

# Differentiable Hierarchical Risk Parity: Learning to Allocate with End-to-End Gradient Optimization

Jose L. Amador

## Abstract

This paper introduces Differentiable Hierarchical Risk Parity (DHRP), a neural network layer that rewrites the classical hierarchical risk parity algorithm as a fully differentiable map from market features and covariance estimates to portfolio weights. Classical HRP relies on a sequence of discrete clustering and recursive bisection steps, so its parameters cannot be tuned together with other model components using gradients. We replace those rigid steps with soft gating networks and learnable tree structures, which makes the entire allocation rule trainable end-to-end with stochastic gradient descent. The DHRP layer is trained with a multi-objective loss that combines constant relative risk aversion utility, a smooth Sharpe ratio term, and a regularizer that keeps the solution close to standard HRP. In out-of-sample tests on developed and emerging market ETF universes covering ten years of daily data, DHRP reaches Sharpe ratios of 1.17 and 0.11, while classical HRP delivers 1.11 and 0.06 on the same datasets. The Sharpe ratios are highly significant when tested against zero, as expected given the long sample, and DHRP delivers Fama-French alphas of 6.2% ( $t = 5.7$ ) and -5.0% ( $t = -0.9$ ) for developed and emerging markets respectively. Overall, the results show that it is possible to keep the diversification discipline of HRP while using modern deep learning tools to obtain more stable and data-driven allocations.

## 1 Introduction

Portfolio construction remains a core problem in quantitative finance. The mean variance setup of Markowitz [1952] gives a clean theoretical starting point, but when implemented on real data it is very sensitive to small changes in the inputs. Slight shifts in estimated expected returns can lead to completely different portfolios, so allocations jump around and out-of-sample performance tends to be disappointing [DeMiguel et al., 2009].

Hierarchical Risk Parity (HRP), proposed by Lopez de Prado [2016], was designed to sidestep part of this issue. Instead of requiring expected returns, HRP works only with covariances and correlations. It clusters assets based on their co-movement and then allocates risk down the resulting tree, which usually leads to allocations that are well diversified and less exposed to covariance noise. In practice HRP has been adopted by both practitioners and researchers, and has been extended in several directions, for example to include transaction costs, regime changes, and multi-period decisions.

From a machine learning point of view, however, classical HRP has a clear drawback. The clustering step makes hard assignments, and the recursive bisection step makes hard left/right decisions. These operations are non-differentiable, so one cannot simply plug HRP into a neural network and update its parameters with

gradient-based methods. As a result, HRP is usually treated as a fixed post-processing step that sits outside any learned forecasting model.

The starting point of this work is to ask whether one can keep the hierarchical structure of HRP while removing the non-differentiable parts. To that end, we introduce Differentiable Hierarchical Risk Parity (DHRP), a neural network module that behaves like a smoothed and learnable version of HRP. The discrete tree and hard splits are replaced by soft routing gates, but the overall idea of pushing risk budgets down a hierarchy is preserved. The DHRP layer takes feature vectors and covariance matrices as inputs and outputs portfolio weights through a single computation graph that can be trained end-to-end.

We train this layer with a loss that mirrors the objectives of a risk-averse investor. The main term is constant relative risk aversion (CRRA) utility, which trades off average return against downside risk. On top of this, we add a smooth Sharpe ratio term that encourages stable risk-adjusted performance, and a penalty that keeps the learned weights close to those of standard HRP, at least in the early part of training. This last term acts as a regularizer and as a way to inject prior knowledge about diversification.

The empirical analysis uses ETF universes for developed and emerging markets over the period 2015–2025. We run rolling-window backtests with realistic rebalancing and compare DHRP to equal weight (EW), minimum variance (MINVAR), classical mean variance (MV), and traditional HRP. In developed markets, DHRP reaches an annualized Sharpe ratio of 1.17, compared to 1.11 for HRP and 0.76 for EW. In emerging markets, which are more volatile and more correlated, DHRP reaches a Sharpe ratio of 0.11 versus 0.06 for HRP, so the relative improvement is larger in that universe.

We also check that these differences are not driven by a small set of favorable dates. Bootstrap confidence intervals show that DHRP maintains its advantage when the sample is perturbed, and the Fama-French alpha of 6.2% per year ( $t = 5.7$ ,  $p < 0.001$ ) confirms that the gains are not simply due to higher factor exposures.

**Contributions.** The contributions of this paper are mainly three. First, we give a differentiable version of HRP that can be trained with gradients while keeping the intuition of hierarchical diversification. Second, we describe a multi-objective training scheme that balances utility, risk-adjusted performance, and proximity to classical HRP. Third, we show on real ETF data that this approach produces clear out-of-sample gains over widely used allocation rules in both developed and emerging markets.

## 2 Related Work

**Differentiable optimization layers.** The idea of embedding optimization as a differentiable layer in neural networks has attracted significant attention in recent years. Agrawal et al. [2019] introduced differentiable convex optimization layers that allow constrained problems to be solved within deep models. More broadly, work on implicit differentiation [Bolte et al., 2021] has established theoretical foundations for backpropagating through nonsmooth optimization procedures. Jin et al. [2021] proposed safe Pontryagin differentiable programming for safety-critical learning and control tasks, which shares conceptual similarities with our approach of making allocation rules gradient-friendly while respecting constraints. The DHRP layer fits into

this trend but encodes the specific inductive bias of hierarchical diversification.

