December Experiments:

New forcefield calculator (SCINE DFB0)

- Cannot use optimization

```
============================================================
SUMMARY: Mean negative eigenvalues by geometry and calculator
============================================================

HIP:
  midpoint_rt         : mean=3.55 (n=100) | distribution: 0:4, 1:5, 2:20, 3:22, 4:17, 5:22, 6:6, 7:2, 8:2
  reactant            : mean=0.00 (n=100) | distribution: 0:100
  product             : mean=0.10 (n=100) | distribution: 0:91, 1:8, 2:1
  midpoint_rt_noise1A : mean=9.55 (n=100) | distribution: 3:1, 4:2, 5:1, 6:4, 7:7, 8:11, 9:22, 10:21, 11:14, 12:8, 13:6, 14:3
  midpoint_rt_noise2A : mean=13.42 (n=100) | distribution: 2:1, 3:1, 5:1, 7:1, 8:3, 9:2, 10:9, 11:11, 12:6, 13:20, 14:6, 15:5, 16:17, 17:6, 18:4, 19:5, 20:2

SCINE:
  midpoint_rt         : mean=4.84 (n=100) | distribution: 1:2, 2:8, 3:17, 4:22, 5:16, 6:17, 7:9, 8:3, 9:2, 10:4
  reactant            : mean=1.80 (n=100) | distribution: 0:5, 1:28, 2:49, 3:18
  product             : mean=2.18 (n=100) | distribution: 0:9, 1:25, 2:36, 3:15, 4:3, 5:9, 6:2, 7:1
  midpoint_rt_noise1A : mean=11.36 (n=100) | distribution: 3:2, 4:2, 5:1, 7:2, 8:4, 9:7, 10:13, 11:19, 12:11, 13:20, 14:10, 15:7, 16:2
  midpoint_rt_noise2A : mean=10.07 (n=100) | distribution: 2:1, 3:1, 4:1, 5:1, 6:2, 7:6, 8:11, 9:9, 10:22, 11:19, 12:14, 13:9, 14:4

============================================================
COMPARISON: HIP vs SCINE agreement
============================================================
  midpoint_rt         : 23/100 agree (23.0%)
  reactant            : 5/100 agree (5.0%)
  product             : 10/100 agree (10.0%)
  midpoint_rt_noise1A : 11/100 agree (11.0%)
  midpoint_rt_noise2A : 10/100 agree (10.0%)

Results saved to: /home/memoozd/scratch/ts-tools-output/eigenvalue_classification/eigenvalue_classification_results.json
```

Table 1: Summary of Mean Negative Eigenvalues by Geometry and Calculator ($n = 100$)

Geometry	HIP Mean	SCINE Mean	Agreement
midpoint_rt	3.55	4.84	23 (23.0%)
reactant	0.00	1.80	5 (5.0%)
product	0.10	2.18	10 (10.0%)
midpoint_rt_noise1A	9.55	11.36	11 (11.0%)
midpoint_rt_noise2A	13.42	10.07	10 (10.0%)

Experiments:

Pure GAD:

- 100% Convergence on non-minima samples
- Doesn't converge on minima. Really easy fix, just kick → 100% convergence on minima.
- https://wandb.ai/memo-ozdincer-university-of-toronto/gad-ts-search-scine/runs/90j51toq
-

```
[Overall Statistics]
  λ₀: -0.074139 ± 0.079542 eV/Å²
  λ₁: 0.044880 ± 0.062035 eV/Å²
  λ₀·λ₁: -5.443389e-03 ± 7.901079e-03
  Steps taken: 18.2 ± 28.8
  Steps to TS: 18.2 ± 28.8
  Final time: 0.132 ± 0.182
```

-
```
[Transition Distribution]
  0neg-to-1neg: 5 samples
  1neg-to-1neg: 11 samples
  2neg-to-1neg: 14 samples
```

HIP vs. SCINE:

- For reacant geometries: 95% disagreeance (SCINE only sees 5% of reactants HIP sees)
  - Cannot escape reactants (5%)
  - This is a super easy fix of just applying a climbing algorithm when a minima is detected…

L-BFGS Energy Minimizer

Since our algorithms converge 100% of the time from minima, why don't we just get them to minima, then we can run our scripts?

