Topic

Learning to Switch: A DRL Meta-Controller over HRP/HERC for Multi-Market Portfolio Allocation

Abstract (About 300 Words)

Markets flip regimes; correlations and volatilities slide around. Classical mean–variance tools lean on fragile covariance inversion and can swing weights unrealistically. This proposal takes a practical middle path: keep the fast, robust structure of *hierarchical* allocation, and add a careful layer of reinforcement learning that decides *which* allocator to trust—and when to simply hold.

At the **base layer**, I use Hierarchical Risk Parity (HRP) and Hierarchical Equal Risk Contribution (HERC) to build diversified sleeves: US/EU/EM equities; Treasury duration buckets (2–5y, 7–10y, 20y+); investment-grade and high-yield credit; gold; and broad commodities. These methods cluster assets by correlation and allocate top–down, spreading risk, curbing concentration, and avoiding noisy matrix inversion.

At the **meta layer**, a discrete-action PPO agent chooses the HRP/HERC variant (lookback, risk metric) that governs the next period, with an explicit *Hold* action for broad selloffs. The state is intentionally lean: recent performance of each base allocator plus a few regime cues (realized volatility, a simple trend filter, a cross-asset stress flag). The reward is net P&L with explicit transaction-cost and turnover penalties and a gentle drawdown term to avoid whipsaw. Guardrails include per-sleeve caps, a turnover budget, and optional cash as a sleeve.

The main question: can a frictions-aware switcher beat static hierarchical baselines and a 60/40 reference on risk-adjusted metrics while keeping turnover truly tradable? I will evaluate daily and hourly settings with walk-forward splits and a reversed-chronology stress. Ablations compare (i) the published discrete switcher, (ii) a regime-augmented switcher, and (iii) a soft mixture that blends multiple HRP/HERC outputs under a simplex constraint. Success means higher Sharpe/Sortino, lower CVaR@5%, stable turnover, and transparent behavior—*which base model the agent chooses and when*.

Relevant Papers (MLA Citation Format)

- López de Prado, Marcos. "Building Diversified Portfolios that Outperform Out of Sample." *The Journal of Portfolio Management*, vol. 42, no. 4, 2016, pp. 59–69. doi:10.3905/jpm.2016.42.4.059
- Raffinot, Thomas. "The Hierarchical Equal Risk Contribution Portfolio." *SSRN Electronic Journal*, 2018. doi:10.2139/ssrn.3237540
- Millea, Adrian, and Abbas Edalat. "Using Deep Reinforcement Learning with Hierarchical Risk Parity for Portfolio Optimization." *International Journal of Financial Studies*, vol. 11, no. 1, 2023, Article 10. doi:10.3390/ijfs11010010

Student(s) Information

Name: Payam Taebi
Student ID: 400104867
Email: xpayamtaebix@gmail.com