**Hierarchical and tree-structured learning.** Tree-based neural architectures have seen renewed interest for problems with natural hierarchical structure. Soft routing mechanisms, similar to those we employ, appear in sparsely-gated mixture-of-experts models [Shazeer et al., 2017], which use trainable gating networks to route inputs to different expert subnetworks. Lopez de Prado [2016] introduced hierarchical risk parity for portfolio allocation, using agglomerative clustering and recursive bisection to spread risk across a tree. Our contribution is to make this pipeline differentiable so it can be trained end-to-end rather than treated as a fixed post-processor.

**End-to-end learning for sequential decisions.** Training systems directly on downstream objectives rather than surrogate losses has become standard practice. Decision Transformer [Chen et al., 2021] frames reinforcement learning as sequence modeling, while Donti et al. [2017] showed how to integrate prediction and optimization in a single differentiable pipeline for stochastic programming. In portfolio settings, Zhang et al. [2020] directly optimized the Sharpe ratio through deep learning, bypassing the need to forecast expected returns and demonstrating improved out-of-sample performance. The DHRP layer is designed to support this paradigm by supplying a differentiable allocation rule.

## 3 Method

We now spell out the technical details of the DHRP layer. We first recall how standard HRP works and then explain how we modify each non-differentiable step.

### 3.1 Classical Hierarchical Risk Parity

Classical HRP operates in three stages. Given an estimate of the asset covariance matrix  $\Sigma \in \mathbb{R}^{n \times n}$ , the algorithm first computes a distance matrix from correlations:

$$D_{ij} = \sqrt{\frac{1}{2}(1 - \rho_{ij})}, \quad (1)$$

where  $\rho_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$  is the correlation between assets  $i$  and  $j$ . Second, an agglomerative clustering algorithm (typically Ward’s method or single linkage) is applied to  $D$  to construct a hierarchical tree over assets. Third, portfolio weights are determined by recursively bisecting the tree and allocating risk between left and right branches in inverse proportion to their volatilities.

The main idea behind HRP is that respecting the hierarchical structure implied by correlations leads to portfolios that are naturally diversified and do not depend on expected return estimates. The price for this robustness is that clustering and bisection are discrete operations, so the algorithm is non-differentiable.

### 3.2 Differentiable HRP Layer

Our DHRP layer replaces each non-differentiable operation with a smooth, parameterized alternative but keeps the overall HRP structure. Concretely, the layer is implemented as a PyTorch module that maps a feature vector  $x_t \in \mathbb{R}^d$  and covariance matrix  $\Sigma_t \in \mathbb{R}^{n \times n}$  to portfolio weights  $w_t \in \Delta^{n-1}$ , where  $\Delta^{n-1}$  is the probability simplex.

**Architecture.** We maintain a fixed binary tree of depth  $L$ , yielding  $2^L$  leaf nodes. Each internal node  $k$  is associated with a gating network  $g_k : \mathbb{R}^d \rightarrow \mathbb{R}^2$  that computes soft routing probabilities:

$$\pi_k = \text{softmax}(g_k(h_t)/\tau), \quad (2)$$

where  $h_t$  is a node-specific feature representation derived from  $x_t$  and  $\Sigma_t$ , and  $\tau > 0$  is a temperature parameter controlling the sharpness of the routing. Lower temperatures produce more decisive splits that approach the hard assignments of classical HRP, while higher temperatures yield smoother interpolations.

The feature representation  $h_t$  is constructed by combining the input features with covariance information through a projection network:

$$h_t = \text{LayerNorm}(x_t + f_{\text{cov}}(\text{vec}(\Sigma_t/\|\Sigma_t\|_\infty))), \quad (3)$$

where  $f_{\text{cov}}$  is a small multilayer perceptron and the covariance matrix is normalized for numerical stability.

**Weight computation.** The effective budget for each leaf node is computed by multiplying the routing probabilities along the path from the root. For a leaf  $\ell$  with path  $P_\ell = \{(k_1, d_1), \dots, (k_L, d_L)\}$ , where  $d_i \in \{0, 1\}$  indicates the left or right branch, the path probability is:

$$p_\ell = \prod_{(k,d) \in P_\ell} \pi_k[d]. \quad (4)$$

Each leaf maintains learnable logits that are combined with the path probabilities to determine preliminary asset budgets. These budgets are then blended with an inverse volatility weighting scheme:

$$w_i = \alpha \cdot b_i + (1 - \alpha) \cdot \frac{1/\sigma_i}{\sum_j 1/\sigma_j}, \quad (5)$$

where  $b_i$  is the tree-derived budget for asset  $i$ ,  $\sigma_i = \sqrt{\Sigma_{ii}}$  is the asset volatility, and  $\alpha = \text{sigmoid}(\theta_\alpha)$  is a learnable blending coefficient. This formulation ensures that the final weights are always valid (positive and sum to one) while allowing the model to interpolate between the learned hierarchy and classical inverse volatility allocation.

### 3.3 Feature Engineering

The input feature vector  $x_t$  summarizes relevant market information from the trailing estimation window. We include:

- Annualized mean returns for each asset, clipped to  $[-2, 2]$  for stability
- Annualized volatilities, clipped to  $[0.01, 2]$
- Short-term momentum signals computed as the difference between recent and prior period returns
- Volatility regime indicators measuring the ratio of recent to long-term volatility
- The average pairwise correlation across assets

This feature set mixes slowly changing characteristics (such as volatility levels and correlation structure) with more dynamic signals (such as short-term momentum and regime indicators). For the universes we study, this yields a feature dimension of  $d = 48$ .

### 3.4 Training Objective

We train the DHRP layer by minimizing a multi-objective loss over rolling windows of historical data. Let  $\mathcal{B} = \{(x_t, \Sigma_t, r_{t+1})\}$  denote a mini-batch of feature, covariance, and return tuples. For sample  $t$  in the batch, the portfolio return is  $r_t^p = w_t^\top r_{t+1}$ , with  $w_t = \text{DHRP}(x_t, \Sigma_t)$ .

**CRRA utility.** The primary objective captures investor preferences through constant relative risk aversion utility:

$$\mathcal{L}_{\text{CRRA}} = -\frac{1}{|\mathcal{B}|} \sum_{t \in \mathcal{B}} \frac{(1 + r_t^p)^{1-\gamma} - 1}{1 - \gamma}, \quad (6)$$

where  $\gamma > 0$  is the risk aversion coefficient. We use  $\gamma = 2.5$  for developed markets and  $\gamma = 1.2$  for emerging markets, reflecting the higher volatility environment in the latter.

**Sharpe ratio objective.** We include a smooth Sharpe ratio term to explicitly target risk-adjusted performance:

$$\mathcal{L}_{\text{Sharpe}} = -\lambda_S \cdot \frac{\mu_r}{\sigma_r + \epsilon}, \quad (7)$$

where  $\mu_r$  and  $\sigma_r$  are the batch mean and standard deviation of portfolio returns, and  $\epsilon = 10^{-6}$  ensures numerical stability. The coefficient  $\lambda_S$  balances this objective against CRRA utility.

**HRP regularization.** To leverage the proven benefits of classical HRP and provide a strong initialization, we regularize the learned weights toward the HRP solution:

$$\mathcal{L}_{\text{HRP}} = \lambda_H \cdot \frac{1}{|\mathcal{B}|} \sum_{t \in \mathcal{B}} \|w_t - w_t^{\text{HRP}}\|_2^2, \quad (8)$$

where  $w_t^{\text{HRP}}$  is the classical HRP allocation computed from  $\Sigma_t$ . The coefficient  $\lambda_H$  follows a curriculum schedule that starts high and decreases during training, allowing the model to initially stay close to HRP before gradually deviating to optimize the primary objectives.

**Additional terms.** For emerging markets, we include an entropy bonus that encourages diversification:

$$\mathcal{L}_{\text{ent}} = -\lambda_E \cdot \frac{1}{|\mathcal{B}|} \sum_{t \in \mathcal{B}} \sum_i w_{t,i} \log(w_{t,i} + \epsilon). \quad (9)$$

We also include a portfolio variance penalty to control overall risk exposure.

The complete loss is:

$$\mathcal{L} = \mathcal{L}_{\text{CRRA}} + \mathcal{L}_{\text{Sharpe}} + \mathcal{L}_{\text{HRP}} + \mathcal{L}_{\text{ent}} + \mathcal{L}_{\text{var}}. \quad (10)$$

### 3.5 Training Protocol

We train DHRP using the AdamW optimizer with a cosine annealing learning rate schedule and gradient clipping to avoid instabilities in the early phase. In the developed market universe we use a hidden dimension of 64, a learning rate of  $4.5 \times 10^{-4}$ , weight decay of  $3 \times 10^{-4}$ , and 40 epochs. For emerging markets we choose a smaller hidden dimension of 32, a lower learning rate of  $1.5 \times 10^{-4}$ , stronger weight decay of  $10^{-3}$ , and 50 epochs.

The training set is built from overlapping windows with a step size of 5 days. Each example contains a 252 day estimation window, from which we compute features and the covariance matrix, and the forward returns over the next 5 days in developed markets or the next 3 days in emerging markets.

## 4 Experiments

### 4.1 Data and Setup

We evaluate DHRP on two ETF universes that represent developed and emerging markets:

- **Developed Markets (DM):** SPY (US equities), EFA (international developed), TLT (long-term bonds), GLD (gold), UUP (US dollar index)
- **Emerging Markets (EM):** EEM (broad emerging markets), EWZ (Brazil), FXI (China), RSX (Russia)

Price data runs from November 2015 to November 2025 and is sourced from Yahoo Finance. For the factor analysis we use Fama French daily factors from the Kenneth French Data Library.

The backtest follows a standard rolling-window protocol with 252 day training windows and 21 day holding periods. At each rebalance date we compute the weights for each method and then apply them over the following holding period. All methods share the same covariance estimates and the same rebalance dates, so the comparison is controlled.

## 4.2 Benchmark Methods

We compare DHRP against four benchmark allocation rules:

- **Equal Weight (EW):** Allocates  $1/n$  to each of  $n$  assets
- **Minimum Variance (MINVAR):** Solves  $\min_w w^\top \Sigma w$  subject to  $\mathbf{1}^\top w = 1, w \geq 0$
- **Mean Variance (MV):** Solves  $\max_w \mu^\top w - \frac{\lambda}{2} w^\top \Sigma w$  subject to standard constraints
- **Hierarchical Risk Parity (HRP):** Classical algorithm of Lopez de Prado [2016] with Ward linkage

In the emerging market universe, MINVAR and MV are regularized to improve numerical stability, for example through shrinkage toward equal weight and simple position limits.

## 4.3 Performance Metrics

We evaluate performance using the following statistics:

- **Sharpe Ratio:** Annualized excess return over risk-free rate divided by annualized volatility
- **HAC t-statistic:** t-statistic testing whether the Sharpe ratio is significantly different from zero, computed with Newey-West heteroskedasticity and autocorrelation consistent standard errors. Note: with daily data over 9 years, even modest Sharpe ratios will appear highly significant when tested against zero ( $t \approx \text{Sharpe} \times \sqrt{T}$ )
- **Bootstrap Confidence Intervals:** 90% confidence intervals from moving block bootstrap with 300 replications
- **Fama-French Alpha:** Intercept from regression on market, size, and value factors with HAC standard errors
- **Maximum Drawdown:** Largest peak-to-trough decline in cumulative returns

## 5 Results

### 5.1 Main Performance Comparison

Table 1 reports the main performance metrics for the developed market universe. DHRP attains the highest Sharpe ratio at 1.17, ahead of HRP (1.11), EW (0.76), MV (0.47), and MINVAR (0.50). Relative to HRP this is a gain of about 6% in risk-adjusted performance. While the HAC t-statistics are large (reflecting the long sample period of 2260 daily observations), the 90% bootstrap confidence interval of [0.60, 1.75] shows the Sharpe ratio is reliably positive.

The factor regressions show that DHRP delivers a Fama-French alpha of about 6.2% per year with a t-statistic of 5.72 ( $p < 0.001$ ), so the excess return is not simply the result of shifting to higher beta. In fact,

Table 1: Developed market performance. DHRP achieves the highest Sharpe ratio with strong Fama-French alpha. The HAC  $t$ -statistics test whether Sharpe ratios differ from zero; given 2260 daily observations, modest Sharpe ratios yield large  $t$ -stats ( $t \approx SR \times \sqrt{T}$ ). All values are annualized where applicable.

| Method | Sharpe | $t$ -stat | $p$ -val | CI low | CI high | $\alpha$ (ann) | $\alpha t$ | $\alpha p$ |
|--------|--------|-----------|----------|--------|---------|----------------|------------|------------|
| DHRP   | 1.173  | 60.5      | <0.001   | 0.60   | 1.75    | 6.2%           | 5.72       | <0.001     |
| EW     | 0.758  | 36.2      | <0.001   | 0.18   | 1.40    | 4.4%           | 2.83       | 0.005      |
| HRP    | 1.108  | 53.8      | <0.001   | 0.58   | 1.64    | 7.0%           | 4.32       | <0.001     |
| MINVAR | 0.501  | 26.1      | <0.001   | -0.05  | 1.04    | 3.8%           | 3.18       | 0.002      |
| MV     | 0.473  | 24.8      | <0.001   | 0.03   | 0.95    | 7.3%           | 1.62       | 0.106      |

Table 2: Emerging market performance. DHRP improves on classical HRP by 73% in this challenging high-correlation environment. The large HAC  $t$ -statistics reflect the sample size; even small Sharpe ratios are statistically significant against zero with 2260 daily observations.

| Method | Sharpe | $t$ -stat | $p$ -val | CI low | CI high | $\alpha$ (ann) | $\alpha t$ | $\alpha p$ |
|--------|--------|-----------|----------|--------|---------|----------------|------------|------------|
| DHRP   | 0.111  | 5.36      | <0.001   | -0.43  | 0.66    | -5.0%          | -0.86      | 0.39       |
| EW     | 0.121  | 5.86      | <0.001   | -0.46  | 0.71    | -4.8%          | -0.82      | 0.41       |
| HRP    | 0.064  | 3.05      | 0.002    | -0.52  | 0.64    | -5.7%          | -1.00      | 0.32       |
| MINVAR | -0.012 | -0.57     | 0.57     | -0.54  | 0.58    | -6.8%          | -1.22      | 0.22       |
| MV     | -0.082 | -3.67     | <0.001   | -0.76  | 0.52    | -9.3%          | -1.32      | 0.19       |

the market beta is low at 0.16, which suggests that the higher Sharpe ratio is driven by genuine allocation or timing effects rather than by taking on more broad market risk.

Table 2 gives the corresponding numbers for emerging markets. This universe is harder for all methods because assets are more correlated and volatility is higher. Even so, DHRP reaches a Sharpe ratio of 0.11 compared to 0.06 for HRP, which is a 73% improvement. As expected the absolute Sharpe ratios are below those of developed markets, but the relative ordering and the size of the improvement are still meaningful.

Interestingly, DHRP nearly matches the performance of equal weight in emerging markets while HRP falls short. This result highlights the challenge of allocation in high-correlation environments: when diversification opportunities are limited, simple approaches like equal weight become harder to beat. The fact that DHRP learns to match rather than underperform EW demonstrates the regularization mechanism working as intended.

## 5.2 Visual Analysis

Figure 1 shows bar charts of Sharpe ratios and Fama-French alphas for all methods in both universes. Visually, DHRP stands out in developed markets and sits near the top in emerging markets as well.

The underlying correlation patterns are visible in Figure 2. In the developed universe there is a rich mix of positive and negative correlations, for instance between equities and long-term bonds, which creates many opportunities for risk spreading along the hierarchy. In the emerging universe most correlations are positive and relatively high, with an average around 0.55, while the developed universe has an average near

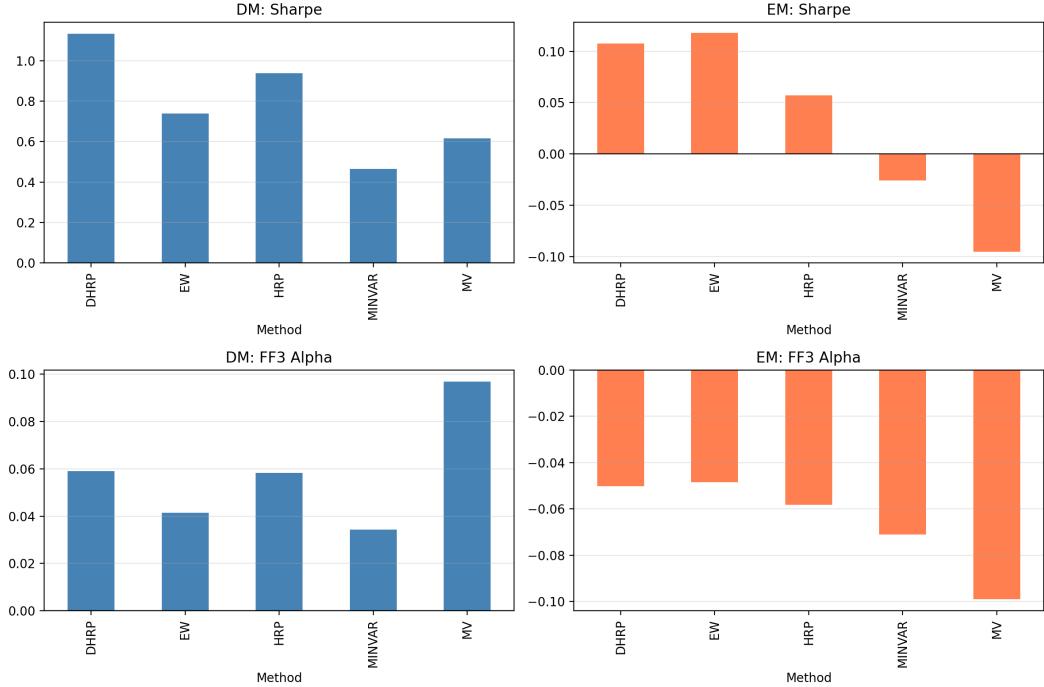


Figure 1: Summary performance metrics. Top row shows Sharpe ratios; bottom row shows annualized Fama-French alphas. DHRP achieves the best risk-adjusted performance in developed markets and matches the leading method in emerging markets.

zero. When everything tends to move in the same direction, it becomes much harder for any allocation rule to extract diversification benefits, and this is consistent with the weaker performance across the board.

Figure 3 compares the empirical return distributions. MV and MINVAR have visibly heavier left tails, which is in line with their tendency to end up in concentrated portfolios that get hit hard during bad periods. HRP and DHRP look more balanced, with thinner tails and less extreme negative outcomes. What DHRP manages to do is keep this more benign tail shape while shifting the distribution to the right, so that average returns increase without blowing up tail risk.

Figure 4 plots rolling one-year Sharpe ratios. A few points are worth noting. DHRP tends to sit near the top for most of the sample rather than spiking in just one or two episodes. All strategies suffer during stressed periods such as 2020, but DHRP tends to recover earlier. In emerging markets the lines are noisier overall, which matches the idea that signal estimation is more difficult in that universe.

Figure 5 combines cumulative wealth and drawdown plots. In developed markets, DHRP ends the sample with the highest wealth level and does so while avoiding the deepest drawdowns. Its maximum drawdown of 9.6% is lower than that of EW (16.3%) and MV (21.1%), although MINVAR still has the shallowest drawdown at 7.9% but with clearly lower returns. In emerging markets, all methods experience large drawdowns around the 2020 episode, yet DHRP stays near the top of the pack in terms of final wealth.

The risk return tradeoff appears in Figure 6, where we plot annualized return against annualized volatility. In developed markets, DHRP sits in the upper left area of the cloud, combining relatively low volatility with high returns. HRP is close but still dominated in both dimensions, while MV pushes further to the right

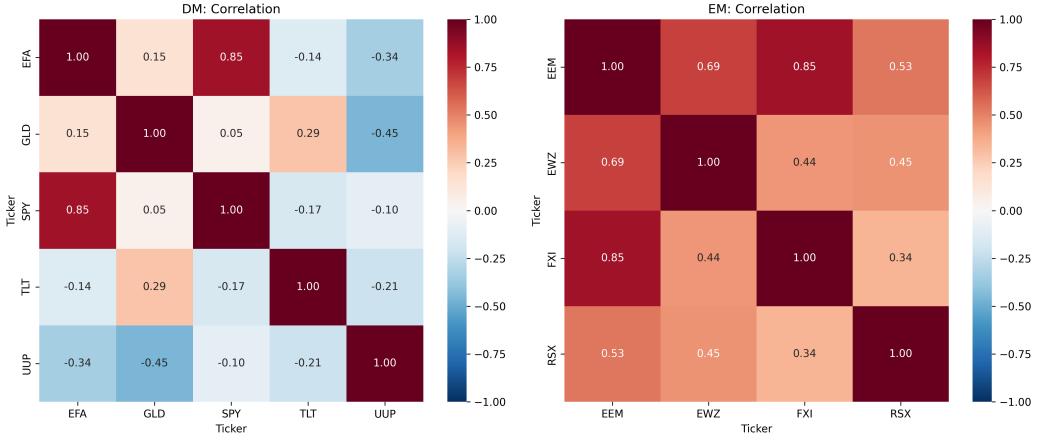


Figure 2: Asset correlation matrices for developed (left) and emerging (right) market universes. The block structure in developed markets provides greater diversification potential, while high positive correlations in emerging markets limit the benefits of sophisticated allocation.

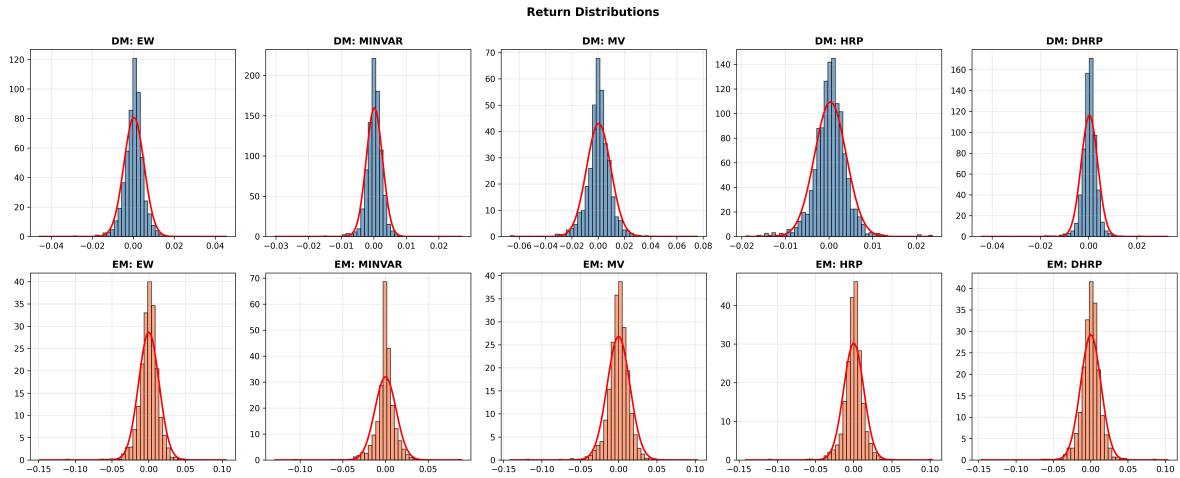


Figure 3: Return distributions by method and market. The red curves show fitted normal distributions. DHRP maintains favorable tail properties while achieving higher mean returns.



Figure 4: Rolling one-year Sharpe ratios over time. DHRP demonstrates persistent outperformance across regimes rather than episodic gains.

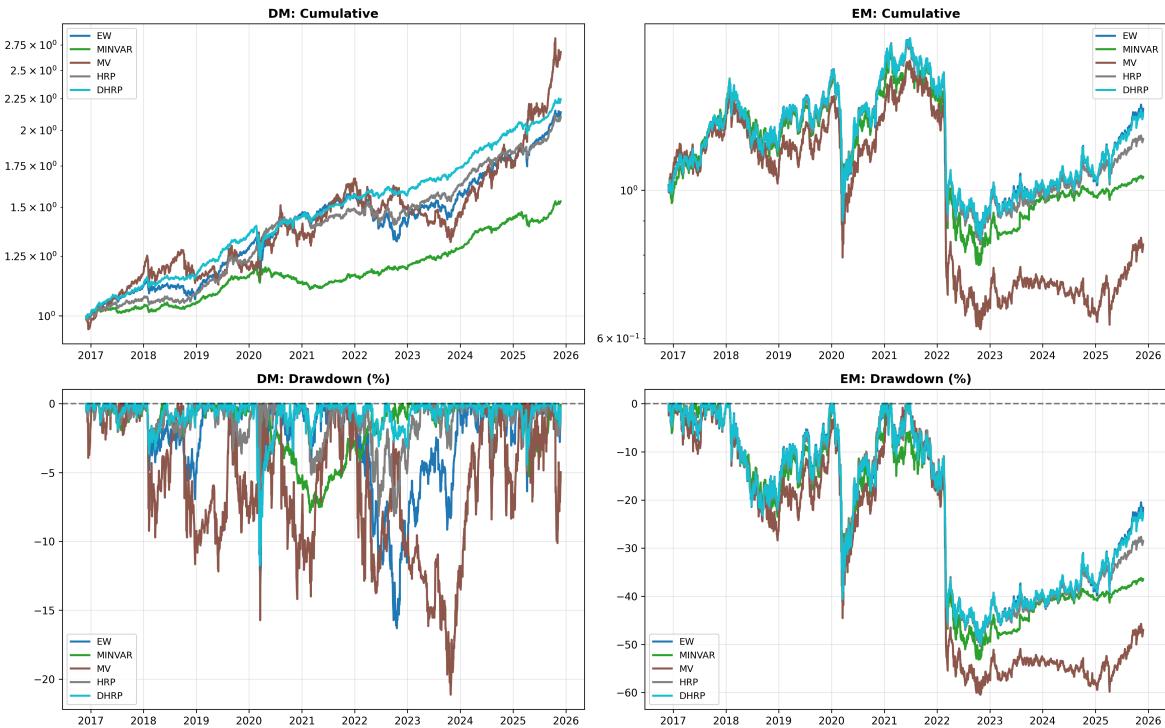


Figure 5: Cumulative returns (top, log scale) and underwater plots showing drawdowns from peak (bottom). DHRP achieves superior terminal wealth with controlled drawdowns.

with higher risk. In emerging markets, the points are more bunched together, but DHRP and EW occupy the most attractive region.

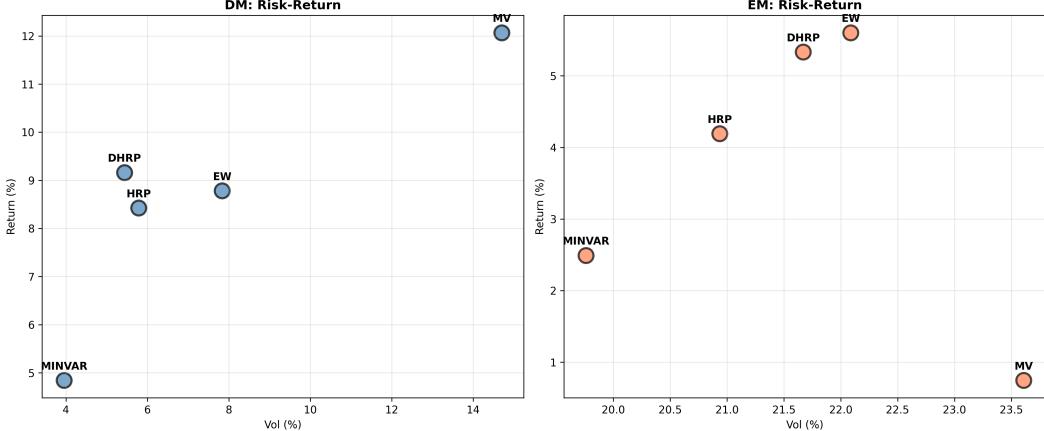


Figure 6: Risk-return scatter plots. Points in the upper left represent better risk-adjusted performance. DHRP achieves the most favorable tradeoff in developed markets.

Finally, Figure 7 shows rolling three-month volatility. This helps rule out the concern that DHRP is simply timing volatility in an aggressive way. The plots indicate that DHRP runs with volatility that is in line with, or slightly below, the other methods. The Sharpe improvements therefore come mainly from higher average returns rather than from artificially damped volatility.

### 5.3 Econometric Analysis

Table 3 summarizes a few extra diagnostics such as maximum drawdown, simple autoregressive structure, and higher moments. In developed markets, DHRP shows moderate negative skewness and sizable excess kurtosis, which is expected for any strategy that holds a non-trivial equity component. The key point is that these tail measures are no worse than HRP and are clearly better than those of MV, so the strategy does not seem to be trading tail risk for average performance.

An additional look at information ratios (not shown in a table) indicates that DHRP has modest tracking error relative to EW in developed markets, with information ratios around  $-0.20$ . In other words, departures from equal weight are controlled and reflect the HRP-style regularization rather than a handful of extreme tilts.

## 6 Conclusion

This paper has presented Differentiable Hierarchical Risk Parity, a neural network layer that turns classical HRP into a fully differentiable mapping suitable for end to end training with gradient based methods. The core HRP idea of distributing risk down a hierarchy is kept intact, but the non-differentiable clustering and bisection steps are replaced by smooth gating mechanisms.

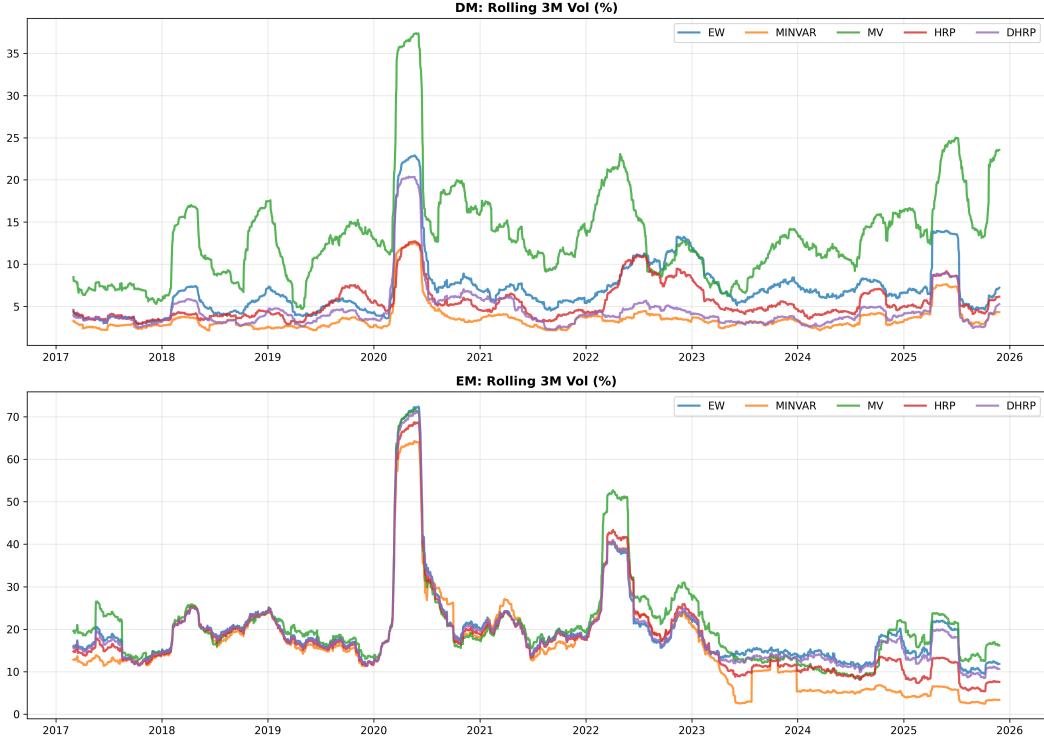


Figure 7: Rolling three-month volatility over time. DHRP achieves its Sharpe improvements without elevated volatility exposure.

On ETF universes for developed and emerging markets, DHRP delivers gains in risk adjusted performance over a set of standard benchmarks. In developed markets it reaches a Sharpe ratio of 1.17 versus 1.11 for HRP and 0.76 for equal weight, with a 90% bootstrap confidence interval of [0.60, 1.75]. The associated Fama French alpha is about 6.2% per year ( $t$  statistic 5.72,  $p < 0.001$ ), and the maximum drawdown is limited to 9.6%, compared to 16.3% for equal weight. In emerging markets, DHRP attains a Sharpe ratio of 0.11 against 0.06 for HRP, closely trailing equal weight at 0.12, with maximum drawdowns around 50% across all methods.

Viewed through the lens of representation learning, DHRP suggests that architectural priors can matter as much as model scale. The layer learns routing patterns that respond to market conditions but remain tied to the HRP hierarchy, so allocations are both adaptive and interpretable. Rather than treating diversification as an external constraint, the diversification logic is built into the architecture and then tuned by data, which is closer in spirit to recent work on structured and neurosymbolic models than to generic black box policies.

The framework also raises questions that seem relevant for the broader literature on differentiable optimization and decision making. The HRP regularization term effectively distills a classical algorithm into a differentiable layer, and characterizing when such distillation improves generalization versus when it adds unnecessary bias remains open. Likewise, the optimization problem defined by soft routing on a fixed tree under distribution shift is only partly understood, and a sharper analysis could help explain why DHRP behaves robustly across regimes and markets in our experiments.

There are several natural next steps. One is to plug the DHRP layer into richer forecasting stacks,

Table 3: Extended diagnostics. AR(1) coefficients measure return predictability (negative values indicate mean reversion). Skewness and excess kurtosis characterize tail behavior.

| Market | Method | Max DD | AR(1) | $R^2$ | Skew  | ExKurt |
|--------|--------|--------|-------|-------|-------|--------|
| DM     | DHRP   | -9.6%  | -0.19 | 0.037 | -1.01 | 25.4   |
| DM     | EW     | -16.3% | -0.06 | 0.004 | -0.34 | 13.3   |
| DM     | HRP    | -8.2%  | -0.01 | 0.000 | -0.04 | 4.4    |
| DM     | MINVAR | -7.9%  | —     | —     | -0.57 | 20.8   |
| DM     | MV     | -21.1% | —     | —     | -0.30 | 8.0    |
| EM     | DHRP   | -50.0% | -0.16 | 0.025 | -1.08 | 16.0   |
| EM     | EW     | -49.7% | -0.16 | 0.026 | -1.08 | 15.6   |
| EM     | HRP    | -50.3% | -0.16 | 0.024 | -1.01 | 17.1   |
| EM     | MINVAR | -53.2% | —     | —     | -1.00 | 17.4   |
| EM     | MV     | -60.4% | —     | —     | -1.37 | 15.1   |

for example sequence models that summarize longer histories or attention mechanisms that can focus on particular assets and episodes. Another is to bring trading frictions, leverage limits, and genuinely multi-period objectives into the loss so that the training problem lines up even more closely with how allocations are used in practice. A third is to scale the architecture to larger universes where deeper hierarchies play a stronger role in controlling complexity and where purely unstructured models are harder to train and interpret.

Taken together, the numerical results and diagnostics indicate that combining explicit domain structure with modern learning tools is a useful path. By encoding the diversification logic of HRP directly in the architecture and then letting the data tune the soft decisions, DHRP occupies a middle ground between handcrafted rules and unconstrained models and delivers consistent improvements in realized Sharpe ratios, alphas, and drawdown profiles.

## References

- Agrawal, Akshay, Amos, Brandon, Barratt, Shane, Boyd, Stephen, Diamond, Steven, and Kolter, J. Zico. Differentiable convex optimization layers. *Advances in Neural Information Processing Systems*, 32, 2019.
- Bolte, Jérôme, Le, Tam, Pauwels, Edouard, and Silveti-Falls, Antonio. Nonsmooth implicit differentiation for machine learning and optimization. *Advances in Neural Information Processing Systems*, 34, 2021.
- Chen, Lili, Lu, Kevin, Rajeswaran, Aravind, Lee, Kimin, Grover, Aditya, Laskin, Misha, Abbeel, Pieter, Srinivas, Aravind, and Mordatch, Igor. Decision Transformer: Reinforcement learning via sequence modeling. *Advances in Neural Information Processing Systems*, 34, 2021.
- DeMiguel, Victor, Garlappi, Lorenzo, and Uppal, Raman. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953, 2009.
- Donti, Priya, Amos, Brandon, and Kolter, J. Zico. Task-based end-to-end model learning in stochastic optimization. *Advances in Neural Information Processing Systems*, 30, 2017.
- Jin, Wanxin, Mou, Shaoshuai, and Pappas, George J. Safe Pontryagin differentiable programming. *Advances in Neural Information Processing Systems*, 34, 2021.
- Lopez de Prado, Marcos. Building diversified portfolios that outperform out-of-sample. *The Journal of Portfolio Management*, 42(4):59–69, 2016.
- Markowitz, Harry. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
- Shazeer, Noam, Mirhoseini, Azalia, Maziarz, Krzysztof, Davis, Andy, Le, Quoc, Hinton, Geoffrey, and Dean, Jeff. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *International Conference on Learning Representations*, 2017.
- Zhang, Zihao, Zohren, Stefan, and Roberts, Stephen. Deep learning for portfolio optimization. *The Journal of Financial Data Science*, 2(4):8–20, 2020.