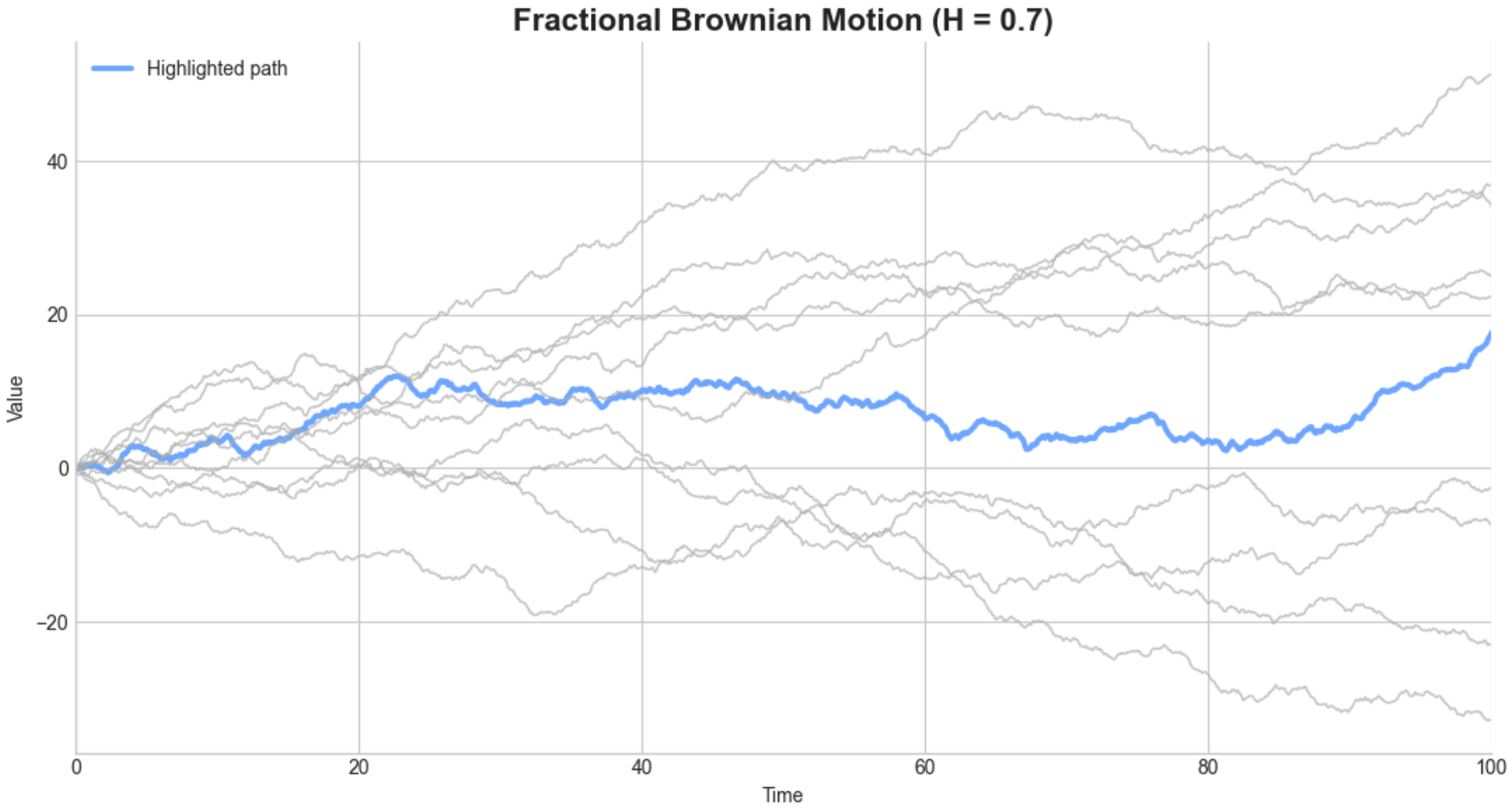
Algorithmic Trading with Reinforcement Learning