SCINE Results:

https://wandb.ai/adversarial-ozd/experiment-lbfgs-minimize/runs/wp1b3pq7 (15k)
https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-lbfgs-minimize/runs/gbq5nu0l

- 10/100 Convergence (transition state)
- Pretty good reduction of order
- 2.5s-10s per sample
- Really good reduction if you start above 10 eigenvalues
- Many plateau at ~8 negative eigenvalues

HIP Results:

https://wandb.ai/adversarial-ozd/experiment-lbfgs-minimize/runs/vl6irj7g (15k)
https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-lbfgs-minimize/runs/ouhjusq6

- 0/100 Convergence
- Number of negeigs plateau, even though displacement may not. Changes in energy become very negligible
- Really good reduction of order
- 0.8s-4.5s range per sample

Trying to get pure GAD to work on noisy geometries

- More elegant
- Should work in theory

Experiment 1 - Kicking (plateau detection):

- GAD algorithm
- If, for 20 steps, the number of negative eigenvalues doesn't change, increase step size.
- Have a stepsize floor (500μÅ)
- Code snippet (or pseudocode)

Typical failure graphs:

- Start at floor displacement, good movement until around 600 steps, then starts oscillating (~7 neg)
- A few graphs with ranges around 12neg - 4 neg: Explodes at around 1300 steps, then can't convergence.

SCINE Results:

Results: 40% convergence

https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-gad-plateau/runs/z51k2hmb

```
[Overall Statistics]
  λ₀: -3.494852 ± 3.185547 eV/Å²
  λ₁: -0.208413 ± 0.377295 eV/Å²
  λ₀·λ₁: 1.025070e+00 ± 2.484256e+00
  Steps taken: 1134.5 ± 471.4
  Steps to TS: 586.2 ± 233.6
  Final time: 9.922 ± 4.290
```

```
[Transition Distribution]
  11neg-to-1neg: 2 samples
  11neg-to-4neg: 2 samples
  11neg-to-8neg: 1 samples
  12neg-to-1neg: 1 samples
  7neg-to-6neg: 1 samples
  8neg-to-1neg: 1 samples
  9neg-to-3neg: 1 samples
  9neg-to-6neg: 1 samples
```

HIP Results:

40% convergence again
https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-gad-plateau/runs/95k5s7fj

```
[Overall Statistics]
  λ₀: -10.851352 ± 4.490373 eV/Å²
  λ₁: -0.425083 ± 1.005568 eV/Å²
  λ₀·λ₁: 5.867434e+00 ± 1.442755e+01
  Steps taken: 1092.0 ± 516.8
  Steps to TS: 480.0 ± 208.4
  Final time: 75.695 ± 35.906
```

```
[Transition Distribution]
  10neg-to-1neg: 2 samples
  10neg-to-4neg: 1 samples
  12neg-to-4neg: 1 samples
  12neg-to-7neg: 1 samples
  5neg-to-3neg: 1 samples
  8neg-to-5neg: 1 samples
  9neg-to-1neg: 2 samples
  9neg-to-4neg: 1 samples
```

Higher-order GAD:

Clearly, kicking in the GAD direction wasn't working too well. So, I changed the algorithm. I knew that the "plateaus" we were seeing didn't make sense, because the geometry was still unstable. I found that the eigenvalue we were minimizing was converging (or rather, it's force norm was becoming extremely small).

So, when I detect that the first order eigenvalue has converged, all I need to do is move to the next, and start moving in that direction. This led to "Higher-Order GAD"

Code Snippet

SCINE Results:

- Immediately achieved 40%, just like the kicking algorithm (on the same samples too). However, this time, the failure graphs looked very very different, they weren't plateauing, they were oscillating.
    a. Before: they were just moving in wrong directions because it was a naive algorithm
    b. Now, the 2-3 smallest eigenvalues are oscillating with eachother. So, my graph tracking just 1 looked like it was oscillating, and they never converged.

```
[Overall Statistics]
  λ₀: -5.598661 ± 9.811996 eV/Å²
  λ₁: -0.252119 ± 0.518562 eV/Å²
  λ₀·λ₁: 4.216977e+00 ± 1.448229e+01
  Steps taken: 365.7 ± 82.3
  Steps to TS: 365.5 ± 78.7
  Final time: 2.838 ± 0.546
```

```
[Transition Distribution]
  10neg-to-1neg: 1 samples
  11neg-to-1neg: 1 samples
  11neg-to-3neg: 2 samples
  11neg-to-7neg: 2 samples
  12neg-to-1neg: 1 samples
  4neg-to-1neg: 1 samples
  4neg-to-3neg: 1 samples
  4neg-to-4neg: 1 samples
  5neg-to-1neg: 1 samples
  7neg-to-6neg: 1 samples
  8neg-to-1neg: 1 samples
  9neg-to-3neg: 1 samples
  9neg-to-5neg: 1 samples
```

https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-multi-mode/runs/rslhdnmx

HIP Results:

https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-multi-mode/runs/g5omatz9

