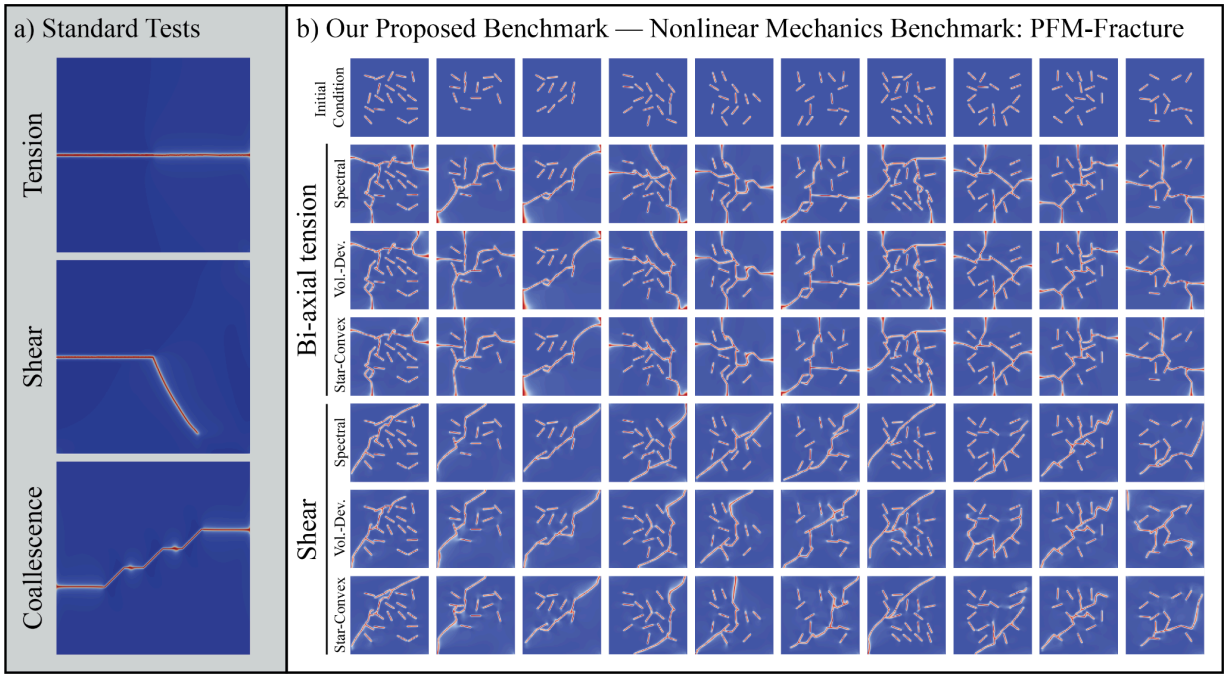
Nonlinear Mechanics Benchmark: PFM-Fracture



Abstract

The Nonlinear Mechanics Benchmarks: PFM-Fracture dataset contains high resolution finite element simulations of quasi-static brittle fracture using the phase field method. Each sample corresponds to the evolution of the crack field and displacement field inside a homogeneous material domain with varying initial conditions, boundary conditions, and energy decomposition methods. The simulations are designed to capture a wide range of crack propagation behaviors, offering a challenging benchmark for machine learning models. Each sample is defined on a 2x2 square domain with a random initial crack configuration. Two types of boundary conditions are considered: Bi-axial tension and pure shear. For each boundary condition, we have performed the simulations with three widely used energy decomposition methods known as Spectral decomposition, Volumetric-Deviatoric decomposition, and Star-Convex decomposition. Each dataset comprises 1,000 simulations for each variant. For each simulation, we store 100 simulation steps capturing the evolution of the phase field and displacement field. All the simulations are reported on the original simulation grid and a downsampled 128x128 version more suitable for machine learning applications. This dataset is intended to serve as a robust benchmark for evaluating surrogate modeling approaches in fracture mechanics. All simulations were conducted using FEniCSx, an open-source finite element software. The full dataset and the code for generating the data and training the baseline models are available on https://dataverse.harvard.edu/dataverse/PFM-Fracture, https://github.com/erfanhamdi/pfm_dataset and https://github.com/erfanhamdi/pfm_bench.

You can find the accompanying manuscript here: https://www.arxiv.org/abs/2507.07237

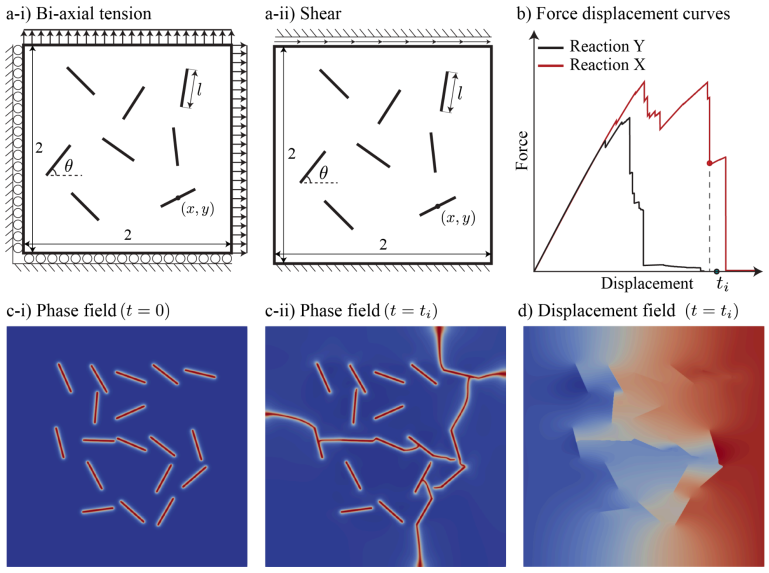
If you use this dataset in your research, please cite it as follows:

```
@misc{hamdi2025robustsurrogatemodelsbenchmarking,
      title={Towards Robust Surrogate Models: Benchmarking Machine Learning
Approaches to Expediting Phase Field Simulations of Brittle Fracture},
      author={Erfan Hamdi and Emma Lejeune},
      year={2025},
      eprint={2507.07237},
      archivePrefix={arXiv},
      primaryClass={cs.LG},
      url={https://arxiv.org/abs/2507.07237},
}
```

Dataset Files:

Each dataset contains 1,000 samples that are stored in `hdf5` format. Each sample is named after the seed used to generate the initial crack pattern. For the 128x128 version of the dataset, the keys required to access the data is as follows:

`2519.hdf5`
|
|__ [`00002519`]
|
|____ [`data`]
|_____ (3, 101, 128, 128)
|
|____ [`grid`]
|_____ [`t` , `x`, `y`]
|
|____ [`force_disp_x`, `force_disp_y`]
|_____ (#sim_steps, 2)
|
|____ [`init`]
|_____(#initial_cracsk, 2, 2)



You can find the file names and links to the datasets in Tables 1 and 2

Table 1: Full resolution version of dataset

Boundary Condition	Energy Decomposition	File Name	DOI Link
Bi-axial tension	Spectral	tension-spect-fullres-part-1	https://doi.org/10.7910/DVN/JCY2AD
		tension-spect-fullres-part-2	https://doi.org/10.7910/DVN/XRKFLH
	Volumetric-Deviatoric	tension-vol-fullres-part-1	https://doi.org/10.7910/DVN/R2EK5L
		tension-vol-fullres-part-2	https://doi.org/10.7910/DVN/R6NUAV
	Star-Convex	tension-star-fullres-part-1	https://doi.org/10.7910/DVN/FHQEG8
		tension-star-fullres-part-2	https://doi.org/10.7910/DVN/YV62WD
Shear	Spectral	shear-spect-fullres-part-1	https://doi.org/10.7910/DVN/D1FOBA
		shear-spect-fullres-part-2	https://doi.org/10.7910/DVN/CLC9IZ
	Volumetric-Deviatoric	shear-vol-fullres-part-1	https://doi.org/10.7910/DVN/M8DAUE
		shear-vol-fullres-part-2	https://doi.org/10.7910/DVN/XQ7NUK
	Star-Convex	shear-star-fullres-part-1	https://doi.org/10.7910/DVN/OEKCNY
		shear-star-fullres-part-2	https://doi.org/10.7910/DVN/ZBNDJO

Table 2: Downsampled to 128x128 version of the dataset

Boundary Condition	Energy Decomposition	File Name	DOI Link
Bi-axial tension	Spectral	tension-spect-128	https://doi.org/10.7910/DVN/YLQGUO
	Volumetric-Deviatoric	tension-vol-128	https://doi.org/10.7910/DVN/G5DLI7

	Star-Convex	tension-star-128	https://doi.org/10.7910/DVN/9URYI1
Shear	Spectral	shear-spect-128	https://doi.org/10.7910/DVN/KZDRUE
	Volumetric-Deviatoric	shear-vol-128	https://doi.org/10.7910/DVN/OCVQJ1
	Star-Convex	shear-star-128	https://doi.org/10.7910/DVN/APUKE5

How to access the data:

This example code shows how to access the phase field, displacement fields, force displacement curves, grid data and the coordinates of the initial cracks of each sample in the dataset:

```
# Import the necessary packages
import h5py
import numpy as np
file_path = "2519.hdf5"
with h5py.File(file_path, "r") as f:
    # The key is the seed used to generate the initial crack
    print(f.keys())
    #>> <KeysViewHDF5 ['00002519']>
    # Now you can access the data using the key
    print(f['00002519'].keys())
    #>> <KeysViewHDF5 ['data', 'force_disp_x', 'force_disp_y', 'grid',
'init']>
    print(f['00002519']['data'].shape)
    #>> (3, 101, 128, 128)
    # The first dimension is the number of fields (damage, displacement x,
displacement y)
    # The second dimension is the number of steps
    # The third and fourth dimensions are the spatial dimensions
```