First semester report

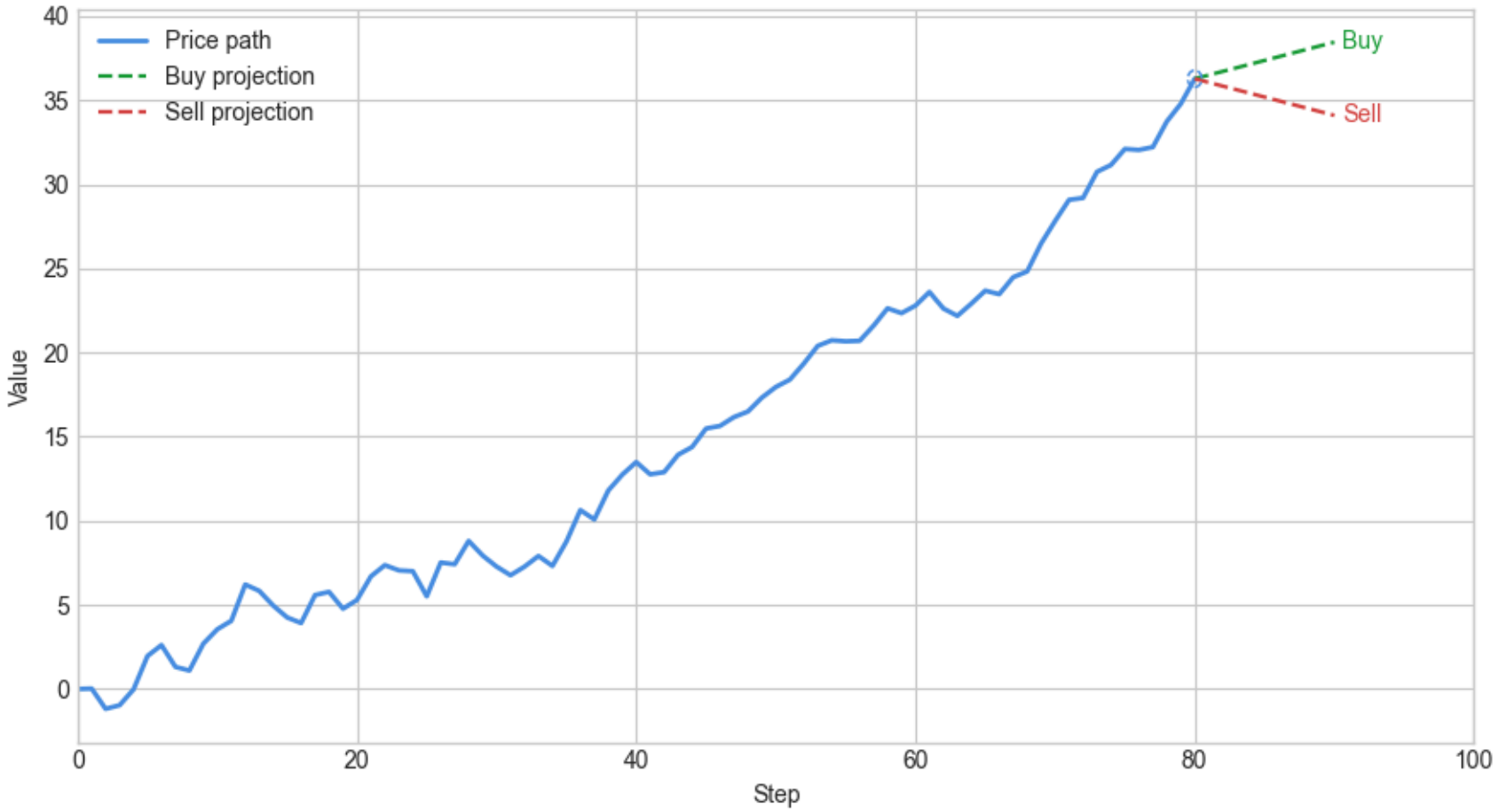
Leonardo Toffalini

2026-01-08

Problem statement



Problem statement



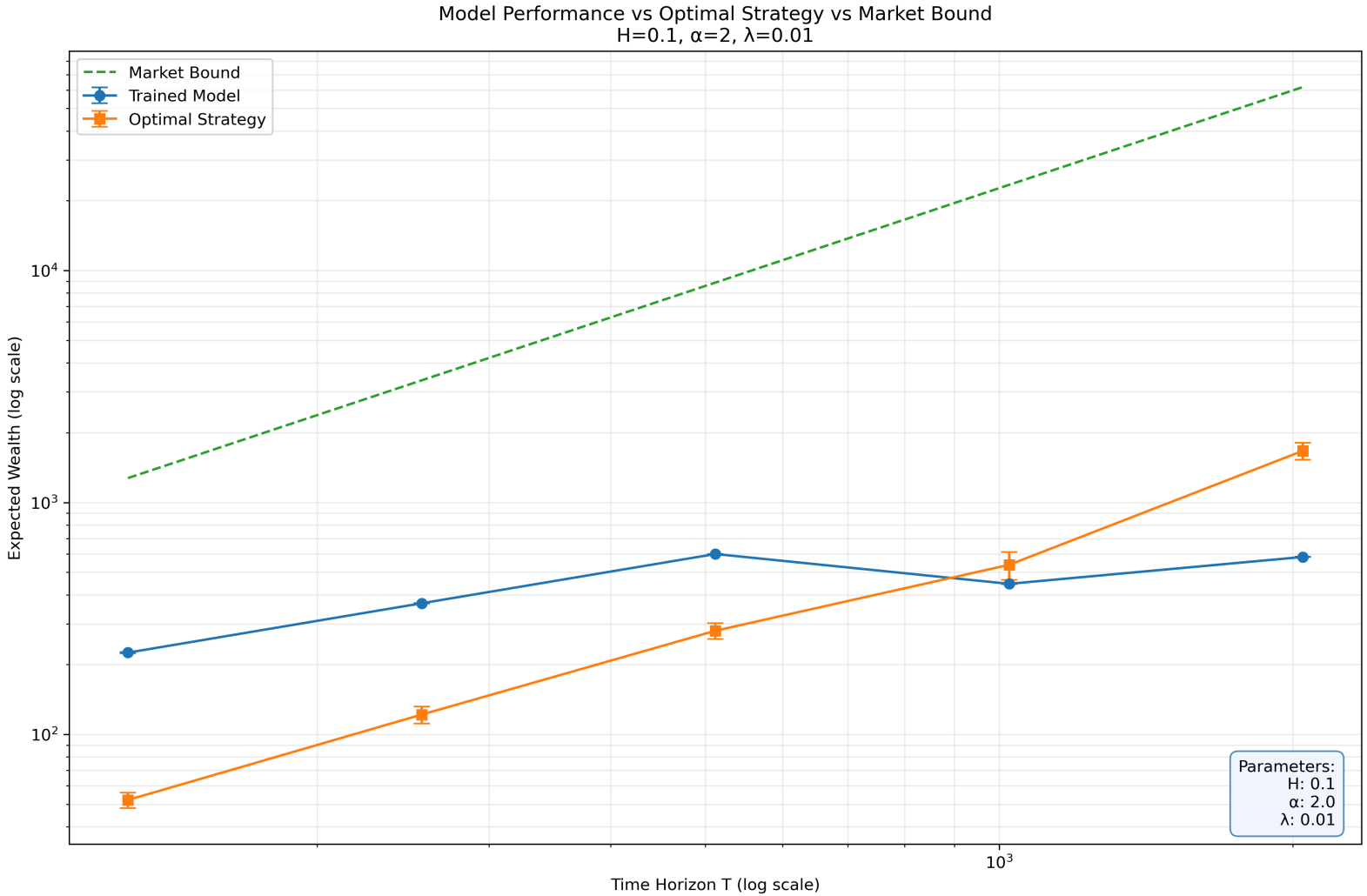
Problem statement

$$X_t^1(\phi) := z^1 + \int_0^t \phi_u \, \mathrm{d}u \qquad \text{(risky)}$$

$$X_t^0(\phi) := z^0 - \int_0^t \phi_u S_u \, \mathrm{d}u - \int_0^t \lambda |\phi_u|^\alpha \, \mathrm{d}u \qquad \text{(riskless)}$$

$$\max_{\phi \in S(t)} \mathbb{E}\left[X_T^0(\phi)\right]$$

Previous work



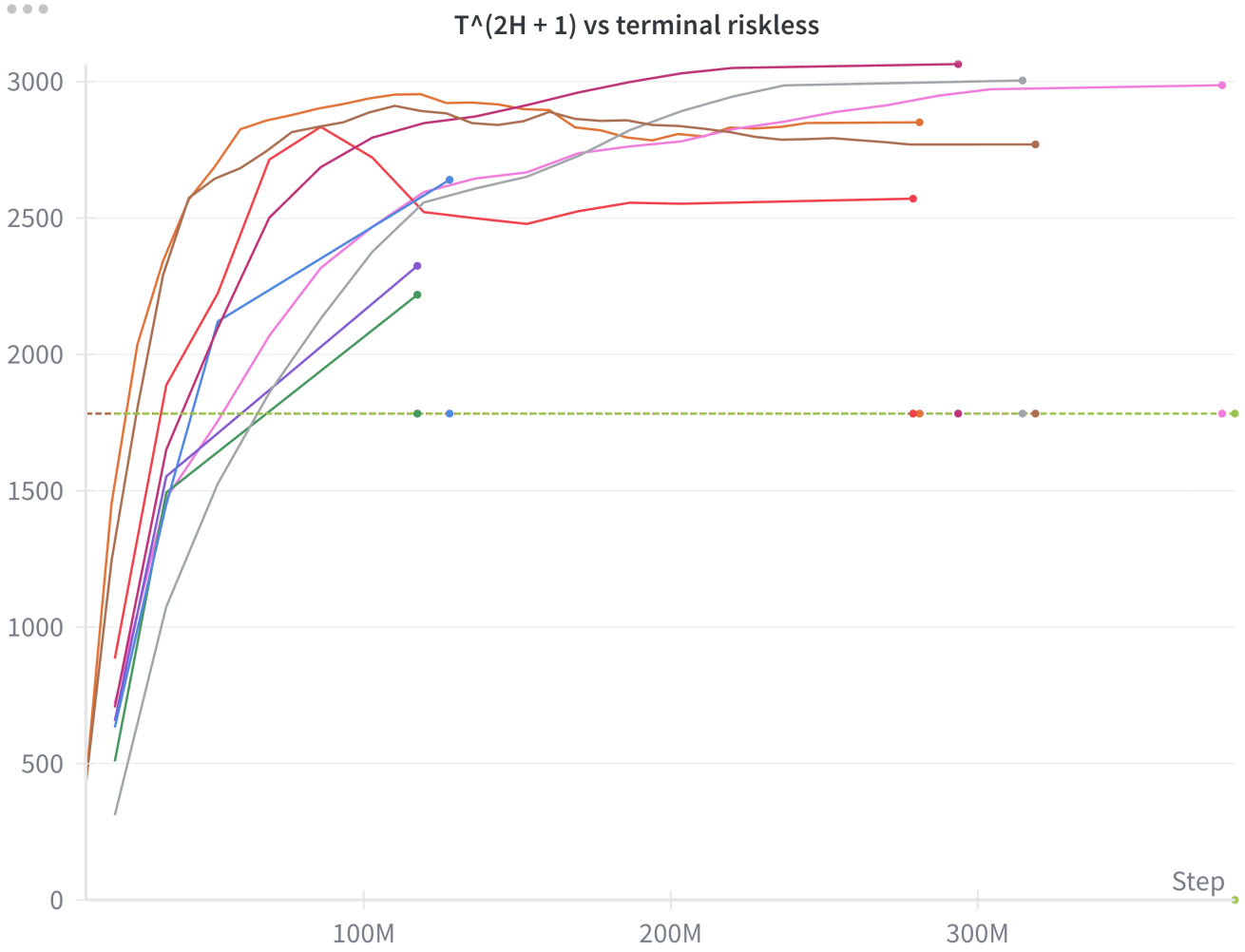
Current work

- The code base was rewritten in C from Python
- Liquidation strategy is now a linear schedule
- Smarter rewards: anticipating liquidation value

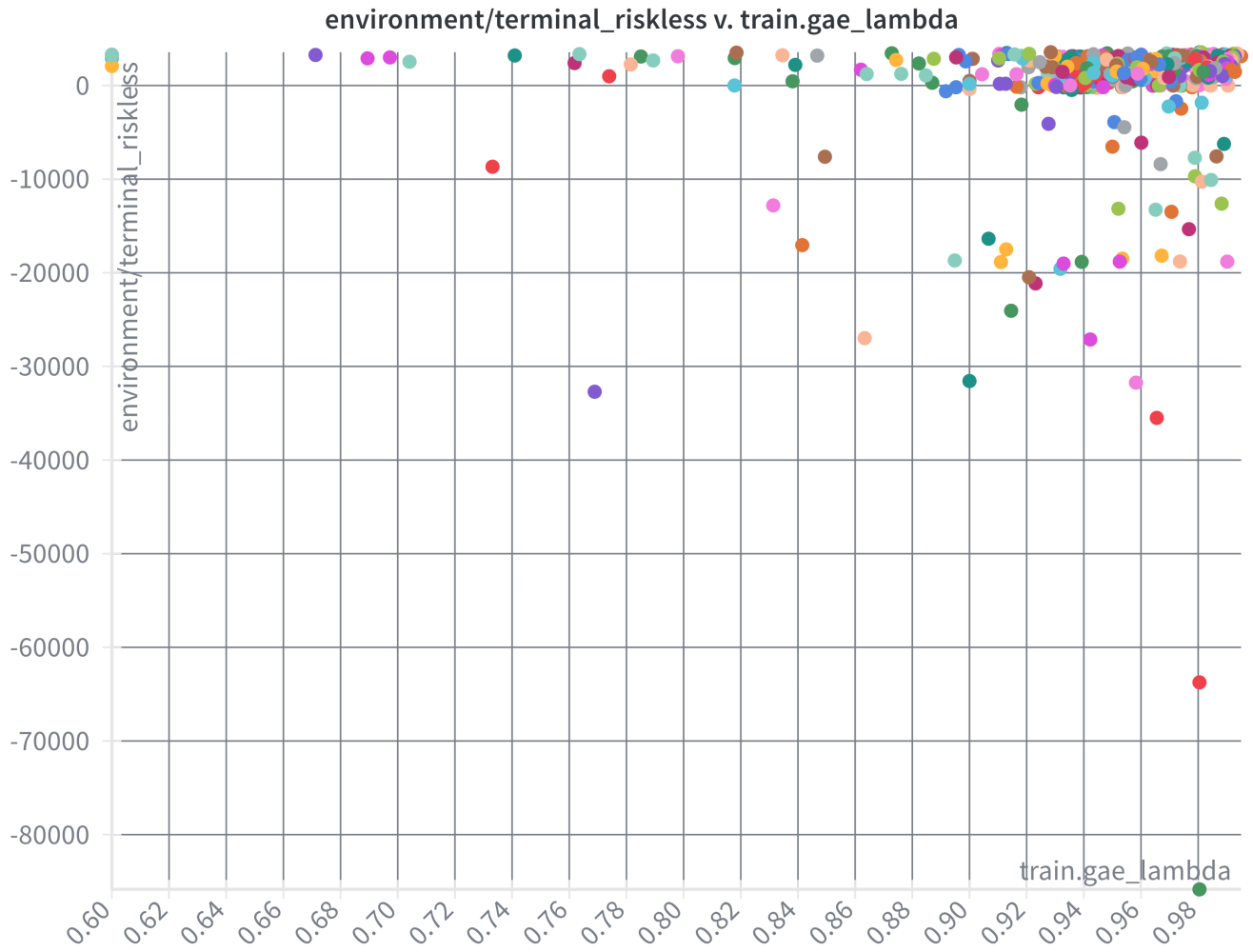
Results

- Three orders of magnitude speedup (1.5k $\Rightarrow$ 1.5M SPS)
- Large-scale hyperparameter search with CARBS
- So far better performance for fix time horizons

Example hyperparameter sweep (T = 512)



Example hyperparameter sweep (T = 512)



Usage of AI tools

- Perplexity – Research
- Cursor – Programming