- Same comments

```
[Overall Statistics]
  λ₀: -4.017393 ± 4.179038 eV/Å²
  λ₁: -0.712270 ± 1.204521 eV/Å²
  λ₀·λ₁: 4.577574e+00 ± 8.091011e+00
  Steps taken: 344.1 ± 175.6
  Steps to TS: 202.5 ± 173.7
  Final time: 24.187 ± 12.299
```

```
[Transition Distribution]
  10neg-to-1neg: 1 samples
  10neg-to-3neg: 1 samples
  10neg-to-6neg: 1 samples
  11neg-to-4neg: 1 samples
  12neg-to-10neg: 1 samples
  12neg-to-4neg: 1 samples
  1neg-to-1neg: 1 samples
  4neg-to-1neg: 2 samples
  5neg-to-1neg: 1 samples
  5neg-to-7neg: 1 samples
  8neg-to-6neg: 1 samples
  9neg-to-1neg: 1 samples
  9neg-to-2neg: 1 samples
  9neg-to-9neg: 1 samples
```

HIP ERROR:
Error edge_vec_0_distance: 1.0767344065243378e-05
Error edge_vec_0_distance: 5.7138218835461885e-05
Error edge_vec_0_distance: 5.7138218835461885e-05
Error edge_vec_0_distance: 2.1210158593021333e-05

Error edge_vec_0_distance: 2.1210158593021333e-05
Error edge_vec_0_distance: 2.8712902349070646e-06
Error edge_vec_0_distance: 2.8712902349070646e-06

Mode Tracking + Trust Radiuses + Improvements on mmGAD:

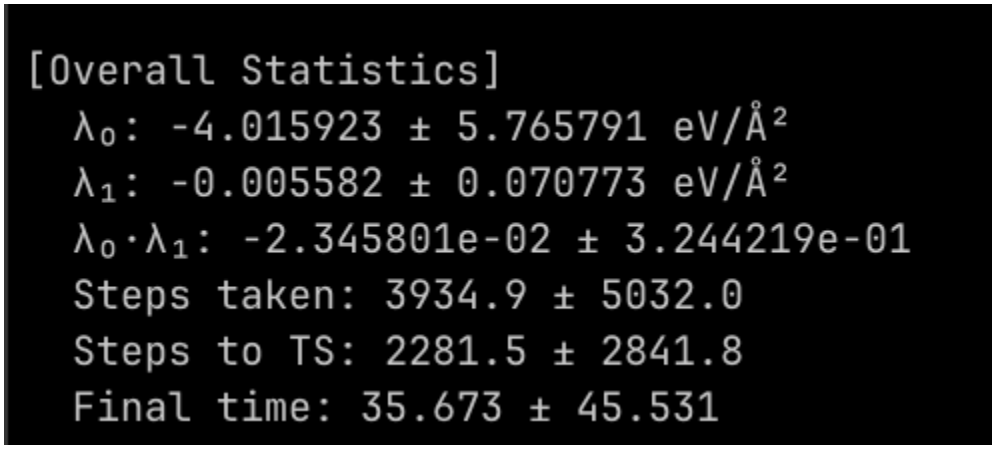
Code Snippet:
Explanation

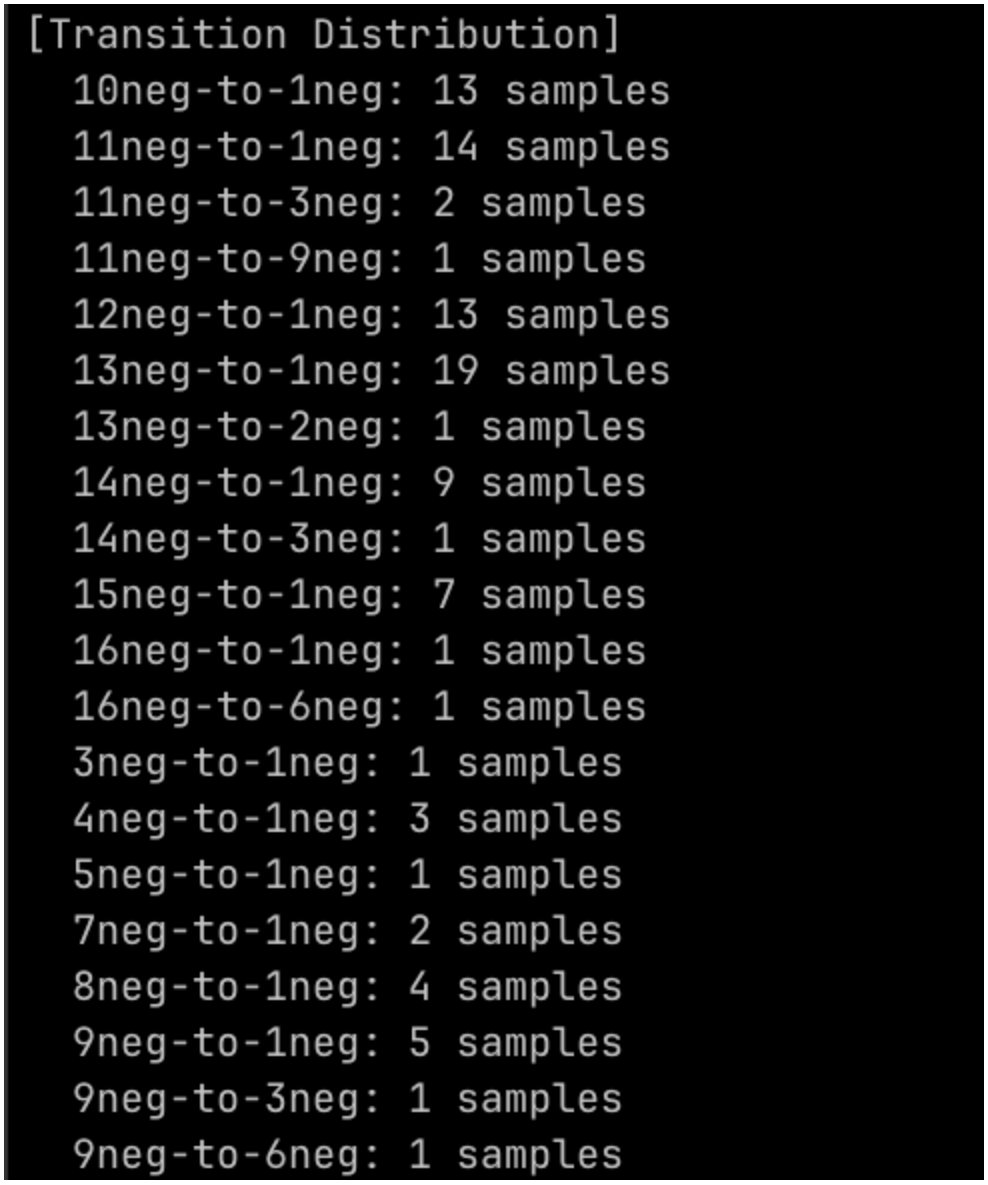
1. Mode tracking removed the oscillating non-convergences
2. Trust radiuses stopped us from exploding, and allowed me to implement proper kicking, to have some additional force in the secondary eigenvector direction

The combination of these improvements made the mmGAD algorithm finally work to its full capacity:

SCINE Results (92%):

https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-multi-mode-eckartmw/runs/phbzn2o0





Some discussion of failure graphs:

HIP Results (71%):

https://wandb.ai/memo-ozdincer-university-of-toronto/experiment-multi-mode-eckartmw/runs/awd9sj4k

```
[Overall Statistics]
  λ₀: -5.223560 ± 5.432469 eV/Å²
  λ₁: -0.296980 ± 0.835196 eV/Å²
  λ₀·λ₁: 2.025604e+00 ± 7.727223e+00
  Steps taken: 3858.8 ± 5636.1
  Steps to TS: 1632.0 ± 2450.6
  Final time: 305.750 ± 390.018
```

```
[Transition Distribution]
  10neg-to--1neg: 3 samples
  10neg-to-10neg: 2 samples
  10neg-to-1neg: 17 samples
  10neg-to-2neg: 1 samples
  10neg-to-4neg: 1 samples
  10neg-to-5neg: 2 samples
  10neg-to-7neg: 1 samples
  11neg-to--1neg: 1 samples
  11neg-to-1neg: 12 samples
  11neg-to-2neg: 1 samples
  11neg-to-4neg: 1 samples
  11neg-to-6neg: 2 samples
  11neg-to-7neg: 1 samples
  12neg-to--1neg: 2 samples
  12neg-to-1neg: 12 samples
  12neg-to-3neg: 1 samples
  12neg-to-4neg: 1 samples
  12neg-to-5neg: 1 samples
  12neg-to-6neg: 1 samples
  13neg-to--1neg: 2 samples
  13neg-to-1neg: 2 samples
  13neg-to-6neg: 1 samples
  14neg-to-1neg: 1 samples
  15neg-to--1neg: 1 samples
  15neg-to-4neg: 1 samples
  17neg-to--1neg: 1 samples
  1neg-to-1neg: 1 samples
  4neg-to-1neg: 2 samples
  5neg-to-1neg: 2 samples
  6neg-to-1neg: 2 samples
  7neg-to--1neg: 1 samples
  7neg-to-1neg: 2 samples
  7neg-to-2neg: 1 samples
  7neg-to-5neg: 1 samples
  8neg-to-1neg: 5 samples
  8neg-to-6neg: 1 samples
  9neg-to-1neg: 10 samples
```

Some discussion of failure graphs: