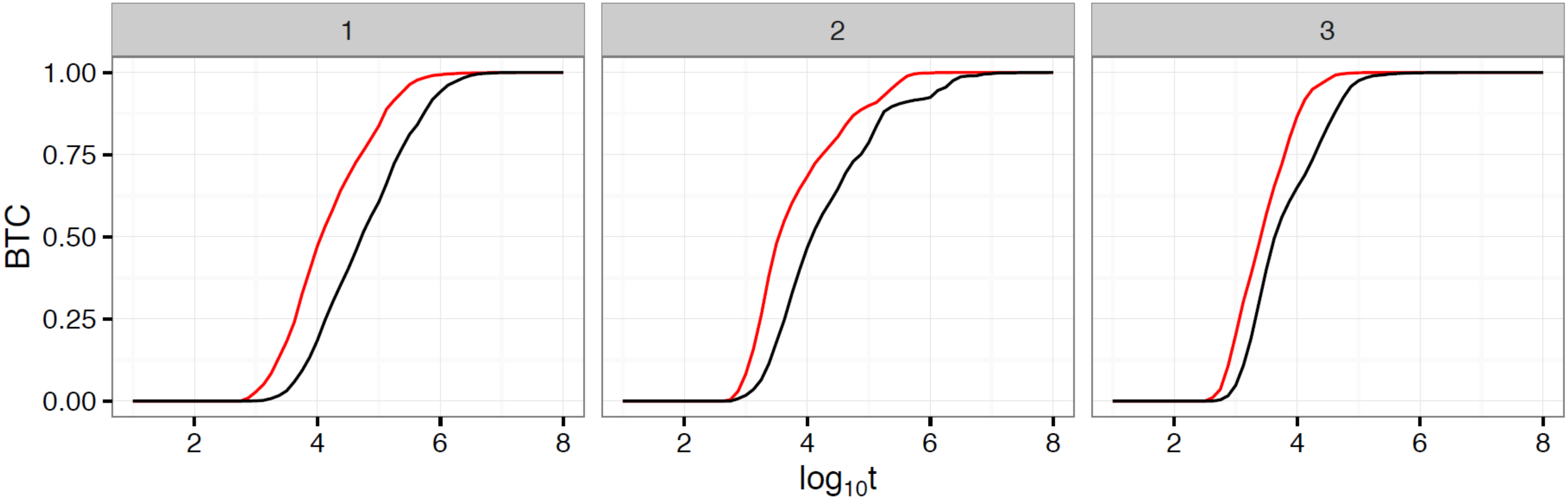
Physics-based Pruning Results





S. Srinivasan, J.D Hyman, S. Karra, H. Viswanathan, G. Srinivasan System Reduction of Discrete Fracture Networks through flow-physics on graph representations, Advances in Water Resources. 2018

Transport on the Graph with Multiple Realizations



The discrepancy from DFN and graph transport is systematic

S. Karra, D. O'Malley, J. D. Hyman, H.S. Viswanathan, and G. Srinivasan. Modeling flow and transport in fracture networks using graphs. *Phys. Rev. E*, 2018

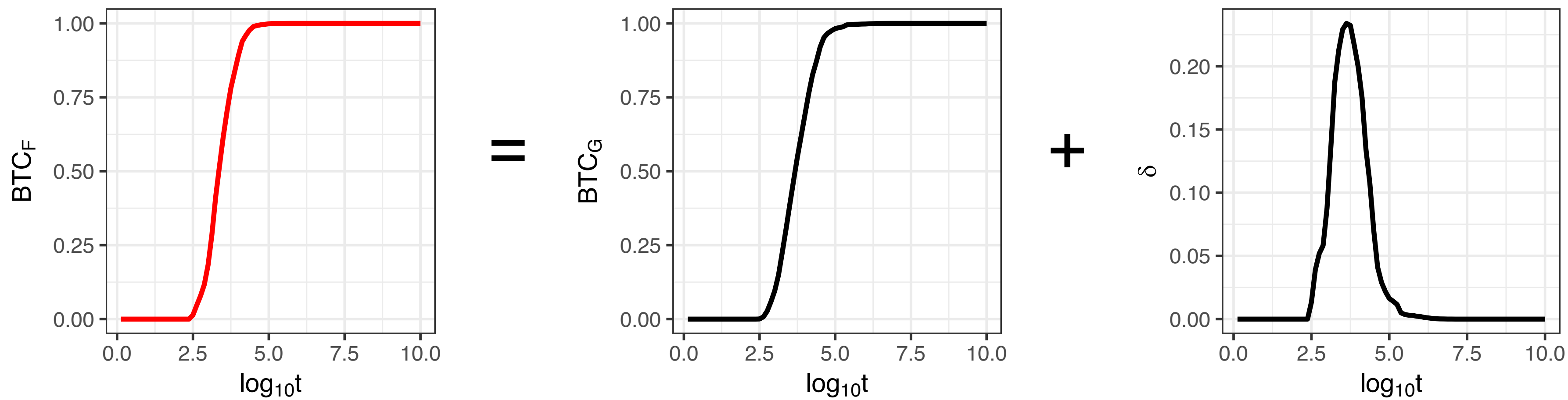
Los Alamos NATIONAL LABORATORY

UQ Model

$$\color{red}{BTC_F(t)} = BTC_G(t + \theta) + \delta(t)$$

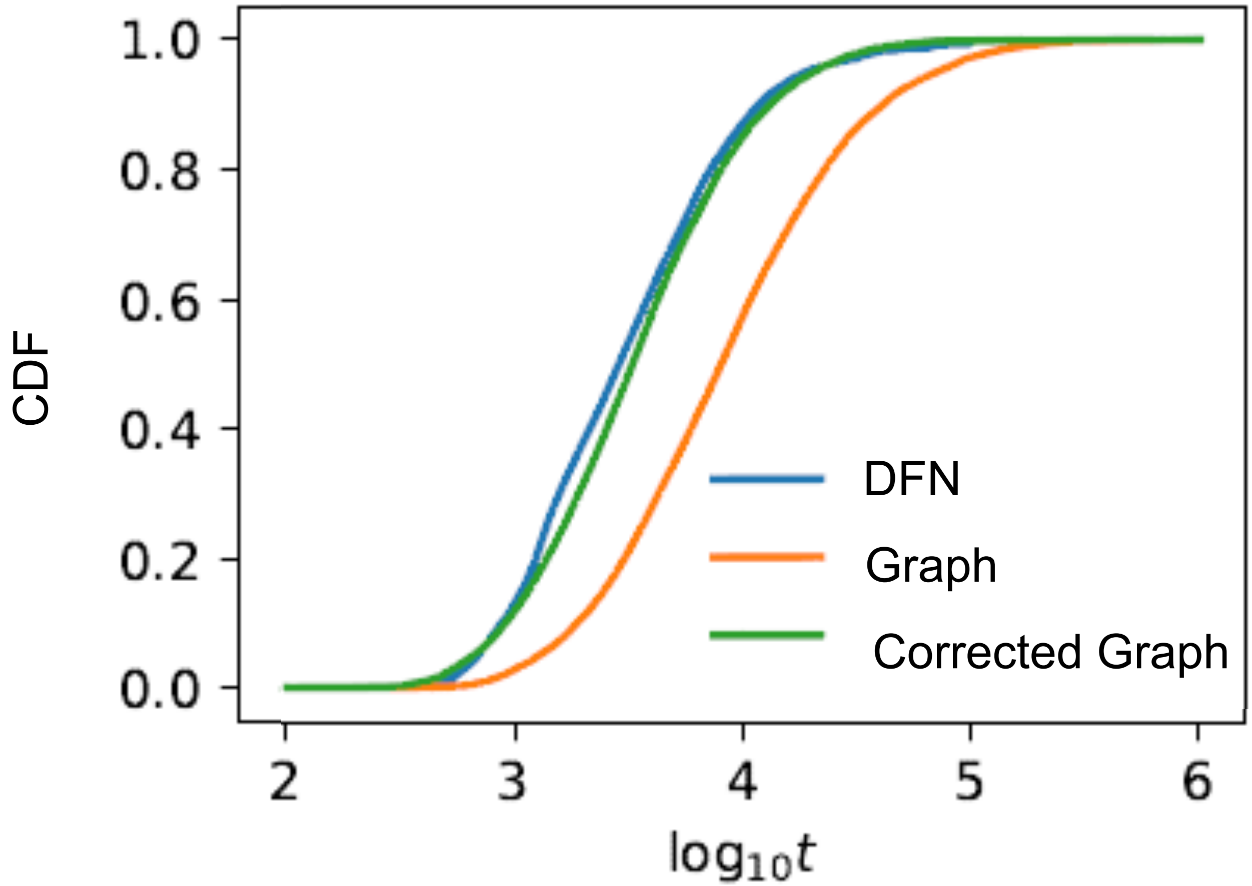
Observation Calibration Discrepancy

Bayesian UQ method -- Kennedy & O'Hagan, 2001



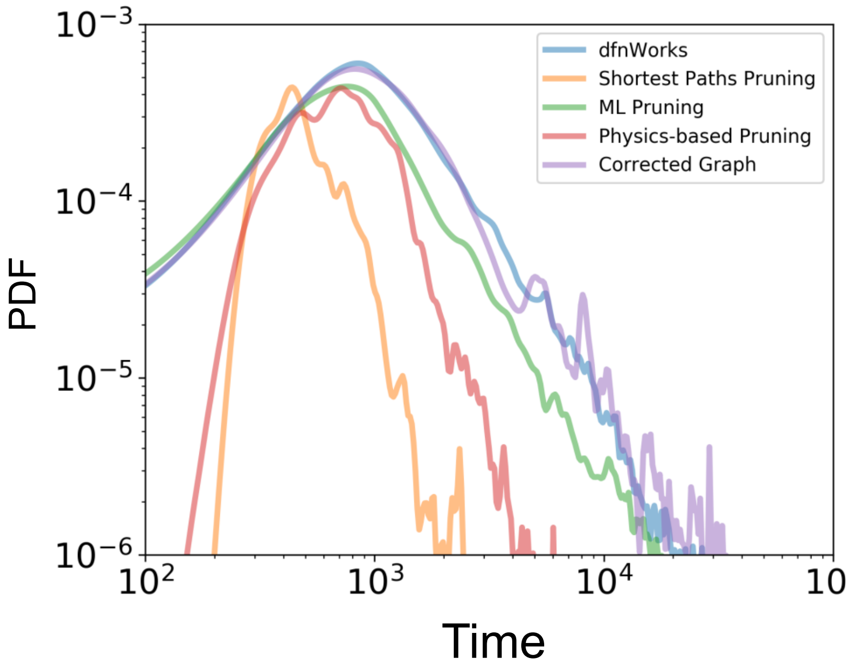
G. Srinivasan, J.D.Hyman, D.Osthus, B.Moore, D.O'Malley, S.Karra, E.Rougier, A.Hagberg, A.Hunter, and H. Viswanathan. Quantifying topological uncertainty in fractured systems using graph theory and machine learning, *Scientific Reports*, 2018

Correcting Systematic Deviation

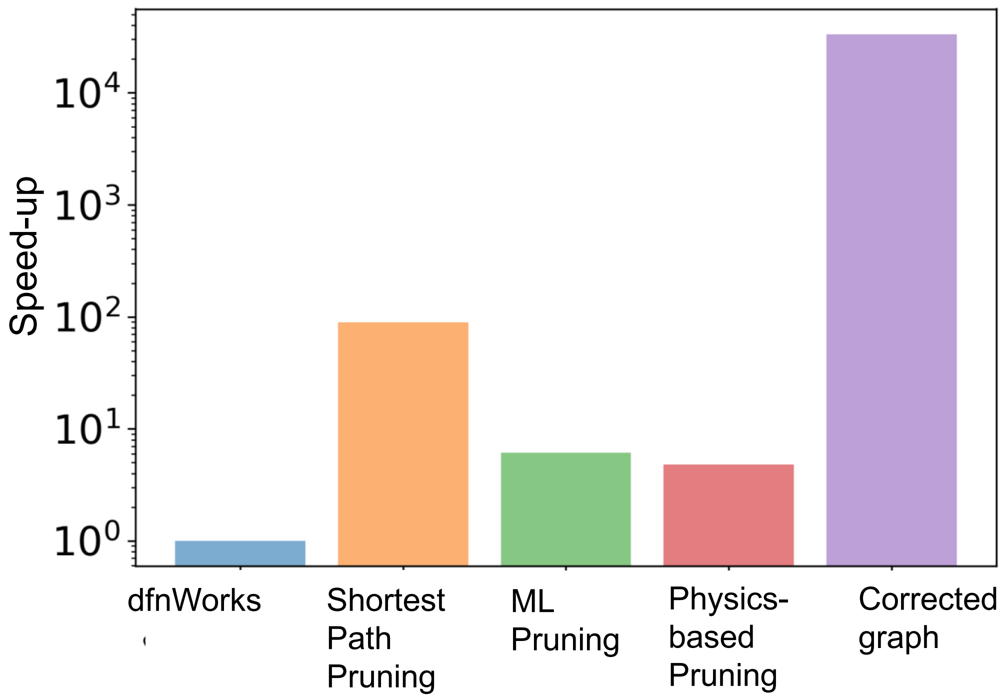


Summary

QOI: Breakthrough Curves



Computational Performance



We can tailor the reduced order model depending on the QOI:

» Quick shortest path calculation if only early arrival is needed

» ML or physics-based pruning is effective but still requires mapping back to DFN(10X-100X speedup)

» Transport on the graph is 4 orders of magnitude faster but accurate for more complex cases?

H. S. Viswanathan, J. D. Hyman, S. Karra, D. O'Malley, S. Srinivasan, A. Hagberg, and G. Srinivasan. Advancing graph-based algorithms for predicting flow and transport in fractured rock. *Water Resour. Res.*, 2018

Summary – Subsurface Flow and Transport

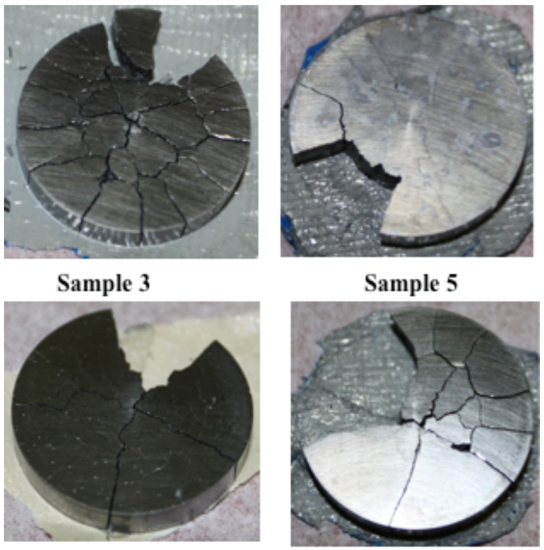
» We can tailor reduced order models to mimic high fidelity DFNs based on the quantity of interest

» Breakthrough curve QOI is quite forgiving so exact percolating path is need not be replicated by reduced order model

» For single phase flow and no in-fracture variability data driven and physics-based pruning are effective but graph transport is far more efficient and systematic deviations can be corrected

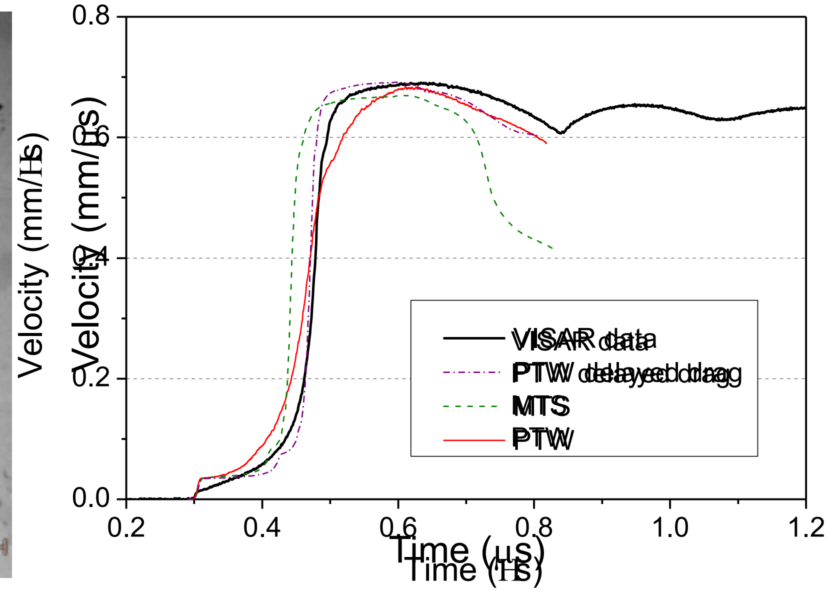
Brittle Failure Modeling

Goal: **Represent physics of fracture evolution and coalescence resulting in an improved continuum-scale damage model for use in hydro-codes**



Sample 3 Sample 5

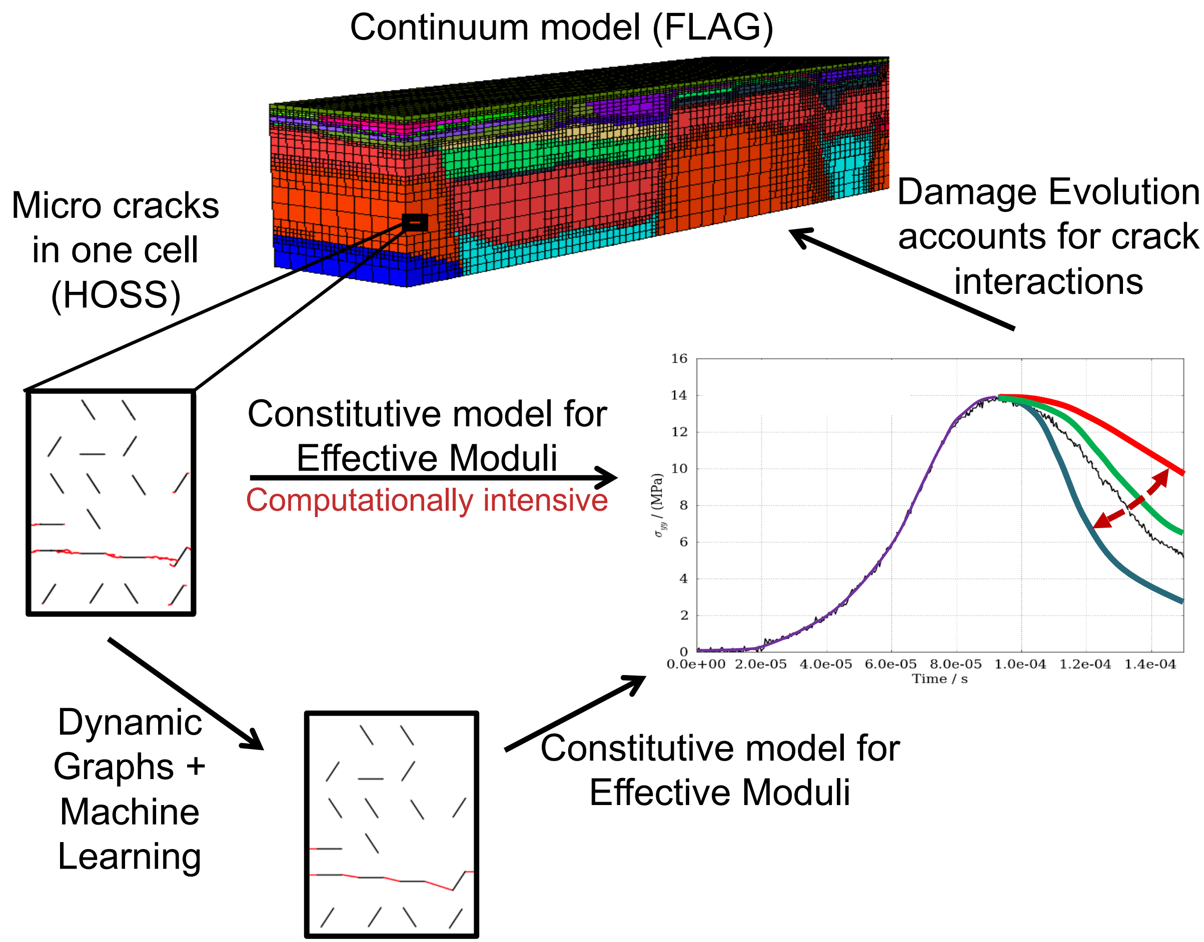
Sample 6 Sample 7



LANL Brittle Fracture Experiment, Cady, et al. LA-UR 11-06976 (2011)

Quantities of Interest: Time to Failure, Crack statistics and Damage Evolution

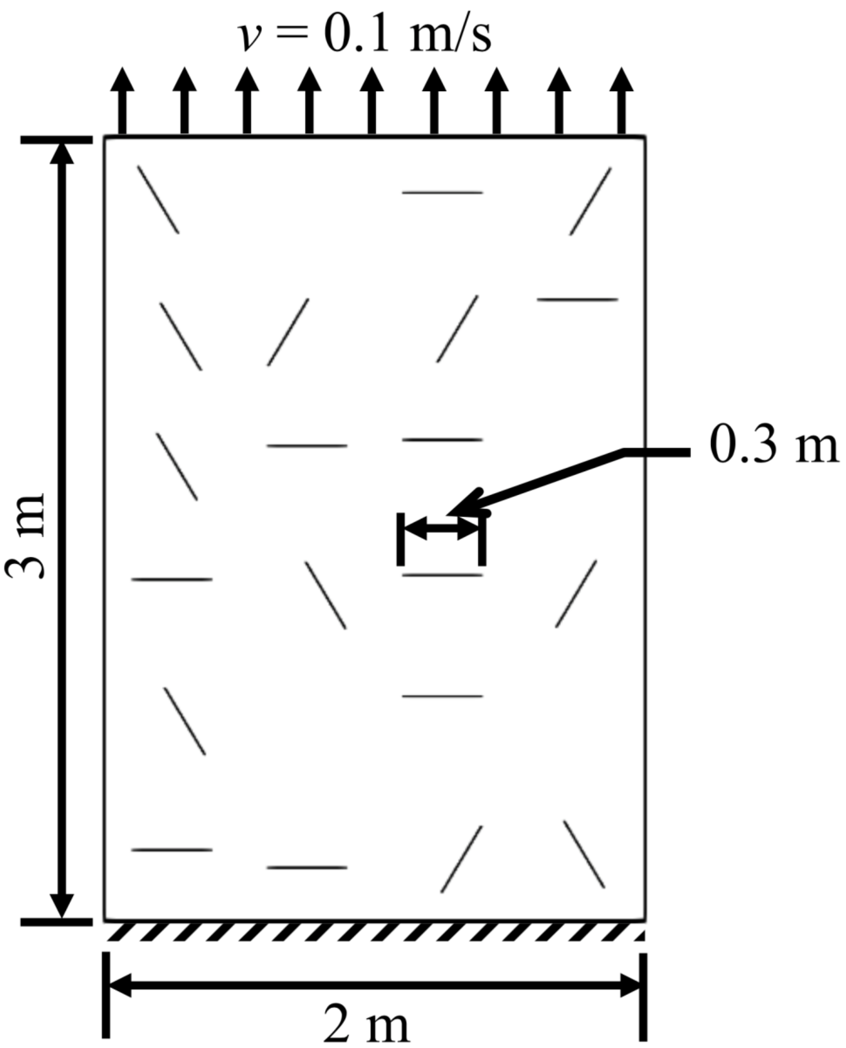
Brittle Failure Modeling Workflow



G. Srinivasan, J.D.Hyman, D.Osthus, B.Moore, D.O'Malley, S.Karra, E.Rougier, A.Hagberg, A.Hunter, and H. Viswanathan. Quantifying topological uncertainty in fractured systems using graph theory and machine learning. Nature Scientific Reports, 2018.

HOSS runs: training/testing data

- 150 HOSS runs were performed for training (~1600 CPU hours each)
- 35 HOSS runs were used for testing
- 2D grid on a 2m-by-3m domain
- Bottom boundary is fixed, top boundary moves up at 0.1m/s, side boundaries move freely
- Each HOSS run contains 20 cracks
- Each crack is…
  - …0.3m in length
  - …randomly oriented at either 0, π/3 or 2π/3 radians
  - …randomly located within a 6-by-4 grid



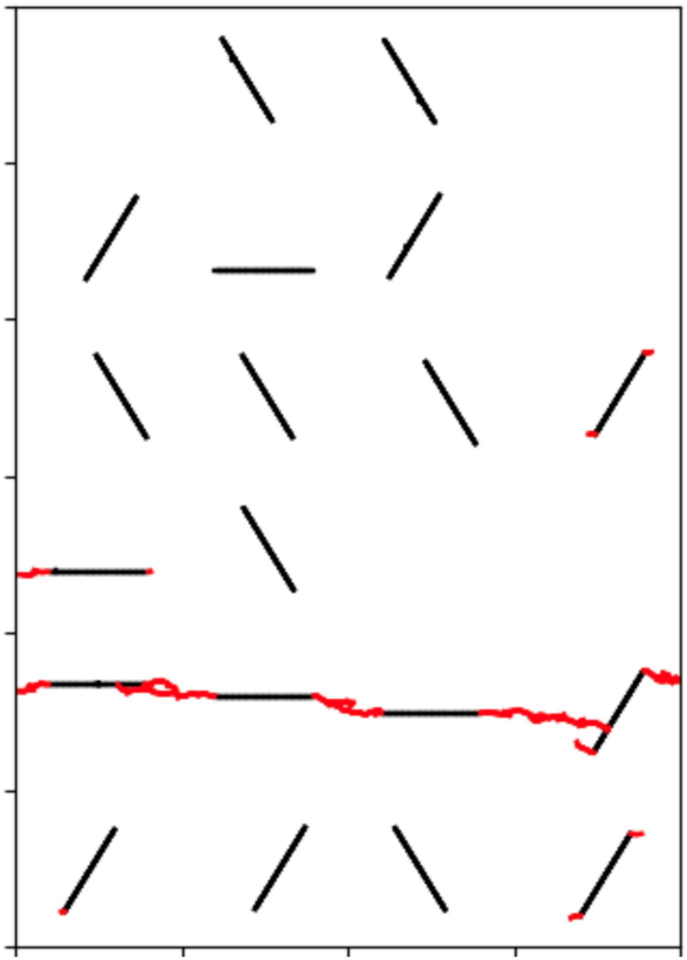
Overview of methods

- **A diverse set of methods were explored that enabled us to find what works and what does not with cross-pollination between the methods**
  - Dynamic graphs
  - Machine learning/data driven
  - Simplified physics
- **All of these methods are orders-of-magnitude faster than HOSS**
  - Typically run in O(CPU seconds) or less in comparison to HOSS which runs in O(CPU months)
  - Our goals have been to understand rather than optimize, so there is room for further speed-ups
- **There are no machine learning methods in the literature that bridge the gap between micro-scale and macro-scale codes**

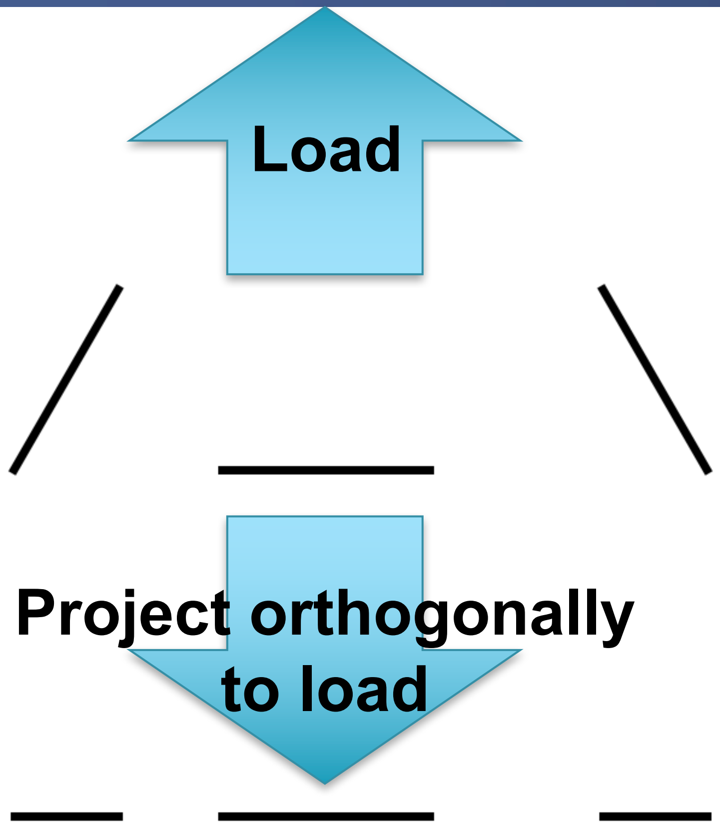
A. Hunter, B. Moore, M. Mudunuru, V. Chau, R. Miller, R. Tchoua, C. Nyshadham, S. Karra, D. O'Malley, E. Rougier, H. Viswanathan, G. Srinivasan. Reduced-Order Modeling through Machine Learning Approaches for Brittle Fracture Applications., Journal of Computational Materials Science, accepted.

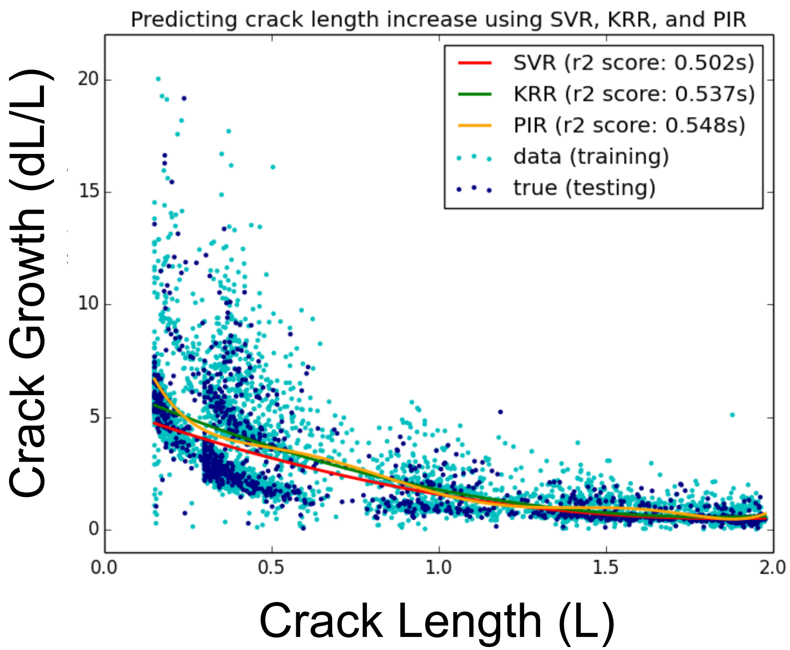
Evaluating the models

- **We evaluate the models on two criteria**
  - The ability to predict the failure path
  - The ability to predict the time at which failure occurs (for some methods)
- **Failure means a crack has formed that connects the left and right boundaries**
- **Predicting the failure path means identifying the set of initial cracks that are part of the crack connecting the left and right boundaries**
  - This metric is harsher than typical ML metrics (e.g., “human level performance”) and gives no ”partial credit”



Orthogonal Projection Approach





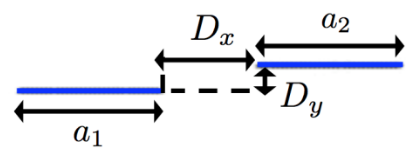
Train to predict dL based on (projected) crack length

» Method was originally developed for a dataset with all horizontal fractures

» Adapted to this dataset based on the idea that Mode I (tensile) failure is the dominant crack propagation mechanism in these scenarios

Micro-crack Pair Informed Coalescence (McPIC) Approach

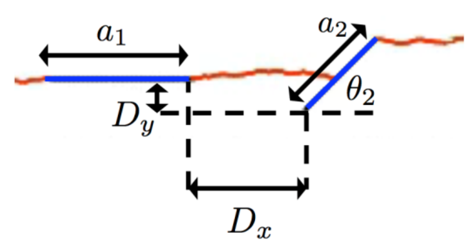
- Uses a two-step approach
- For each pair of cracks
  - Step 1: Classify crack pairs as either coalescing or not
  - Step 2: If coalescing, predict the time at which they coalesce
- The K_i are stress intensity factors
- D_B is the distance to the nearer of the left and right boundaries
- f(…) has been implemented via neural nets, random forest, and decision trees



Feature Vector: $f\left(a_1, a_2, \theta_1, \theta_2, D_x, D_y, K_1, K_2, D_B\right)$

↓

Coalescence Classification: **No**



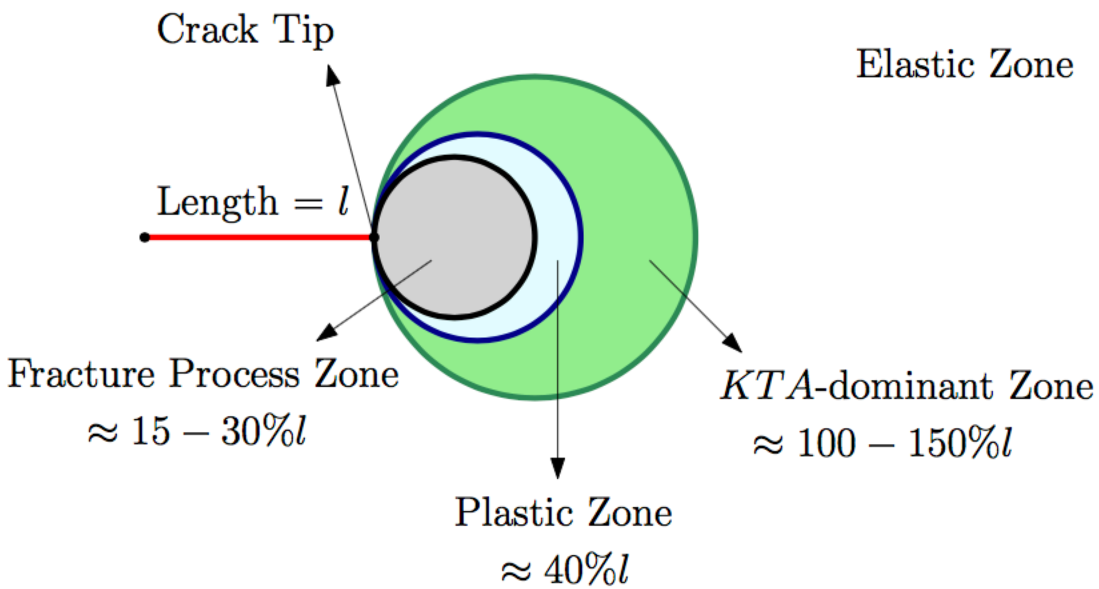
Feature Vector: $f\left(a_1, a_2, \theta_1, \theta_2, D_x, D_y, K_1, K_2, D_B\right)$

↓

Coalescence Classification: **Yes** — Regression

B. Moore, E. Rougier; G. Srinivasan; D. O'Malley; A. Hunter and H. Viswanathan, Predictive modeling of dynamic fracture growth in brittle materials with machine learning, *Computational Material Science*, 2018

Fracture Process Zone



» The fracture process zone (FPZ) is the zone where damage accumulates as cracks evolve

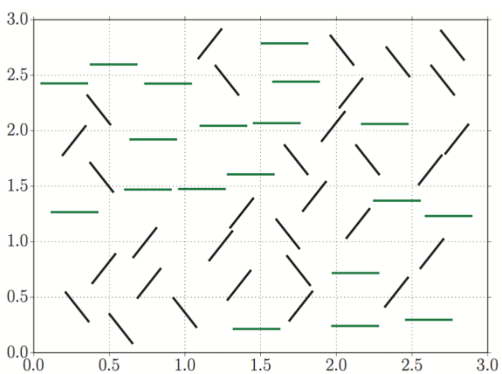
» FPZ contains micro-cracks near the crack tip

» These micro-cracks merge and extend the crack

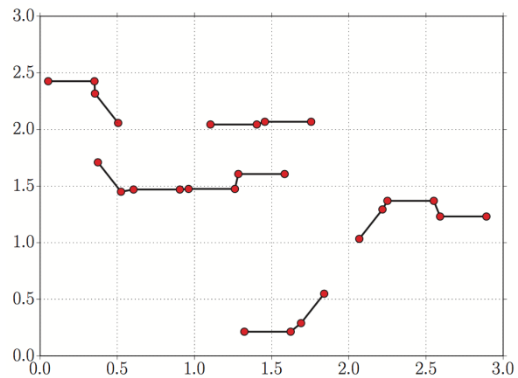
» KTA-dominant zone is an asymptotic elastic zone that transitions from elastic to plastic

M. K. Mudunuru, S. Karra, G. Srinivasan, E. Rougier, H. S. Viswanathan, and L. Margolin, Predicting paths of least resistance and time to failure for dynamic fracture using network theory, for Applied Soft Computing (In Prep.)

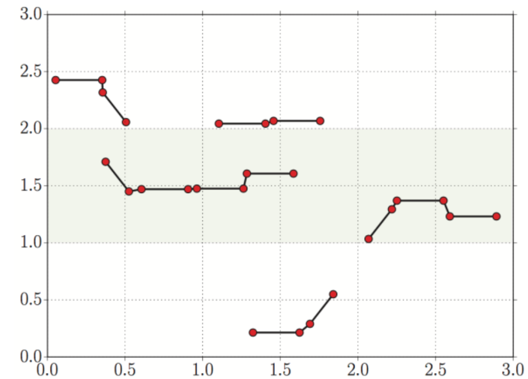
Network-based Fracture Process Zone Approach



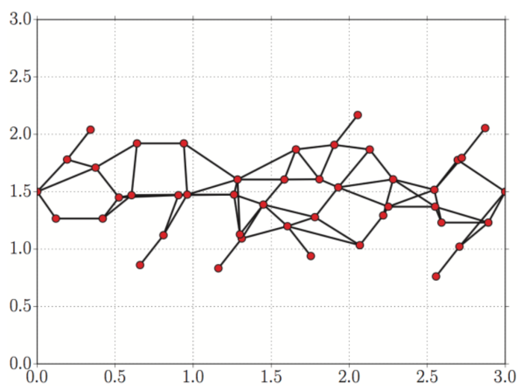
1. Identify cracks orthogonal to loading



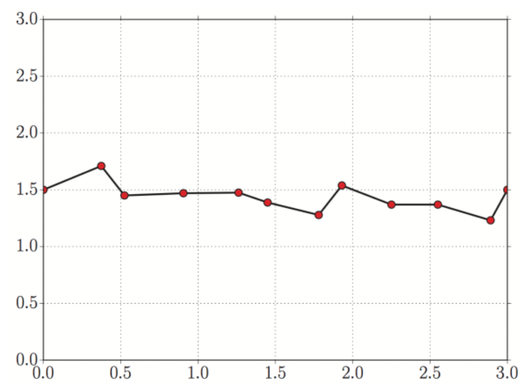
2. Use FPZ to identify longest crack



3. Identify failure region



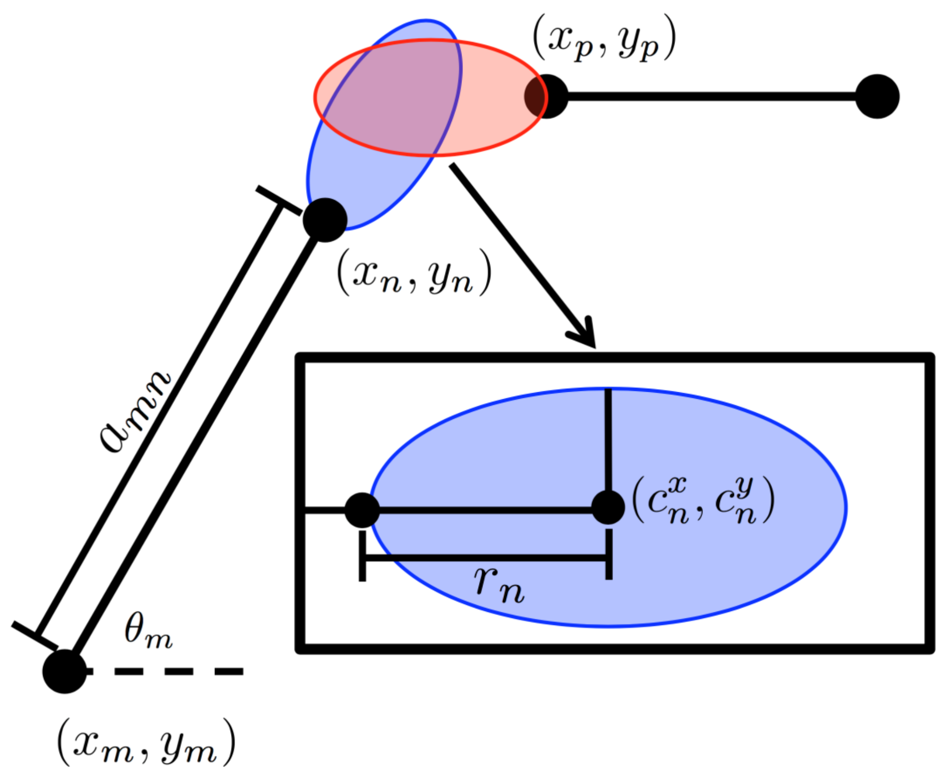
4. Prune via nearest neighbors analysis



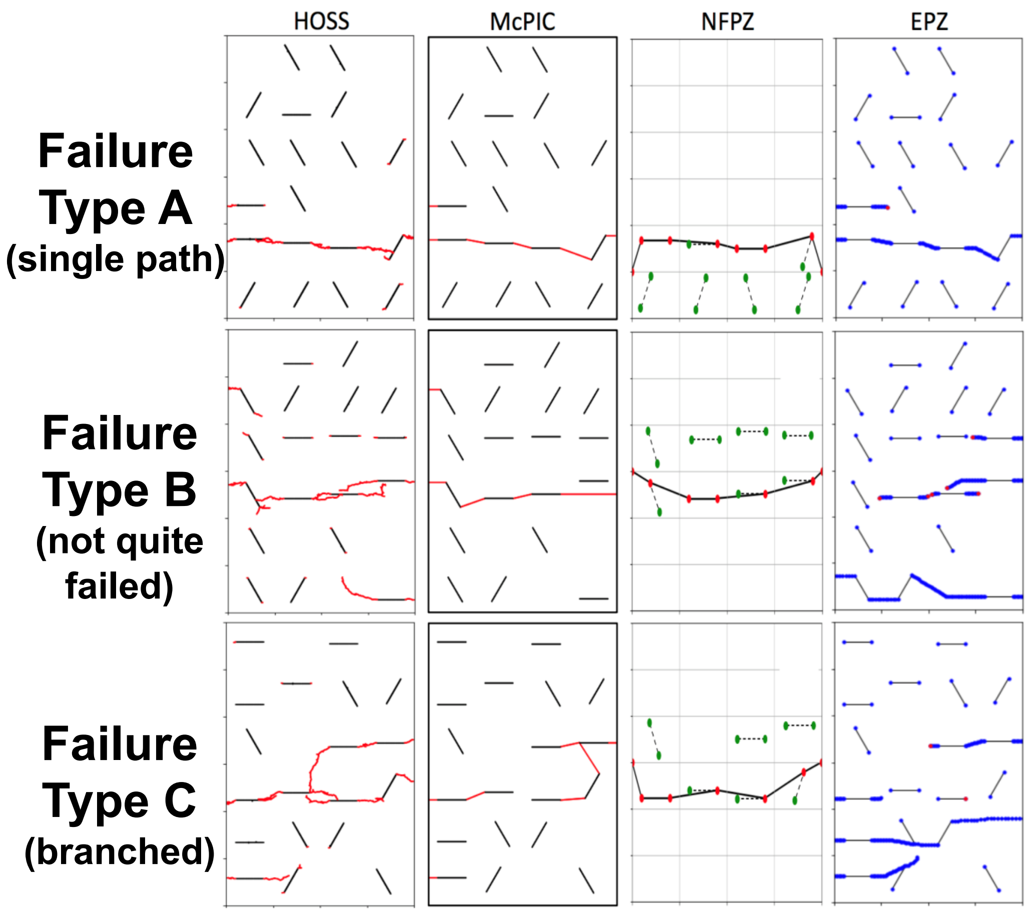
5. Identify failure path with shortest path

Elliptical Process Zone Approach

- Uses an elliptical process zone
- When two ellipses overlap, the cracks will coalesce
- $r_n = \gamma f(p) a_{mn}$ where p is a vector of material parameters
- Parameters such as γ and ellipse eccentricity are learned from data
- θ_m will be updated at each time step to orient the crack tip at (x_n, y_n) toward (x_p, y_p) as the crack tip grows

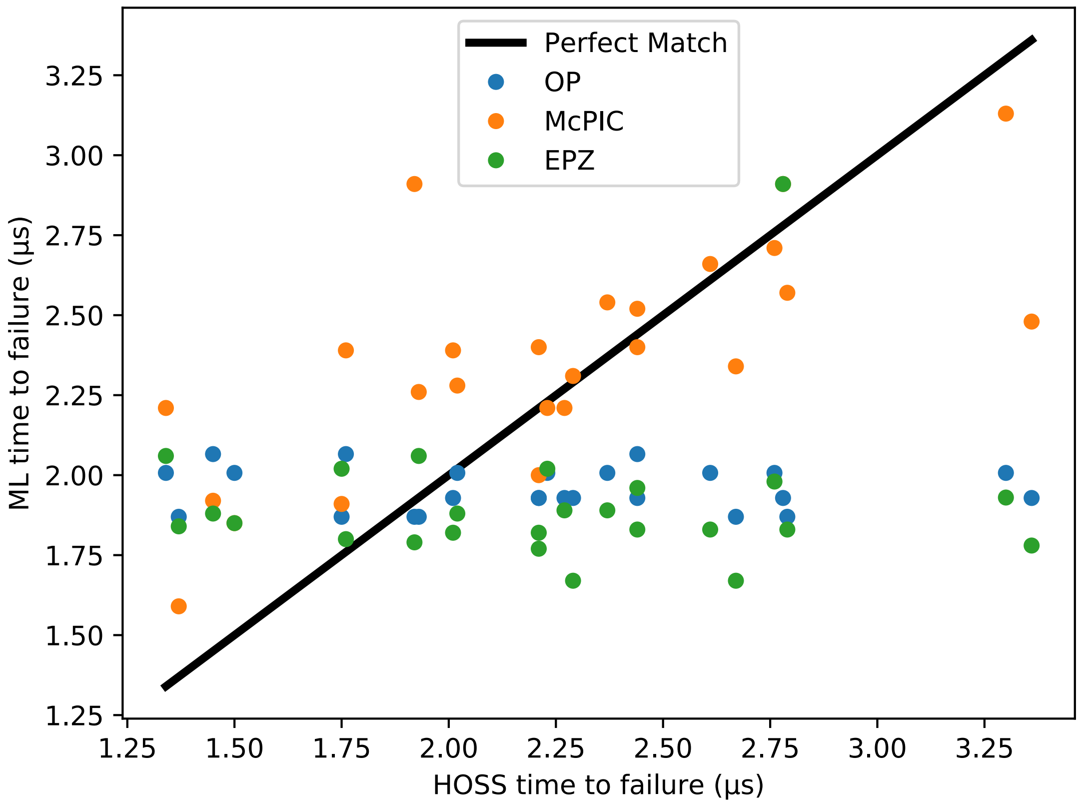


How well did the methods predict the path to failure?



Method	Type A	Type B	Type C
Orthog. Projection	9/20	2/10	0/5
McPIC	9/20	5/10	0/5
Network FPZ	6/20	2/10	0/5
Ellipse FPZ	4/20	1/10	0/5

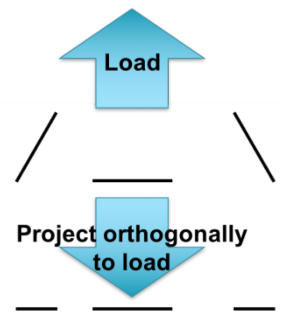
How well did the methods predict the time of failure?



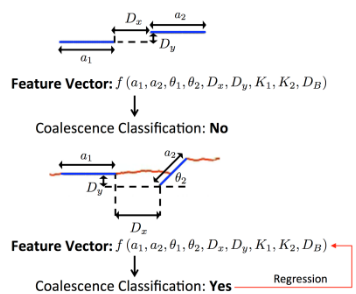
Only the orthogonal projection (OP), micro-crack pair informed coalescence (McPIC) and elliptical process zone (EPZ) approaches predict the time to failure

Lessons Learned: Goldilocks and the 4 models

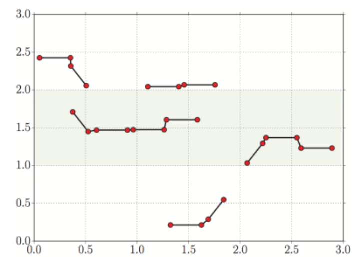
Orthogonal Projection



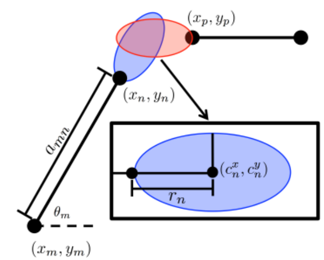
McPIC



Network FPZ



Elliptical FPZ



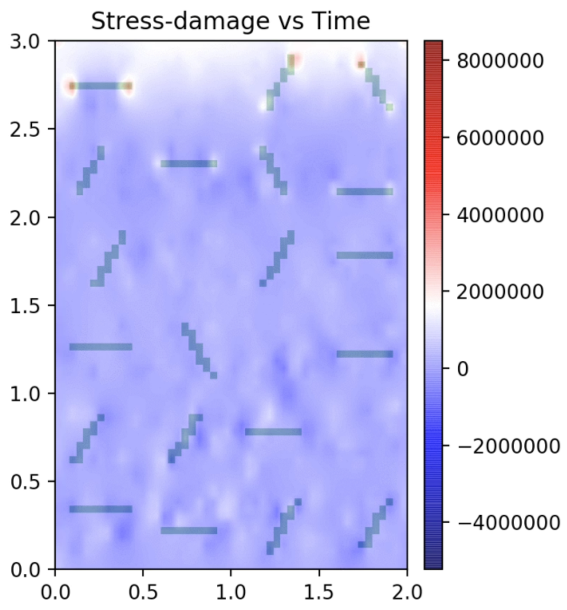
Not enough learning from data

Richer feature set enables learning, some heuristic physics

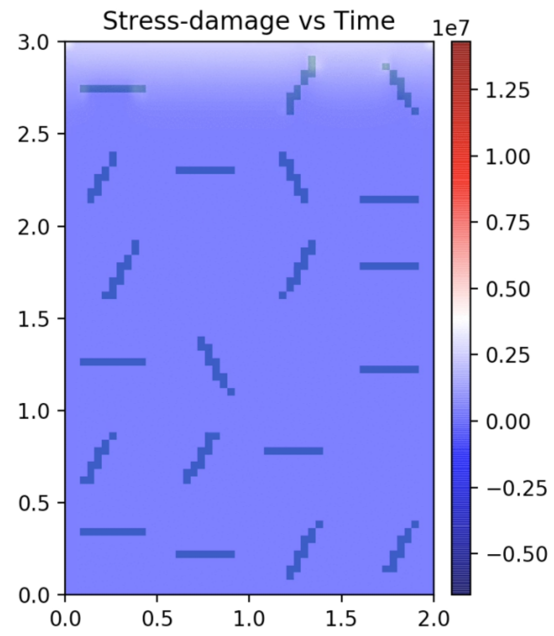
Too much heuristic physics

Emulating Stress from HOSS through ML

Emulated Stress



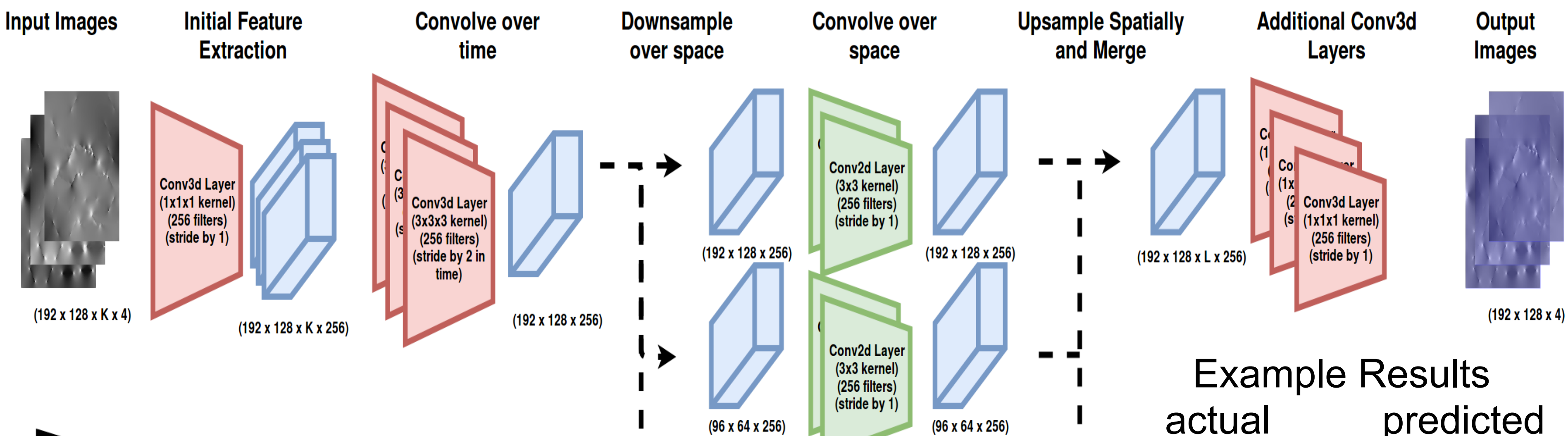
True Stress



Using ML, we predict stress field in a HOSS simulation.

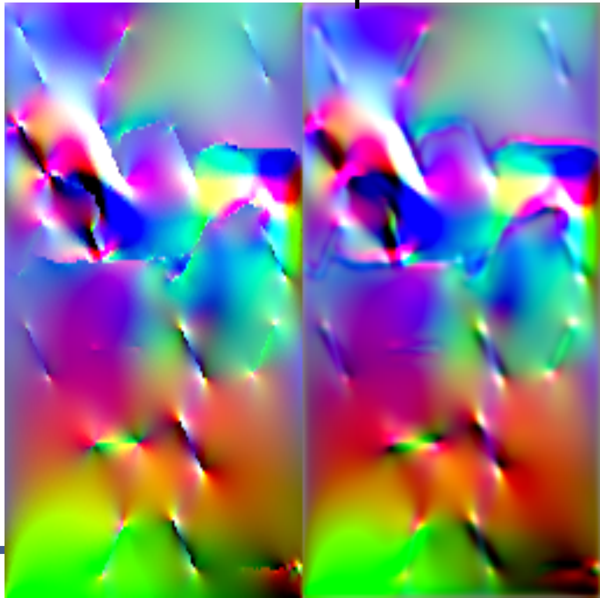
- Stress is predicted using a kernel PCA and "time stacked regression" from training data sets.
- The green lines indicate damage. Currently, work is being done to predict damage as well to get true HOSS emulation.

Prediction of Stress Field with Physics Informed Convolutional Neural Networks



Example Results

actual predicted



- 0.02 Mean-squared-error on held-out validation set
- Multi-scale parallel models combined with a convolutional layer – less grainy AND sharpness in high-stress areas

Summary – Brittle Failure

- Dynamic graphs and ML enable the development of efficient emulators of crack propagation in brittle materials
- The winning algorithm is based on a combination of physics informed and data driven approaches
- QOI like time to failure, failure path and crack statistics can be predicted by our approach
- Since each HOSS simulation takes ~4 hours on 400 processors, ML algorithms that run in seconds to a few minutes offer promising speedup

Conclusions

- QOI in fractured systems are heavily dependent on topological uncertainties, so graph representations are a great way to retain the underlying structure using fewer dof.
- The successful algorithms that emulate detailed high fidelity behavior are based on a combination of physics informed and ML approaches.
- Graph-based ML based emulators can be trained to reproduce key QOI in the fractured systems considered here at significant computational savings.
- UQ framework utilizes 1000’s of runs of the cheap emulators and can bound uncertainties on QOI, which was previously out of reach due to computational burden.

Thank you!

Questions?