Networks and Heterogeneous Media

https://www.aimspress.com/journal/nhm

NHM, 21(3): 1099–1145.
DOI: 10.3934/nhm.2026045
Received: 27 March 2026
Revised: 01 May 2026
Accepted: 14 May 2026
Published: 25 May 2026

Research article

A multiscale differential principal-feature architecture for portfolio optimization: Mathematical formulation and empirical results

Muhammad Hilal Alkhudaydi*

Department of Mathematics and Statistics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

* **Correspondence:** Email: mh.ayedh@tu.edu.sa.

Abstract: This paper proposed a multiscale differential principal-feature (MDPF) architecture for portfolio optimization in which the central contribution is an interpretable feature-extraction layer rather than a claim of unconditional benchmark dominance. The model combines the discrete wavelet transform (DWT), ordinary differential equation (ODE) identification, and principal component analysis (PCA) to transform rolling asset-return histories into low-dimensional dynamic signatures that are subsequently mapped into constrained portfolio overlays. In the U.S. mega-cap equity universe, the strongest MDPF overlay obtained a Sharpe ratio of 1.117 with average ℓ_1 turnover of 0.132, compared with 1.138 for the equal-weight benchmark and 1.187 for raw mean–variance with substantially higher ℓ_1 turnover of 0.464. In the global multi-asset ETF universe, the strongest MDPF overlay obtained a Sharpe ratio of 0.801, close to the equal-weight benchmark at 0.804 but below raw mean–variance at 0.927, again with lower ℓ_1 turnover than the raw mean–variance allocation. These results indicated that the proposed architecture is most useful when it is used to extract structured multiscale dynamic information and to express that information as controlled active tilts around a stable base portfolio. Supplementary ablation, statistical, regime, sensitivity, ODE-diagnostic, PCA-interpretability, feature-to-allocation, optimization-refinement, and transaction-cost diagnostics further clarify why the representation is economically meaningful while showing that benchmark dominance is not universal and that higher-order nonlinear components require further refinement. Accordingly, the paper positions MDPF as a transparent feature-extraction framework for dynamic portfolio construction and as a basis for future work on regime-aware multiscale allocation.

Keywords: multiscale portfolio optimization; discrete wavelet transform; ordinary differential equation identification; principal component analysis; dynamic asset allocation

Abbreviations and notation convention

Throughout the paper, *multiscale differential principal feature* (MDPF) denotes the proposed architecture, while DWT, ODE, and PCA denote the *discrete wavelet transform*, *ordinary differential*

equation, and *principal component analysis*, respectively. To preserve compatibility between the formal derivation and the empirical results, the overlay-intensity notation β_{ov} in the methodological formulation is equivalent to β_{overlay} in the empirical tables and sensitivity figures. The portfolio labels are used consistently as follows: MDPF-Ov denotes the MDPF overlay, MDPF-Full denotes the full-allocation version, WaveletPCA-Ov denotes the wavelet-only principal-component overlay, EqW denotes equal weight, InvVol denotes inverse volatility, MinVar denotes minimum variance, and RawMV denotes raw mean–variance. Performance columns are also used consistently: SR denotes the annualized Sharpe ratio, CAGR denotes compound annual growth rate, MDD denotes maximum drawdown, Avg. ℓ_1 turnover denotes the sum of absolute portfolio-weight changes at rebalancing, and Avg. TN denotes average traded notional, equal to one half of the ℓ_1 turnover used for transaction-cost accounting.

1. Introduction and research motivation

Portfolio optimization based directly on rolling returns remains difficult because the input data are noisy, high-dimensional, and highly sensitive to the estimation window. These problems are amplified when useful information is distributed across horizons. Short-run shocks, medium-term adjustment, and longer-horizon persistence can coexist in the same return window, creating signal aliasing across time scales and making it difficult for a conventional allocator to distinguish transient noise from persistent structures. A feature-extraction layer is therefore valuable when it can separate horizon-specific information, describe how the resulting scale components interact dynamically, and compress the extracted signal into a small set of stable variables that can be used inside a disciplined allocation rule.

The proposed MDPF architecture is designed around this feature-extraction problem. The DWT layer first separates each rolling return history into multiscale state components, enabling short- and longer-horizon movements to be represented separately. The ODE layer then estimates a local dynamic law in this multiscale state space, so that the features describe not only the magnitude of scale components but also their local cross-scale propagation, nonlinear interaction, common forcing, residual uncertainty, and instability. PCA is then applied to these structured ODE–wavelet signatures, rather than directly to raw returns, to obtain a low-dimensional set of principal features. Finally, the portfolio layer uses these principal features as an overlay signal around a stable base allocation. This design is consistent with evidence that multiscale financial information, deep feature learning, and constrained allocation interfaces are most useful when the feature layer and the portfolio decision layer are evaluated together rather than in isolation [1–5].

The empirical design is deliberately balanced. The model is not presented as a universal replacement for simple diversified portfolios or raw mean–variance allocation. Instead, the main question is whether a multiscale dynamic representation can produce economically interpretable features that remain useful when translated into controlled portfolio tilts. This interpretation is particularly important in finite samples, where simple benchmarks such as equal weight can be difficult to dominate, and aggressive mean–variance portfolios can achieve high Sharpe ratios at the cost of materially higher turnover. Accordingly, the empirical analysis therefore evaluates performance and representation quality, with particular attention to whether the extracted features are interpretable, whether the overlay design controls estimation risk, and whether the architecture remains informative across two investment universes.

The study is organized around four research questions that guide the empirical evaluation. First, can wavelet decomposition convert noisy rolling returns into horizon-resolved state variables that are useful for portfolio construction? Second, does local ODE identification add information beyond static multiscale summaries by capturing cross-scale dynamics? Third, can PCA compression preserve the economic structure of these dynamic signatures while making them suitable for constrained allocation? Fourth, is the resulting representation more effective as a controlled overlay than as a full stand-alone allocator? The empirical analysis answers these questions using U.S. mega-cap equities and global multi-asset ETFs, with additional diagnostics reported in Appendix A.

2. Literature review

The relevant literature is best understood as a technical chain rather than as three isolated methods. MDPF begins with multiscale decomposition, uses dynamic modeling to describe interactions among the extracted scales, applies dimensionality reduction to obtain a compact representation, and finally embeds the representation in a constrained portfolio overlay. In this section, we follow that sequence and then position the proposed architecture relative to recent feature-based allocation methods.

2.1. Multiscale financial feature extraction

Wavelet methods are widely used in financial time-series analysis because they separate non-stationary return variation into horizon-specific components. Empirical studies show that wavelet and MODWT decompositions can reveal co-movement, volatility persistence, and diversification patterns that are less visible in aggregate returns [1, 2, 6]. Related applications use wavelet analysis to study geopolitical-risk transmission, macro-financial dependence, and the behavior of developed and emerging markets across frequency bands [7, 8]. In this paper, the wavelet layer is not treated as a stand-alone forecast engine. Its role is to provide a structured multiscale state in which short-horizon noise and longer-horizon persistence can be represented separately before dynamic modeling and portfolio allocation are applied.

2.2. Dynamic modeling of multiscale states

Static multiscale summaries describe the level or energy of scale components, but they do not describe how these components evolve locally. Dynamic-system approaches, including ordinary and fractional differential-equation models, are useful because they focus on propagation, interaction, and local instability in complex time series [9, 10]. In financial settings, nonlinear and fractional dynamics have been used to study chaos, regime changes, and noise-affected systems [11–13]. MDPF adapts this idea by fitting a local ODE dictionary in wavelet space. The intercept block represents local drift, the linear block represents cross-scale propagation, the quadratic block represents pairwise nonlinear interaction, the common-forcing block represents exposure to the cross-sectional multiscale state, and the residual and instability terms quantify approximation error and local expansion risk.

2.3. Dimensionality reduction and economic interpretability

PCA remains an important tool in asset pricing, yield-curve modeling, and portfolio construction because it converts high-dimensional signals into orthogonal latent factors while preserving the dominant sources of

variation [14, 15]. Recent applications also combine PCA with multivariable time-series forecasting, clustering, and portfolio-control tasks [16, 17]. In MDPF, PCA is not applied blindly to raw returns. It is applied after the wavelet and ODE layers have imposed an economically interpretable feature structure. As a result, the retained components can be interpreted through the loading contributions of wavelet endpoints, wavelet energy, ODE linear dynamics, ODE forcing, nonlinear terms, residual variance, and instability. Appendix Table A9 uses this block-loading logic to document the empirical interpretation of the retained principal features.

2.4. Feature-based portfolio optimization and overlay design

Modern portfolio research increasingly treats forecasting models, feature learners, graph-based representations, causal signals, and adaptive decompositions as inputs to a constrained decision layer rather than as direct replacements for portfolio construction. Recent open-access work illustrates this shift. A two-stage deep-learning portfolio framework uses CNN, self-attention, graph convolution, and nearest-neighbor graph refinement to form a stock pre-selection layer before global minimum-variance allocation [3]. A deep portfolio-optimization framework similarly studies sequence and cross-asset feature modules inside a reinforcement-learning allocation system and uses ablations to show that the choice and fusion of feature modules materially affects portfolio decisions [4]. Graph and label-aware methods are useful for the same reason: They organize cross-sectional information and translate it into portfolio decisions while requiring risk constraints, turnover control, and benchmark awareness [18, 19]. Other studies emphasize that machine-learning and dynamic feature methods are more credible when they are connected to hedging, allocation, or optimization mechanisms that control implementation risk [5, 20–22]. This motivates the overlay form used in this paper. The extracted signal is not allowed to fully determine the portfolio; instead, it is expressed as an active tilt around a diversified base allocation. Table 1 presents how MDPF is positioned in comparison to related feature extraction and allocation methods.

Table 1. Positioning of MDPF relative to related feature-extraction and allocation approaches. The table summarizes the main representation mechanism, the portfolio interface, and the limitation that motivates the proposed architecture.

Approach	Main representation mechanism	Portfolio interface	Limitation addressed by MDPF
Wavelet-only methods	Horizon-specific return or volatility components [1, 2]	Often used for forecasting or diagnostic decomposition	Adds an ODE layer to model cross-scale dynamics rather than only scale magnitudes.
PCA/factor methods	Orthogonal factors extracted from high-dimensional inputs [14, 15]	Factor forecasting, covariance reduction, or allocation inputs	Applies PCA to structured dynamic signatures rather than directly to raw returns.
Dynamic-system methods	Local evolution, nonlinear interaction, and instability [9, 10]	Typically diagnostic or predictive	Embeds the local dynamic law inside a portfolio feature map.
Deep feature-selection and reinforcement-learning methods	CNN, attention, graph, LSTM, or correlation modules for latent market-state extraction [3, 4]	Stock pre-selection, global minimum-variance allocation, or learned portfolio policy	Emphasizes transparent multiscale dynamic features and auditable feature-to-allocation diagnostics.
Graph, causal, and adaptive feature methods	Cross-sectional relations, labels, or adaptive signals [5, 18]	Optimization or decision layer with constraints	Keeps the representation transparent and decomposable through wavelet, ODE, and PCA blocks.
MDPF architecture	Multiscale states, local ODE signatures, and principal-feature compression	Controlled overlay and full-allocation variants	Integrates multiscale, dynamic, and dimensionality-reduction stages while preserving diagnostic interpretability.

3. Contributions

This paper makes four contributions to the quantitative portfolio-optimization literature. First, we formulate MDPF as an integrated feature-extraction architecture rather than as a simple concatenation of mature tools. The novelty lies in the interaction between modules: DWT creates the multiscale state, ODE identification estimates a local dynamic law on that state, PCA compresses the resulting dynamic signatures, and the allocation layer converts the compressed representation into controlled portfolio tilts.

Second, we derive an interpretable feature map that records wavelet endpoints and energy, ODE intercept, linear dynamics, nonlinear interaction, common forcing, sparsity, instability, and residual uncertainty. This block structure makes it possible to diagnose which parts of the representation are active and economically meaningful instead of treating the extracted signal as a black box.

Third, for the empirical analysis, we evaluate the architecture in overlay and full-allocation forms across two universes. The evidence shows that the overlay implementation is the more credible portfolio interface because it limits estimation-error amplification, reduces turnover relative to aggressive full allocation, and keeps the extracted signal close to a diversified base portfolio.

Fourth, we provide supplementary diagnostics that make the feature-extraction claim auditable. Appendix A reports ablation, statistical, regime, wavelet-sensitivity, ODE-diagnostic, PCA-interpretability, feature-to-allocation, optimization-refinement, and transaction-cost analyses. These diagnostics clarify both the strengths of the representation and its current limitations, particularly the lack of universal benchmark dominance and the weak empirical activation of some nonlinear blocks.

4. Outline of the paper

The remainder of the paper is organized as follows: In Section 5, we present the mathematical formulation and sequential solution algorithm, formalizing the transformation from raw returns to multiscale states, ODE signatures, principal features, and portfolio weights. In Section 6, we report the empirical results across the equity and ETF universes, including realized performance, benchmark-relative behavior, feature diagnostics, and component analysis. In Section 7, we discusse the practical interpretation, limitations, and future research directions. In Section 8, we present the conclusion of the study. The appendices provide supplementary robustness checks, feature diagnostics, and notation details.

5. Mathematical formulation and solution algorithm

In this section, we detail the mathematical construction of the proposed multiscale differential principal features architecture. We formalize the sequential transformation from raw return histories to final portfolio weights, integrating the components reviewed in Section 2.

The proposed MDPF architecture is designed to extract structurally informative predictors from asset returns before portfolio weights are chosen. The central modeling objective is not to optimize directly on raw returns, but rather to map each return history into a multiscale state, identify the local law governing that state, compress the resulting dynamical signature into a low-dimensional latent representation, and then use that representation inside a constrained portfolio-allocation problem. This construction couples three complementary mathematical ingredients. First, the DWT resolves the return

path across investment horizons while preserving energy under an orthonormal basis. Second, an ODE is fitted in wavelet space so that the extracted features encode *how* scales interact, rather than only *how large* they are. Third, PCA is applied to the ODE-derived signatures instead of raw returns, thereby producing orthogonal factors that summarize the geometry of multiscale dynamics. The resulting model is therefore a sequential learning-and-allocation system in which feature extraction and portfolio construction are mathematically linked but conceptually separated.

Throughout, let $\mathcal{N} = \{1, \dots, N\}$ index assets, let $\mathcal{T} = \{L, \dots, T-1\}$ index rebalancing dates, let $\mathcal{J} = \{1, \dots, J\}$ denote wavelet detail levels, and let $\mathcal{S} = \{1, \dots, S\}$ denote the normalized interpolation grid used for ODE estimation. Tables 2 and 3 define all symbols appearing in the formulation and algorithm.

Table 2. Notation and abbreviations for the data, wavelet layer, and ODE layer.

Symbol	Description	Type	Unit
$\mathcal{N}$	Set of assets, $\mathcal{N} = \{1, \dots, N\}$	Set	N/A
$\mathcal{T}$	Set of rebalancing dates, $\mathcal{T} = \{L, \dots, T-1\}$	Set	Date index
$\mathcal{J}$	Set of wavelet detail levels, $\mathcal{J} = \{1, \dots, J\}$	Set	Scale index
$\mathcal{S}$	Interpolation grid, $\mathcal{S} = \{1, \dots, S\}$	Set	Grid index
i, t, j, s, k	Asset, date, wavelet scale, interpolation index, and principal-feature index	Index	N/A
N, T, L, J, S	Number of assets, sample length, rolling-window length, wavelet depth, and grid size	Parameter	Count
$r_{i,t}$	Log return of asset i at date t	Variable	Return
$y_{i,t}$	Rolling return window $(r_{i,t-L+1}, \dots, r_{i,t})^\top$	Variable	Return vector
$\mathcal{W}_J$	J-level orthonormal DWT operator	Parameter	N/A
$a_{i,t}^{(J)}$	Approximation coefficients at level J	Variable	Coefficient
$d_{i,t}^{(j)}$	Detail coefficients at level j	Variable	Coefficient
ν_j	Soft-threshold at detail level j	Parameter	Coefficient
$\widetilde{d}_{i,t}^{(j)}$	Thresholded detail coefficients	Variable	Coefficient
R_0, R_j	Inverse reconstruction operators for approximation and detail components	Parameter	N/A
$u_{i,t}^{(0)}, u_{i,t}^{(j)}$	Reconstructed time-domain scale components	Variable	Return vector
$U_{i,t}$	Multiscale state matrix $[u_{i,t}^{(0)}, u_{i,t}^{(1)}, \dots, u_{i,t}^{(J)}]$	Variable	Return matrix
M	Number of multiscale channels, $M = J + 1$	Parameter	Count
$X_{i,t}$	Interpolated multiscale state on the normalized grid	Variable	State matrix
$\dot{X}_{i,t}$	Numerical time derivative of $X_{i,t}$	Variable	State rate
G	Differentiation matrix on the normalized grid	Parameter	N/A
$\overline{X}_t$	Cross-sectional mean multiscale state	Variable	State matrix
$\Phi_{i,t}$	ODE design dictionary built from $X_{i,t}$ and $\overline{X}_t$	Variable	Feature matrix
$\Psi(X_{i,t})$	Quadratic interaction block of pairwise Hadamard products	Variable	Feature matrix
Q_ϕ	Number of columns in the ODE dictionary	Parameter	Count
$B_{i,t}$	ODE coefficient matrix	Decision variable	Coefficient matrix
$\alpha_{i,t}$	Intercept block extracted from $B_{i,t}$	Variable	Coefficient
$B_{\mathrm{lin},i,t}$	Linear dynamics block of $B_{i,t}$	Variable	Coefficient matrix
$B_{\mathrm{quad},i,t}$	Quadratic interaction block of $B_{i,t}$	Variable	Coefficient tensor/matrix
$B_{\mathrm{for},i,t}$	Common-forcing block of $B_{i,t}$	Variable	Coefficient matrix
$\lambda_1^{(\mathrm{ODE})}, \lambda_2^{(\mathrm{ODE})}$	ℓ_1 and ℓ_2 penalties in the ODE estimation stage	Parameter	N/A
$\sigma_{\varepsilon,i,t}^2$	Mean squared ODE residual	Variable	Squared rate
$s_{i,t}$	Sparsity ratio of active ODE coefficients	Variable	Dimensionless
$\kappa_{i,t}$	Local instability score from the symmetric Jacobian part	Variable	Dimensionless
DWT, ODE, PCA	Discrete wavelet transform, ordinary differential equation, and principal component analysis	Abbreviation	N/A

Table 3. Notation and abbreviations for the PCA layer, predictive layer, and portfolio-allocation layer.

Symbol	Description	Type	Unit
$f_{i,t}$	Full ODE-wavelet signature of asset i at date t	Variable	Feature vector
$\overline{f}_t$	Rolling cross-sectional feature mean	Variable	Feature vector
D_t	Diagonal scaling matrix of feature standard deviations	Parameter	N/A
$\widetilde{f}_{i,t}$	Standardized feature vector	Variable	Dimensionless
C_t	Rolling covariance matrix of standardized features	Variable	Covariance matrix
V_t, Λ_t	Eigenvector and eigenvalue matrices of C_t	Variable	N/A
ρ_{PCA}	Explained-variance retention threshold for PCA	Parameter	Proportion
K_t	Number of retained principal features at date t	Variable	Count
$z_{i,t}$	Latent principal-feature score of asset i	Variable	Score vector
Z_t	Matrix of stacked principal-feature scores	Variable	Score matrix
H_f, H_μ, H_Σ	Lookback lengths for PCA, mean estimation, and covariance estimation	Parameter	Dates
$\hat{\alpha}_t, \hat{\beta}_t$	Coefficients of the predictive ridge regression	Decision variable	Coefficient
η	Ridge penalty in the predictive regression	Parameter	N/A
$\hat{\mu}_t$	Forecast return vector generated by the extracted features	Variable	Return vector
$\hat{g}_{t+1}$	Realized feature-factor return from cross-sectional projection	Variable	Return vector
$\hat{\Omega}_t$	Covariance matrix of feature-factor returns	Variable	Covariance matrix
$\hat{D}_t$	Idiosyncratic covariance matrix	Variable	Covariance matrix
$\hat{\Sigma}_t$	Asset covariance forecast	Variable	Covariance matrix
w_t^{base}	Exogenous base portfolio used by the overlay formulation	Parameter	Weight
δ_t	Active tilt chosen by the optimization model	Decision variable	Weight
$w_t^\star$	Final portfolio weights	Decision variable	Weight
β_{ov}	Overlay intensity multiplying the active tilt	Parameter	Dimensionless
γ	Risk-aversion coefficient in the allocation problem	Parameter	Dimensionless
λ_{TO}	Turnover penalty coefficient	Parameter	Dimensionless
λ_κ	Penalty coefficient for unstable extracted features	Parameter	Dimensionless
L_δ	Maximum active-tilt leverage budget	Parameter	Weight
ℓ, u	Lower and upper portfolio-weight bounds	Parameter	Weight
A, b	Linear exposure-constraint matrix and target vector	Parameter	N/A
Θ	Hyperparameter set for rolling model selection	Set	N/A
$\varepsilon_{\mathrm{ode}}$	Tolerance for the proximal ODE solver	Parameter	N/A
$\varepsilon_{\mathrm{cvx}}$	Tolerance for the convex allocation solver	Parameter	N/A
ν	Inner proximal-gradient iteration index	Index	N/A
MDPF	Multiscale differential principal features architecture	Abbreviation	N/A

5.1. Model description and assumptions

The model operates under three assumptions that are standard in rolling-horizon portfolio design. First, the return path of each asset is locally regular over a finite window of length L, so that a wavelet decomposition and a smooth interpolation are meaningful. Second, the multiscale state is locally governed by a first-order dynamical law with linear, quadratic, and common-forcing terms. This assumption does not claim a globally correct physical model; rather, it provides a tractable local generator whose coefficients can be interpreted as the dynamic signature of the return process. Third, the extracted signatures are informative for cross-sectional allocation after orthogonal compression. Under these assumptions, the portfolio decision variable is not the raw feature vector but the active tilt δ_t imposed on a stable base allocation w_t^{base}. This separation is important: The wavelet-ODE-PCA pipeline performs representation learning, whereas the final optimization stage determines how aggressively that representation is translated into investable weights.

5.2. Wavelet multiscale embedding

For each asset $i \in \mathcal{N}$ and rebalance date $t \in \mathcal{T}$, define the rolling log-return vector

$$y_{i,t} = \begin{bmatrix} r_{i,t-L+1} & r_{i,t-L+2} & \cdots & r_{i,t} \end{bmatrix}^\top \in \mathbb{R}^L. \tag{5.1}$$

A J-level orthonormal DWT maps $y_{i,t}$ into one approximation block and J detail blocks,

$$\mathcal{W}_J y_{i,t} = \left(a_{i,t}^{(J)}, d_{i,t}^{(J)}, d_{i,t}^{(J-1)}, \ldots, d_{i,t}^{(1)}\right). \tag{5.2}$$

To attenuate high-frequency noise while preserving localized shape information, the detail coefficients are soft-thresholded:

$$\widetilde{d}_{i,t}^{(j)} = \operatorname{sign}\left(d_{i,t}^{(j)}\right)\left(\left|d_{i,t}^{(j)}\right| - \nu_j\right)_+, \qquad j \in \mathcal{J}. \tag{5.3}$$

The approximation and detail channels are then reconstructed in the time domain,

$$u_{i,t}^{(0)} = R_0 a_{i,t}^{(J)}, \qquad u_{i,t}^{(j)} = R_j \widetilde{d}_{i,t}^{(j)}, \quad j \in \mathcal{J}, \tag{5.4}$$

and stacked into the multiscale state matrix

$$U_{i,t} = \left[u_{i,t}^{(0)}, u_{i,t}^{(1)}, \ldots, u_{i,t}^{(J)}\right] \in \mathbb{R}^{L \times M}, \qquad M = J + 1. \tag{5.5}$$

Because the wavelet basis is orthonormal, the original window energy is preserved by construction:

$$\left\|y_{i,t}\right\|_2^2 = \left\|u_{i,t}^{(0)}\right\|_2^2 + \sum_{j=1}^{J} \left\|u_{i,t}^{(j)}\right\|_2^2. \tag{5.6}$$

The energy identity above is important because it shows that the wavelet stage is not an arbitrary smoothing device. It is an energy-preserving multiresolution embedding that keeps the contribution of each investment horizon mathematically identifiable.

To formulate a continuous-time local law, $U_{i,t}$ is interpolated on a normalized grid $s \in \mathcal{S}$:

$$X_{i,t} = \mathcal{I}(U_{i,t}) \in \mathbb{R}^{S \times M}, \qquad \dot{X}_{i,t} = G X_{i,t}, \tag{5.7}$$

where $\mathcal{I}$ is a smoothing interpolation operator and G is a numerical differentiation matrix. The cross-sectional mean state is

$$\overline{X}_t = \frac{1}{N} \sum_{i=1}^{N} X_{i,t}. \tag{5.8}$$

The preceding construction therefore moves from a raw return window to a scale-separated and differentiable state representation. No portfolio weights are chosen at this stage; the purpose is only to construct a horizon-resolved state on which local dynamics can be estimated.

The empirical justification for the wavelet specification is separated from the formal definition. The supplementary grid in Table A7 varies the mother wavelet, decomposition depth, thresholding rule, and threshold multiplier under the same frozen return panel, thereby showing whether the selected multiscale representation is robust to reasonable wavelet-design perturbations.

5.3. ODE identification in wavelet space

The local multiscale dynamics are assumed to satisfy a linear-quadratic first-order ODE in wavelet space. For each asset and date,

$$\dot{X}_{i,t} = \Phi_{i,t} B_{i,t} + E_{i,t}, \tag{5.9}$$

where $E_{i,t}$ is a residual matrix and the design dictionary is

$$\Phi_{i,t} = \left[\mathbf{1}_S, X_{i,t}, \Psi(X_{i,t}), \overline{X}_t\right] \in \mathbb{R}^{S \times Q_\phi}. \tag{5.10}$$

The quadratic block $\Psi(X_{i,t})$ contains all distinct pairwise Hadamard products of the columns of $X_{i,t}$:

$$\Psi(X_{i,t}) = \left[x_{\cdot 1} \odot x_{\cdot 1}, x_{\cdot 1} \odot x_{\cdot 2}, \ldots, x_{\cdot M} \odot x_{\cdot M}\right]. \tag{5.11}$$

Accordingly, $B_{i,t}$ partitions as

$$B_{i,t} = \left[\alpha_{i,t}^\top \quad B_{\text{lin},i,t}^\top \quad B_{\text{quad},i,t}^\top \quad B_{\text{for},i,t}^\top\right]^\top. \tag{5.12}$$

The coefficient matrix is estimated by the elastic-net problem

$$\widehat{B}_{i,t} = \arg\min_{B \in \mathbb{R}^{Q_\phi \times M}} \left\{ \frac{1}{2S} \left\| \dot{X}_{i,t} - \Phi_{i,t} B \right\|_F^2 + \lambda_1^{(\text{ODE})} \|B\|_{1,1} + \frac{\lambda_2^{(\text{ODE})}}{2} \|B\|_F^2 \right\}. \tag{5.13}$$

This stage is mathematically special because it estimates a *local generator* of the multiscale state rather than a static descriptor. The linear block captures first-order propagation across scales, the quadratic block captures nonlinear interactions, and the forcing block records co-movement with the market-wide multiscale state. Consequently, the feature vector will depend on the law of motion of returns in wavelet space, not merely on their levels.

A proximal-gradient iteration for solving Eq (5.13) is

$$B_{i,t}^{(\nu+1)} = \mathcal{S}_{\eta_\nu \lambda_1^{(\text{ODE})}} \left(B_{i,t}^{(\nu)} - \eta_\nu \left[\frac{1}{S} \Phi_{i,t}^\top \left(\Phi_{i,t} B_{i,t}^{(\nu)} - \dot{X}_{i,t} \right) + \lambda_2^{(\text{ODE})} B_{i,t}^{(\nu)} \right] \right), \tag{5.14}$$

where $\mathcal{S}_\tau(\cdot)$ is the elementwise soft-thresholding operator. The iteration stops when

$$\frac{\left\| B_{i,t}^{(\nu+1)} - B_{i,t}^{(\nu)} \right\|_F}{1 + \left\| B_{i,t}^{(\nu)} \right\|_F} \le \varepsilon_{\text{ode}}. \tag{5.15}$$

The nonlinear dictionary is intentionally retained in the formulation because it enables local cross-scale interactions to be measured rather than imposed. Its empirical contribution is assessed separately in Table A8, which reports coefficient-block activity, full-versus-linear fit gain, and instability-score behavior. This diagnostic treatment is important because weak nonlinear activity is a substantive limitation of this sample rather than a hidden implementation detail.

5.4. Construction of ODE-wavelet signatures

The extracted feature vector is assembled from interpretable blocks. Let the endpoint state be

$$\xi_{i,t} = X_{i,t}[S,:]^\top \in \mathbb{R}^M, \tag{5.16}$$

and define the multiscale energy vector by

$$e_{i,t,j} = \frac{1}{S}\sum_{s=1}^{S} X_{i,t}(s,j)^2, \qquad e_{i,t} = \begin{bmatrix} e_{i,t,1} & \cdots & e_{i,t,M} \end{bmatrix}^\top. \tag{5.17}$$

Let

$$\sigma^2_{\varepsilon,i,t} = \frac{1}{S}\left\| \dot{X}_{i,t} - \Phi_{i,t}\widehat{B}_{i,t} \right\|_F^2 \tag{5.18}$$

denote the average ODE misfit, and define the sparsity ratio

$$s_{i,t} = \frac{\left|\{(q,m) : \left|\widehat{B}_{i,t;q,m}\right| > \tau_0\}\right|}{Q_\phi\, M}, \tag{5.19}$$

where $\tau_0 > 0$ is a numerical activity threshold. To measure local instability, consider the Jacobian of the estimated drift map:

$$J^{\mathrm{loc}}_{i,t} = \frac{\partial}{\partial x^\top}\left[\widehat{B}^\top_{i,t}\phi(x,\overline{x}_t)\right]\Big|_{x=\xi_{i,t}}, \tag{5.20}$$

where $\phi(\cdot,\cdot)$ is the row-wise version of the dictionary and $\overline{x}_t = \overline{X}_t[S,:]^\top$. The instability score is

$$\kappa_{i,t} = \max\left\{0, \lambda_{\max}\left(\frac{J^{\mathrm{loc}}_{i,t} + J^{\mathrm{loc}\top}_{i,t}}{2}\right)\right\}. \tag{5.21}$$

The complete ODE-wavelet signature is then

$$f_{i,t} = \begin{bmatrix} \xi_{i,t} \\ e_{i,t} \\ \mathrm{vec}(\alpha_{i,t}) \\ \mathrm{vec}(B_{\mathrm{lin},i,t}) \\ \mathrm{vec}(B_{\mathrm{for},i,t}) \\ \mathrm{vec}(B_{\mathrm{quad},i,t}) \\ s_{i,t} \\ \kappa_{i,t} \\ \log\left(1+\sigma^2_{\varepsilon,i,t}\right) \end{bmatrix} \in \mathbb{R}^Q. \tag{5.22}$$

The complete signature clarifies the mathematical novelty of the representation. The extracted features jointly encode state, scale energy, dynamic propagation, nonlinear interaction, forcing sensitivity, sparsity, instability, and residual uncertainty. The model therefore transforms a return window into a *dynamic multiscale signature* rather than a purely statistical summary.

5.5. Principal compression of dynamic signatures

Let

$$F_t = \begin{bmatrix} f_{1,t}^\top \\ \vdots \\ f_{N,t}^\top \end{bmatrix} \in \mathbb{R}^{N\times Q} \tag{5.23}$$

be the cross-sectional feature matrix at date t. Using the most recent H_f rebalancing dates, compute the rolling mean and diagonal scaling matrix

$$\overline{f}_t = \frac{1}{NH_f} \sum_{\tau=t-H_f+1}^{t} \sum_{i=1}^{N} f_{i,\tau}, \qquad D_t = \mathrm{diag}(d_{1,t}, \ldots, d_{Q,t}), \tag{5.24}$$

where $d_{q,t}$ is the rolling standard deviation of feature q. The standardized feature vector is

$$\widetilde{f}_{i,t} = D_t^{-1} \left(f_{i,t} - \overline{f}_t \right). \tag{5.25}$$

The pooled covariance matrix of standardized dynamic signatures is

$$C_t = \frac{1}{NH_f - 1} \sum_{\tau=t-H_f+1}^{t} \widetilde{F}_\tau^\top \widetilde{F}_\tau, \tag{5.26}$$

where $\widetilde{F}_\tau$ stacks the rows $\widetilde{f}_{i,\tau}^\top$. Let

$$C_t = V_t \Lambda_t V_t^\top, \qquad \Lambda_t = \mathrm{diag}(\lambda_{1,t}, \ldots, \lambda_{Q,t}), \qquad \lambda_{1,t} \geq \cdots \geq \lambda_{Q,t} \geq 0, \tag{5.27}$$

and choose the smallest integer K_t such that

$$\frac{\sum_{q=1}^{K_t} \lambda_{q,t}}{\sum_{q=1}^{Q} \lambda_{q,t}} \geq \rho_{\mathrm{PCA}}. \tag{5.28}$$

The principal dynamic score of asset i is

$$z_{i,t} = V_{t,1:K_t}^\top \widetilde{f}_{i,t} \in \mathbb{R}^{K_t}. \tag{5.29}$$

The role of PCA here is unusual and mathematically important: Orthogonal compression is applied to ODE-wavelet signatures rather than to raw returns or covariance matrices. Hence, the retained components summarize the dominant directions of *dynamic multiscale law variation* across assets.

The PCA step is interpreted through loading-block contributions rather than treated as a purely numerical compression. Table A9 reports the dominant feature block associated with the leading principal directions, while Table A10 links the extracted blocks to active tilts, active alpha scores, and realized top-minus-bottom spreads.

5.6. Predictive layer and portfolio-allocation model

The extracted scores enter a predictive ridge regression for the next-period return:

$$(\widehat{\alpha}_t, \widehat{\beta}_t) = \arg\min_{\alpha,\beta} \left\{ \sum_{\tau=t-H_\mu}^{t-1} \|r_{\tau+1} - \alpha \mathbf{1}_N - Z_\tau \beta\|_2^2 + \eta \|\beta\|_2^2 \right\}, \tag{5.30}$$

where Z_τ stacks the principal-feature scores. The implied mean forecast is

$$\widehat{\mu}_t = \widehat{\alpha}_t \mathbf{1}_N + Z_t \widehat{\beta}_t. \tag{5.31}$$

A feature-factor covariance model is obtained by projecting realized returns onto the feature scores:

$$\widehat{g}_{\tau+1} = \left(Z_\tau^\top Z_\tau + \epsilon I\right)^{-1} Z_\tau^\top \left(r_{\tau+1} - \bar{r}_{\tau+1} \mathbf{1}_N\right), \tag{5.32}$$

$$\widehat{\Omega}_t = \mathrm{Cov}\left(\widehat{g}_{t-H_\Sigma+1}, \ldots, \widehat{g}_t\right), \qquad \widehat{D}_t = \mathrm{diag}\left(\widehat{\mathrm{Var}}(u_1), \ldots, \widehat{\mathrm{Var}}(u_N)\right), \tag{5.33}$$

and

$$\widehat{\Sigma}_t = Z_t \widehat{\Omega}_t Z_t^\top + \widehat{D}_t. \tag{5.34}$$

The predictive and covariance equations convert principal-feature scores into the expected-return and covariance inputs required by the allocation layer. The forecasts are not used as unconstrained trading rules; they are passed only to the overlay problem below, which controls active risk, turnover, and exposure.

To separate representation learning from aggressive reallocation, the extracted signal is injected through an overlay. Let w_t^{base} be a stable base portfolio. The optimization chooses an active tilt δ_t by solving

$$\begin{aligned}
\delta_t^\star = \arg\min_{\delta \in \mathbb{R}^N} \quad & \frac{\gamma}{2} \delta^\top \widehat{\Sigma}_t \delta - \widehat{\mu}_t^\top \delta + \lambda_{\mathrm{TO}} \|\delta - \delta_{t-1}\|_1 + \lambda_\kappa \kappa_t^\top |\delta| && (5.35)\\
\text{s.t.} \quad & \mathbf{1}_N^\top \delta = 0, && (5.36)\\
& \|\delta\|_1 \le L_\delta, && (5.37)\\
& \ell \le w_t^{\mathrm{base}} + \beta_{\mathrm{ov}} \delta \le u, && (5.38)\\
& A^\top \left(w_t^{\mathrm{base}} + \beta_{\mathrm{ov}} \delta\right) = b. && (5.39)
\end{aligned}$$

The final portfolio is

$$w_t^\star = w_t^{\mathrm{base}} + \beta_{\mathrm{ov}} \delta_t^\star. \tag{5.40}$$

Constraint Eq (5.36) enforces self-financing active tilts, Eq (5.37) limits overlay aggressiveness, Eq (5.38) enforces box constraints on the final weights, and Eq (5.39) preserves desired exposure conditions. When $\beta_{\mathrm{ov}} = 1$ and $w_t^{\mathrm{base}} = \mathbf{0}$, the formulation collapses to a full allocator; otherwise, it acts as a shrinkage overlay that converts extracted features into controlled portfolio tilts.

5.7. Solution algorithm

The complete rolling solution routine is summarized in Algorithm 1. The algorithm alternates between multiscale state construction, ODE identification, PCA compression, predictive estimation, and convex portfolio optimization. At each rebalance date, only information available up to that date is used.

Algorithm 1 Rolling Wavelet-ODE-PCA feature extraction and overlay allocation

Require: Return matrix $\{r_{i,t}\}_{i\in\mathcal{N},\,t=1}^{T}$, hyperparameter set Θ, base portfolio rule $\{w_t^{\mathrm{base}}\}_{t\in\mathcal{T}}$, tolerances $\varepsilon_{\mathrm{ode}}$ and $\varepsilon_{\mathrm{cvx}}$

Ensure: Feature scores $\{z_{i,t}\}$, forecasts $\{\widehat{\mu}_t, \widehat{\Sigma}_t\}$, and portfolio weights $\{w_t^\star\}$

Initialize rolling histories, base portfolio, and previous active tilt $\delta_{L-1} = \mathbf{0}$
for $t = L, \ldots, T-1$ **do**
  Select hyperparameters $\Theta_t \in \Theta$ by nested walk-forward validation on historical data only
  for each asset $i \in \mathcal{N}$ **do**
    Form $y_{i,t}$ from the most recent L returns
    Compute the J-level DWT and threshold the detail coefficients
    Reconstruct $U_{i,t}$, interpolate to $X_{i,t}$, and compute $\dot{X}_{i,t} = GX_{i,t}$
    Build the dictionary $\Phi_{i,t}$ using $X_{i,t}$ and $\overline{X}_t$
    Solve the elastic-net ODE problem by proximal gradient until the relative update is at most $\varepsilon_{\mathrm{ode}}$
    Construct the feature vector $f_{i,t}$ from the endpoint, energy, coefficient, sparsity, instability, and residual blocks
  end for
  Standardize the feature panel and compute the rolling covariance matrix C_t
  Obtain V_t and K_t from the explained-variance rule, then compute $z_{i,t}$
  Estimate $\widehat{\mu}_t$ and $\widehat{\Sigma}_t$ from the predictive and feature-factor covariance models
  Solve the convex overlay problem (5.35)–(5.39) to tolerance $\varepsilon_{\mathrm{cvx}}$
  Set $w_t^\star = w_t^{\mathrm{base}} + \beta_{\mathrm{ov}}\delta_t^\star$ and update the realized portfolio state
end for
Stop when the last out-of-sample date $T-1$ is reached
return $\{f_{i,t}\}$, $\{z_{i,t}\}$, $\{\widehat{\mu}_t, \widehat{\Sigma}_t\}$, and $\{w_t^\star\}$

Algorithm 1 shows that the model is solved sequentially rather than as a monolithic nonlinear program. The wavelet stage is deterministic, the ODE stage is a regularized convex problem, the PCA stage is an eigenvalue problem, the predictive stage is ridge regression and factor covariance estimation, and the portfolio stage is a convex overlay optimization. This decomposition is computationally attractive because each block has a stable numerical routine and an interpretable output. More importantly, it clarifies the intellectual contribution of the architecture: Wavelets create a scale-resolved state, ODE estimation turns that state into a local dynamical law, PCA compresses that law into orthogonal latent factors, and the overlay portfolio problem converts those factors into implementable investment tilts.

6. Results

In this section, we evaluate the proposed architecture across two investment universes: A U.S. mega-cap equity cross-section, denoted by $\mathcal{U}_{\mathrm{Eq}}$, and a global multi-asset ETF universe, denoted by $\mathcal{U}_{\mathrm{ETF}}$. The empirical analysis is organized around realized performance (Section 6.1), dynamic benchmark-relative behavior (Section 6.2), feature-extraction evidence (Section 6.3), and component analysis (Section 6.4). Comparator labels follow the notation convention defined earlier: MDPF-Ov, MDPF-Full, WaveletPCA-Ov, EqW, InvVol, MinVar, and RawMV.

The results are interpreted through the feature-extraction objective established in Sections 1–3. The analysis therefore asks whether the architecture produces economically interpretable multiscale dynamic features, whether those features translate into stable active tilts, and whether the overlay interface

controls the allocation risk of the learned signal. Benchmark performance remains important, but it is not the only criterion. This distinction is essential because simple diversified allocations and raw mean–variance portfolios can be strong competitors in finite samples, especially when the latter is enabled to take materially higher turnover.

All empirical diagnostics use a fixed sample window so that the major figures, tabulated results, and supplementary analyses are internally comparable. Appendix A documents the sample window and experiment registry in Tables A2 and A3. It then reports supplementary ablation, statistical, regime, wavelet-sensitivity, ODE-diagnostic, PCA-interpretability, feature-to-allocation, optimization-refinement, and transaction-cost evidence in Tables A4–A12. These supplementary diagnostics are used to explain the major empirical patterns rather than to replace the primary performance figures and tables.

6.1. Cross-universe summary of realized performance

Figure 1 and Table 4 summarize the realized annualized return–risk outcomes across the two universes. The results do not support a claim of universal outperformance. In $\mathcal{U}_{\mathrm{Eq}}$, the strongest MDPF configuration was the ridge-identified overlay, which attained a compound annual growth rate of 20.31%, annualized volatility of 18.02%, and a Sharpe ratio of 1.117. This exceeded the wavelet-only overlay benchmark, which delivered a Sharpe ratio of 1.104, but it remained below EqW and RawMV, whose Sharpe ratios were 1.138 and 1.187, respectively. In $\mathcal{U}_{\mathrm{ETF}}$, the strongest MDPF outcome was the overlay blend with weight 0.75, which reached a Sharpe ratio of 0.801 at annualized volatility of 8.29%. This nearly matched EqW at 0.804 and exceeded WaveletPCA-Ov and InvVol, but it lagged RawMV at 0.927.

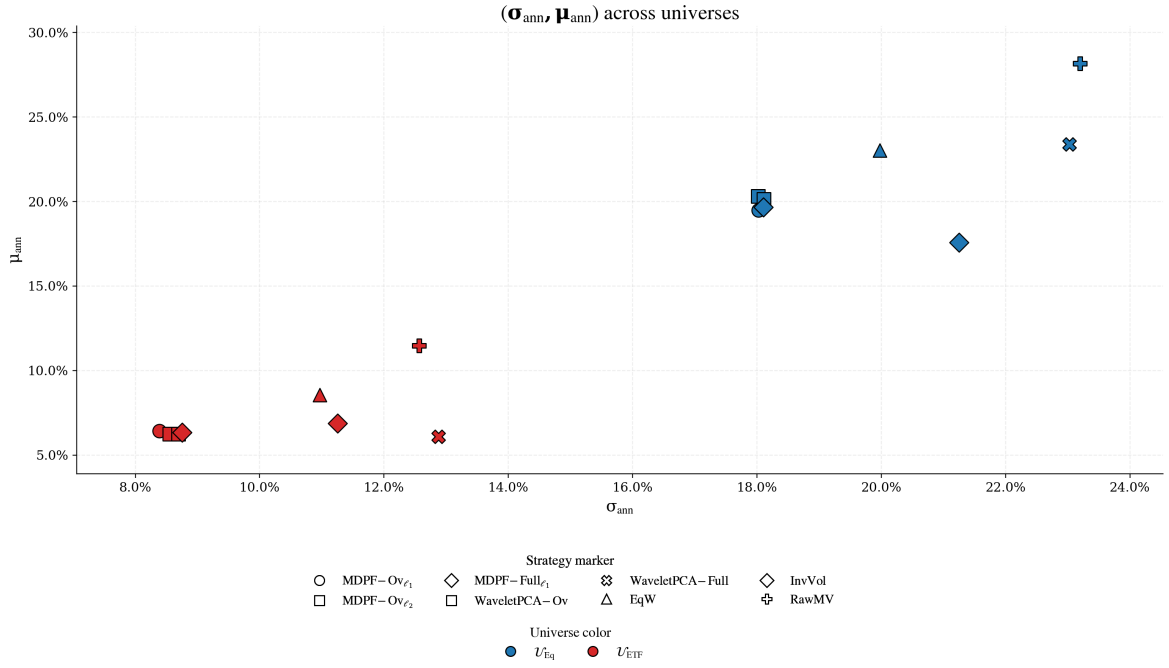


Figure 1. Realized annualized return–risk coordinates across $\mathcal{U}_{\mathrm{Eq}}$ and $\mathcal{U}_{\mathrm{ETF}}$. Marker shape identifies the strategy and marker color identifies the universe.

These results are important for interpretation. The model did not dominate the strongest benchmark in either universe, yet it remained competitive in risk-adjusted terms and, more importantly, did so

while operating through a richer representation of multiscale dynamics. Figure 1 reinforces this point visually: The MDPF points lie close to the core benchmark cloud rather than appearing as outliers, which indicates that the architecture generated benchmark-like return streams while altering the underlying state representation and allocation interface.

Table 4. Selected realized performance across the two evaluation universes. The table reports MDPF-Ov (ℓ_1 ODE), the strongest MDPF variant within each universe, the WaveletPCA-Ov baseline, and the benchmark comparators. Avg. ℓ_1 turnover is the average sum of absolute portfolio-weight changes at each rebalance.

Universe	Strategy	Role	CAGR (%)	Volatility (%)	Sharpe	Max drawdown (%)	Avg. ℓ_1 turnover
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1 ODE)	Main MDPF-Ov (ℓ_1 ODE)	19.49	18.03	1.078	29.59	0.169
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_2 ODE)	Best MDPF variant	20.31	18.02	1.117	30.39	0.132
$\mathcal{U}_{\mathrm{Eq}}$	WaveletPCA-Ov	Wavelet-only overlay baseline	20.14	18.11	1.104	30.08	0.137
$\mathcal{U}_{\mathrm{Eq}}$	EqW	Simple diversification benchmark	23.02	19.98	1.138	31.09	0.0480
$\mathcal{U}_{\mathrm{Eq}}$	InvVol	Risk-scaled benchmark	19.65	18.10	1.082	30.64	0.0620
$\mathcal{U}_{\mathrm{Eq}}$	RawMV	Best-performing benchmark in this run	28.17	23.20	1.187	29.51	0.464
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1 ODE)	Main MDPF-Ov (ℓ_1 ODE)	6.42	8.39	0.784	21.00	0.112
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov ($\beta_{\mathrm{overlay}} = 0.75$)	Best MDPF variant	6.50	8.29	0.801	20.92	0.166
$\mathcal{U}_{\mathrm{ETF}}$	WaveletPCA-Ov	Wavelet-only overlay baseline	6.26	8.69	0.742	21.04	0.106
$\mathcal{U}_{\mathrm{ETF}}$	EqW	Simple diversification benchmark	8.56	10.97	0.804	22.18	0.0230
$\mathcal{U}_{\mathrm{ETF}}$	InvVol	Risk-scaled benchmark	6.34	8.75	0.746	21.13	0.0540
$\mathcal{U}_{\mathrm{ETF}}$	RawMV	Best-performing benchmark in this run	11.47	12.57	0.927	21.68	0.362

6.2. Dynamic benchmark-relative behavior

The benchmark-relative dynamics in Figures 2 and 3 clarify where the main MDPF-Ov (ℓ_1 ODE) specification diverged from the benchmark ensemble over time. In $\mathcal{U}_{\mathrm{Eq}}$, the active cumulative log-wealth path stayed close to zero against WaveletPCA-Ov and InvVol for long subperiods, while the gap against RawMV drifted persistently negative. In $\mathcal{U}_{\mathrm{ETF}}$, the same general pattern remained visible, but the negative gap relative to EqW and RawMV became more pronounced over longer horizons.

The lower panels of Figures 2 and 3 summarize divergence from the benchmark set using the aggregated ℓ_1 and ℓ_2 norms of the benchmark-gap vector. This presentation is useful because it separates two issues that a single cumulative-return line can obscure. The upper panel reveals which individual benchmark dominates or lags MDPF-Ov (ℓ_1 ODE) at a given point in time, while the lower panel reports the overall distance of the MDPF-Ov path from the benchmark ensemble. In both universes, the norms stayed moderate for substantial stretches, indicating that the proposed model often behaved like a structured tilt on a conventional portfolio rather than as a radically different trading rule. From the perspective of feature extraction, this is informative rather than disappointing: The architecture altered the representation meaningfully, yet the resulting return paths remained economically plausible and did not depend on extreme departures from the benchmark opportunity set.

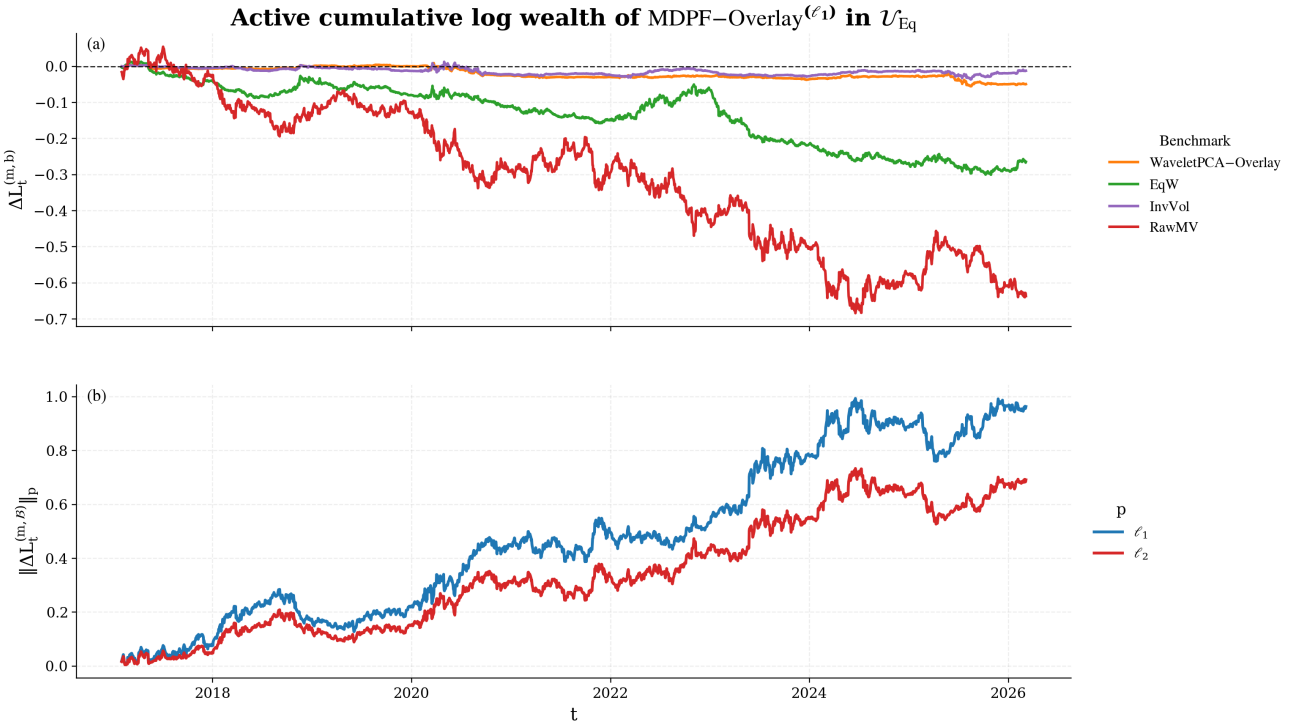


Figure 2. Active cumulative log-wealth gaps of the main MDPF-Ov (ℓ_1 ODE) in $\mathcal{U}_{\mathrm{Eq}}$, together with the aggregate ℓ_1 and ℓ_2 norms over the benchmark set.

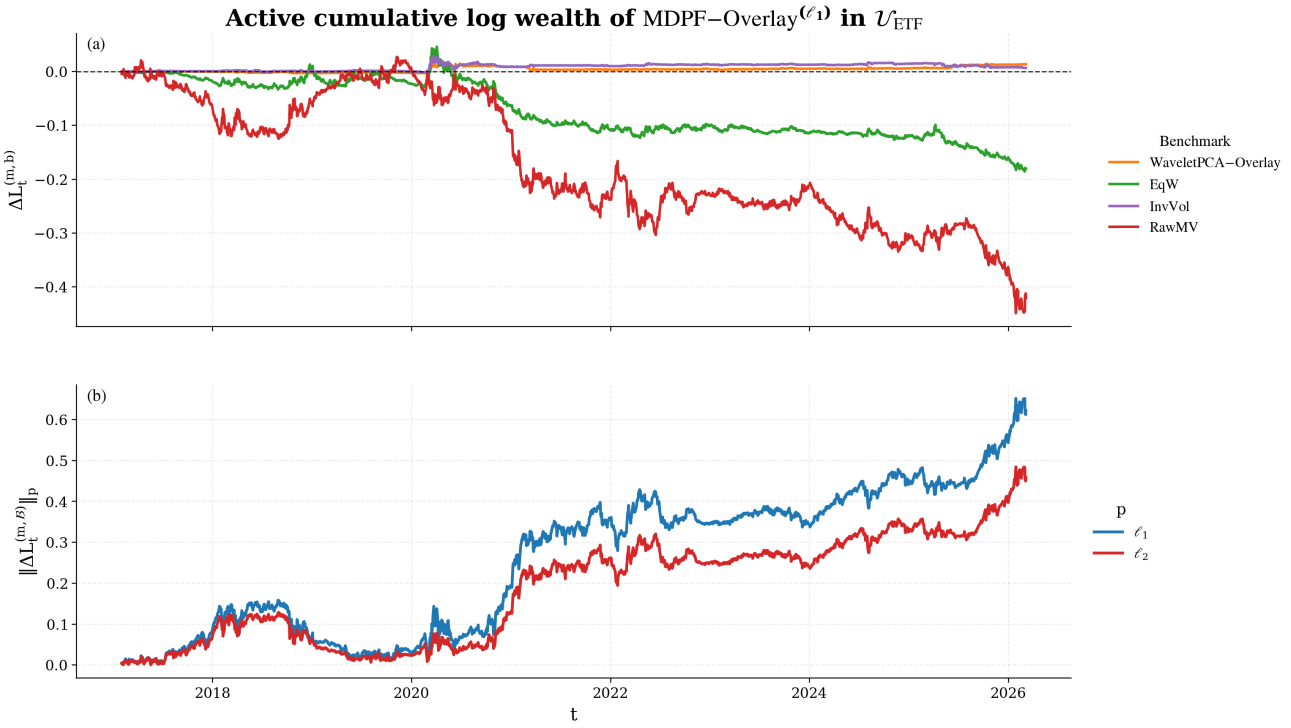


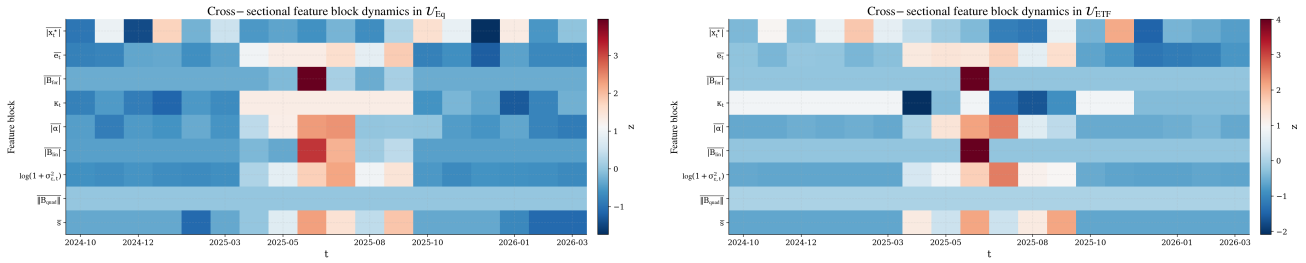
Figure 3. Active cumulative log-wealth gaps of the main MDPF-Ov (ℓ_1 ODE) in $\mathcal{U}_{\mathrm{ETF}}$, together with the aggregate ℓ_1 and ℓ_2 norms over the benchmark set.

6.3. *Evidence from the extracted features*

The strongest support for the proposed architecture comes from the feature diagnostics rather than from the headline performance rankings. Figure 4 reports the feature-block heatmaps for the two universes, while Figure 5 reports the corresponding heatmaps of the principal feature scores after compression. These plots show that the model produces pronounced cross-asset heterogeneity, especially in the residual-dynamic and linear-interaction blocks. The latent scores also remain well separated across assets after dimensionality reduction, which indicates that the PCA layer retained economically differentiated information rather than collapsing the representation into a nearly homogeneous factor.

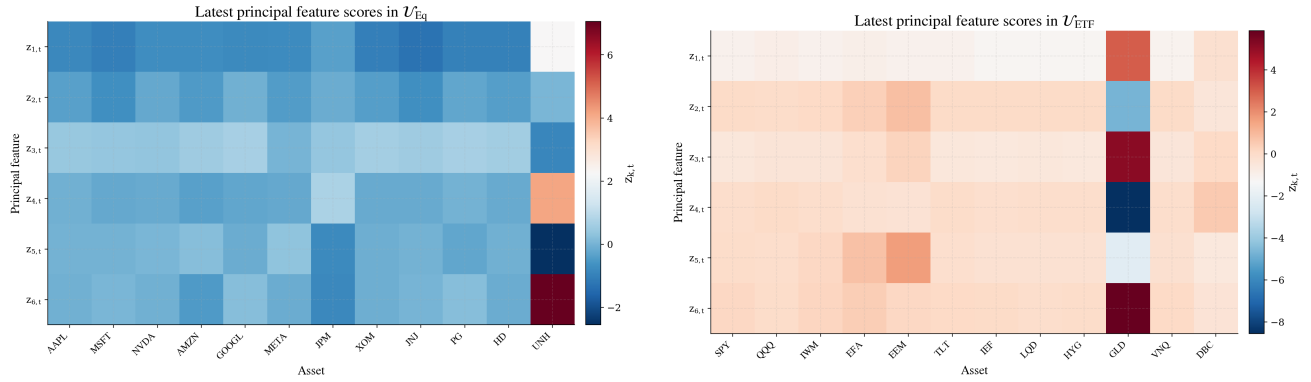
Table 5 quantifies the extracted blocks directly. In $\mathcal{U}_{\mathrm{Eq}}$, the largest historical cross-sectional dispersion appeared in the instability score, log residual variance, and linear interaction magnitude, whereas the common forcing block was inactive at the latest date. In $\mathcal{U}_{\mathrm{ETF}}$, the same ordering was weaker but still informative: Log residual variance and linear interaction magnitude remained the major sources of feature dispersion, while the common forcing magnitude and the quadratic interaction block were effectively dormant in the latest snapshot. The near-zero shares were revealing. Linear interaction magnitudes were sparse in both universes, but especially in $\mathcal{U}_{\mathrm{ETF}}$, where 91.67% of assets were near zero at the latest date. By contrast, log residual variance was active in both universes. This combination suggests that the architecture extracted persistent information from dynamic fit quality even when the higher-order interaction terms were mostly shut down.

Table 6 complements the block-level statistics with a compact cross-universe diagnostic summary. In both universes, the most dispersed block was log residual variance, followed by linear interaction magnitude. The most dispersed latent principal score was z_6 in $\mathcal{U}_{\mathrm{Eq}}$ and z_4 in $\mathcal{U}_{\mathrm{ETF}}$. The interpretation is consistent across the empirical outputs: The architecture is most informative when it captured variation in local dynamic fit and linear cross-scale interactions, not when it emphasized common forcing or highly nonlinear terms. That result is central to the methodological contribution. The model appears to be useful as a representation learner because it transforms return histories into a structured, heterogeneous, and low-dimensional state description even when the more ambitious nonlinear blocks are empirically weak.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure 4. Heatmaps of the extracted MDPF feature-block summaries across assets.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure 5. Heatmaps of the latest principal feature scores after PCA compression of the MDPF blocks.

Table 5. Feature-block activity and dispersion for the extracted representation. Historical dispersion is computed as the time average of the cross-sectional standard deviation of each feature block, while the latest snapshot summarizes the magnitude and activity of the corresponding asset-level features.

Universe	Feature block	Mean historical cross-sectional s.d.	Latest mean $\lvert f\rvert$	Latest cross-sectional s.d.	Near-zero share (%)
$\mathcal{U}_{\mathrm{Eq}}$	Endpoint magnitude	0.0014	0.0013	6.93e-04	0.00
$\mathcal{U}_{\mathrm{Eq}}$	Wavelet energy	1.51e-05	1.24e-05	1.29e-05	0.00
$\mathcal{U}_{\mathrm{Eq}}$	Intercept magnitude	0.0210	0.0112	0.0222	8.33
$\mathcal{U}_{\mathrm{Eq}}$	Linear interaction magnitude	0.162	0.0793	0.265	83.33
$\mathcal{U}_{\mathrm{Eq}}$	Common forcing magnitude	0.0130	0.000	0.000	100.00
$\mathcal{U}_{\mathrm{Eq}}$	Quadratic interaction norm	0.0052	0.0131	0.0454	91.67
$\mathcal{U}_{\mathrm{Eq}}$	Sparsity ratio	0.0320	0.0252	0.0272	8.33
$\mathcal{U}_{\mathrm{Eq}}$	Instability score	6.111	0.551	1.044	0.00
$\mathcal{U}_{\mathrm{Eq}}$	Log residual variance	0.201	0.100	0.246	0.00
$\mathcal{U}_{\mathrm{ETF}}$	Endpoint magnitude	7.46e-04	8.63e-04	9.37e-04	0.00
$\mathcal{U}_{\mathrm{ETF}}$	Wavelet energy	3.38e-06	3.18e-06	4.66e-06	0.00
$\mathcal{U}_{\mathrm{ETF}}$	Intercept magnitude	0.0094	0.0033	0.0060	33.33
$\mathcal{U}_{\mathrm{ETF}}$	Linear interaction magnitude	0.0377	0.0166	0.0577	91.67
$\mathcal{U}_{\mathrm{ETF}}$	Common forcing magnitude	0.0043	0.000	0.000	100.00
$\mathcal{U}_{\mathrm{ETF}}$	Quadratic interaction norm	0.000	0.000	0.000	100.00
$\mathcal{U}_{\mathrm{ETF}}$	Sparsity ratio	0.0185	0.0164	0.0187	33.33
$\mathcal{U}_{\mathrm{ETF}}$	Instability score	0.847	0.553	1.044	0.00
$\mathcal{U}_{\mathrm{ETF}}$	Log residual variance	0.0734	0.0326	0.0786	0.00

Table 6. Cross-universe diagnostics for the extracted features. The table identifies the dominant blocks of cross-sectional heterogeneity and the latent principal scores that are most dispersed across assets.

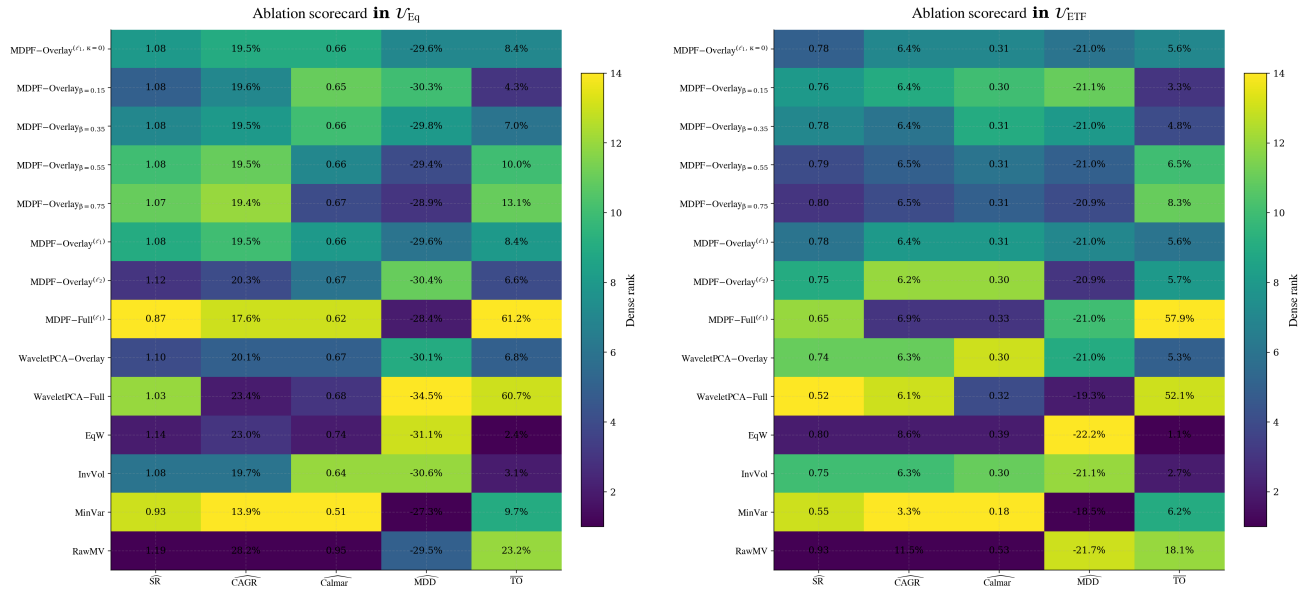
Universe	Most dispersed block	Second block	Dominant latent score	Inactive or near-zero blocks
$\mathcal{U}_{\mathrm{Eq}}$	Log residual variance (0.201)	Linear interaction magnitude (0.162)	z6 (2.091)	Common forcing magnitude
$\mathcal{U}_{\mathrm{ETF}}$	Log residual variance (0.073)	Linear interaction magnitude (0.038)	z4 (2.463)	Common forcing magnitude, Quadratic interaction norm

6.4. *Which architectural components mattered most?*

Table 7 and Figure 6 isolate the contribution of the principal architectural choices. The most robust empirical result is that overlay allocation dominates full allocation. In $\mathcal{U}_{\mathrm{Eq}}$, moving from the sparse full allocator to the sparse overlay increased the Sharpe ratio by 0.210 while reducing average ℓ_1 turnover by 1.055. In $\mathcal{U}_{\mathrm{ETF}}$, the corresponding gains were 0.137 in Sharpe and 1.046 in ℓ_1 turnover reduction. This is a large effect relative to the other architectural levers and indicates that the representation is more effective when it tilts a stable base portfolio than when it carries the entire capital-allocation burden.

The evidence on ODE regularization was more nuanced. In $\mathcal{U}_{\mathrm{Eq}}$, the ridge overlay outperformed the sparse overlay by 0.038 Sharpe, whereas in $\mathcal{U}_{\mathrm{ETF}}$, the sparse overlay outperformed the ridge overlay by 0.033. Hence, the preferred regularization regime depended on the universe. The instability penalty, by contrast, contributed almost nothing in either setting, as removing it changed Sharpe by essentially zero. Table 8 shows that parameter sensitivity was concentrated in a small subset of axes: The $\lambda_1^{(\mathrm{ODE})}$ grid and the mother wavelet family were materially relevant, whereas the turnover penalty and most other hyper-parameters were only mildly influential within the tested grids.

Figure 6 makes the same point visually. The strongest performance differences were aligned with the overlay-versus-full split and not with the instability penalty or modest hyper-parameter perturbations. For the feature-extraction interpretation, this matters because it suggests that the value of the architecture resides primarily in how the learned representation is inserted into the decision layer, rather than in the presence of every individual theoretical component. The advantages and limitations derived from the empirical results are summarized in Table 9.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure 6. Architecture-ablation heatmaps summarizing the relative performance of the principal model variants and benchmark strategies.

Table 7. Architecture-ablation summary. Positive ΔSharpe values favor the first-named specification in the comparison row. Negative ΔAvg. ℓ_1 turnover values indicate a reduction in average ℓ_1 turnover.

Universe	Comparison	ΔSharpe	ΔAvg. ℓ_1 turnover	Evidence
$\mathcal{U}_{\mathrm{Eq}}$	Best MDPF minus best benchmark	-0.0700	-0.332	MDPF-Ov (ℓ_2 ODE) vs RawMV
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1) minus MDPF-Full (ℓ_1)	0.210	-1.055	MDPF-Ov (ℓ_1 ODE) vs MDPF-Full (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_2) minus MDPF-Ov (ℓ_1)	0.0380	-0.0360	MDPF-Ov (ℓ_2 ODE) vs MDPF-Ov (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov ($\kappa = 0$) minus MDPF-Ov (ℓ_1)	0.000	0.000	MDPF-Ov ($\kappa = 0$) vs MDPF-Ov (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1) minus WaveletPCA-Ov	-0.0260	0.0320	MDPF-Ov (ℓ_1 ODE) vs WaveletPCA-Ov
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1 ODE): Sharpe change from 0 to 25 bps	-0.0140	–	Transaction-cost robustness
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Full (ℓ_1 ODE): Sharpe change from 0 to 25 bps	-0.0870	–	Transaction-cost robustness
$\mathcal{U}_{\mathrm{ETF}}$	Best MDPF minus best benchmark	-0.126	-0.196	MDPF-Ov ($\beta_{\mathrm{overlay}} = 0.75$) vs RawMV
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1) minus MDPF-Full (ℓ_1)	0.137	-1.046	MDPF-Ov (ℓ_1 ODE) vs MDPF-Full (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_2) minus MDPF-Ov (ℓ_1)	-0.0330	0.0020	MDPF-Ov (ℓ_2 ODE) vs MDPF-Ov (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov ($\kappa = 0$) minus MDPF-Ov (ℓ_1)	0.000	0.000	MDPF-Ov ($\kappa = 0$) vs MDPF-Ov (ℓ_1 ODE)
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1) minus WaveletPCA-Ov	0.0420	0.0060	MDPF-Ov (ℓ_1 ODE) vs WaveletPCA-Ov
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1 ODE): Sharpe change from 0 to 25 bps	-0.0200	–	Transaction-cost robustness
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Full (ℓ_1 ODE): Sharpe change from 0 to 25 bps	-0.156	–	Transaction-cost robustness

Table 8. Sensitivity summary for the main overlay specification. The Sharpe span reports the range of realized Sharpe ratios across the tested grid for each sensitivity axis.

Universe	Sensitivity axis	Best value or family	Best sharpe	Sharpe span	Assessment
$\mathcal{U}_{\text{Eq}}$	$\lambda_1^{(\text{ODE})}$	0.0	1.124	0.0850	Material
$\mathcal{U}_{\text{Eq}}$	$\lambda_2^{(\text{ODE})}$	0.002	1.088	0.0060	Mild
$\mathcal{U}_{\text{Eq}}$	ρ_{PCA}	0.9	1.088	0.0090	Mild
$\mathcal{U}_{\text{Eq}}$	γ	4.0	1.086	0.0170	Mild
$\mathcal{U}_{\text{Eq}}$	λ_{TO}	0.0005	1.082	0.000	Mild
$\mathcal{U}_{\text{Eq}}$	β_{overlay}	0.15	1.083	0.0070	Mild
$\mathcal{U}_{\text{Eq}}$	Mother wavelet family	db4	1.082	0.0430	Material
$\mathcal{U}_{\text{ETF}}$	$\lambda_1^{(\text{ODE})}$	0.0005	0.785	0.0340	Material
$\mathcal{U}_{\text{ETF}}$	$\lambda_2^{(\text{ODE})}$	0.005	0.791	0.0080	Mild
$\mathcal{U}_{\text{ETF}}$	ρ_{PCA}	0.8	0.805	0.0280	Mild
$\mathcal{U}_{\text{ETF}}$	γ	16.0	0.793	0.0120	Mild
$\mathcal{U}_{\text{ETF}}$	λ_{TO}	0.005	0.785	0.000	Mild
$\mathcal{U}_{\text{ETF}}$	β_{overlay}	0.75	0.802	0.0400	Material
$\mathcal{U}_{\text{ETF}}$	Mother wavelet family	db2	0.789	0.0610	Material

Table 9. Balanced assessment of the principal architectural components. The matrix summarizes the empirical advantages and limitations that emerged from the empirical evaluation framework.

Component	Empirical advantage	Empirical limitation
Wavelet multiscale decomposition	Produces stable cross-asset feature dispersion and universe-dependent but economically interpretable wavelet-family preferences.	Performance differences across wavelet families are moderate rather than transformative; the best family varies by universe.
ODE identification layer	Adds value in overlay form, but the preferred regularization differs by universe, suggesting the layer adapts to market structure.	Sparse ODE is not uniformly superior to ridge ODE, and some nonlinear blocks are weak or inactive, especially in ETFs.
Overlay allocation	Dominates the full-allocation version in both universes and sharply reduces turnover and transaction-cost sensitivity.	Even the best overlay does not surpass the strongest benchmark in either universe.
Instability penalty	Provides a conceptually appealing mechanism for penalising locally unstable dynamics.	Its empirical contribution is negligible in the present run.
Feature-extraction objective	The architecture produces heterogeneous asset-level feature blocks and well-separated latent scores, supporting its usefulness as a representation learner.	The resulting return stream remains benchmark-like, so the representation advantage does not always convert into benchmark dominance.

6.5. Synthesis of supplementary feature-extraction diagnostics

The supplementary diagnostics provide an empirical audit of what the MDPF representation measures and how it enters the portfolio decision. Table A4 decomposes the architecture into raw-return, DWT-only, DWT+PCA, DWT+ODE+PCA, nonlinear-block, forcing-block, no-PCA, overlay, and full-allocation variants. This decomposition is important because it separates the value of the feature representation from the value of the allocation interface. Table A5 then places the major return comparisons under sampling uncertainty using paired block-bootstrap Sharpe differences and Newey–West active-return tests, while Table A6 conditions the same evaluation on bull, normal, bear, high-volatility, low-volatility, drawdown, and shock states. Together, these tables measure whether the extracted features are robust across market environments or whether their usefulness is concentrated in a narrow average-sample condition.

The remaining supplementary tables examine the internal mechanics of the feature-extraction pipeline. Table A7 tests whether the multiscale state is materially affected by the mother wavelet, decomposition depth, thresholding rule, and threshold multiplier. Table A8 measures the empirical activity of the linear, quadratic, common-forcing, residual, and instability blocks, thereby identifying which parts of the local dynamic law are genuinely active. Table A9 links retained principal components back to their dominant feature blocks, and Table A10 connects those blocks to active tilts, alpha scores, next-period returns, and top-minus-bottom feature spreads. Finally, Tables A11 and A12 test how the extracted signal behaves when the overlay intensity, risk aversion, turnover penalty, base allocation, and trading-cost assumptions are varied. These diagnostics support the central interpretation of the paper: MDPF is most valuable as an interpretable multiscale feature-extraction layer for controlled overlays, while its stand-alone allocation performance remains limited by benchmark strength, estimation risk, and weak activation of some nonlinear terms.

7. Discussion

In this section, we interpret the empirical findings and clarify the practical value and limitations of the proposed methodology.

7.1. Interpretation of the empirical evidence

The empirical evidence supports a balanced interpretation of MDPF. The strongest contribution is representational rather than purely performance based. The wavelet–ODE–PCA stack generated heterogeneous asset-level states, separated latent principal-feature scores, and featured blocks in which residual-dynamic information and linear interaction structure carried a substantial empirical signal. Figures 4 and 5, together with Tables 5 and 6, show that the extracted representation is not equivalent to a simple raw-return summary. It contains structured information about scale-specific states, local dynamic interactions, and residual uncertainty.

Furthermore, the results do not support a claim of universal benchmark dominance. RawMV remained the strongest benchmark in both universes, and EqW was difficult to surpass in the equity universe. This limitation is important because it separates the value of the representation from the value of the final capital allocation. A feature map can be informative and interpretable without necessarily producing the highest out-of-sample Sharpe ratio once transaction costs, turnover, constraints, and

benchmark strength are considered. The active-gap figures further show that MDPF-Ov often behaves like a structured benchmark tilt rather than an orthogonal return stream. That behavior is consistent with the intended overlay interpretation: The model extracts state information and expresses it in a controlled way, rather than forcing an aggressive independent allocation.

The component evidence clarifies why the overlay design is the most credible interface. The ablation results in Table 7 and Appendix Table A4 indicate that full allocation tends to amplify estimation error and turnover, whereas the overlay formulation keeps the signal close to a diversified base portfolio. The statistical and regime diagnostics in Tables A5 and A6 further caution against interpreting small Sharpe differences as unconditional superiority. In this sense, the MDPF architecture is better understood as a disciplined representation-and-tilt system than as a complete replacement for traditional allocation rules.

7.2. Practical implications for portfolio construction

The practical use case for MDPF is as a transparent signal layer that can be placed on top of an existing allocation process. A practitioner could use EqW, InvVol, risk-parity, or a constrained mean–variance portfolio as the base allocation and then use MDPF to generate active tilts that reflect multiscale dynamic information. This is operationally attractive because the overlay format limits the size of the active position, reduces the likelihood that weak signals dominate the full portfolio, and permits the investor to inspect which feature blocks are driving the tilt.

The supplementary diagnostics are useful from an implementation perspective. Table A10 links extracted feature blocks to tilts, model scores, and realized return spreads, which helps determine whether the signal is economically connected to portfolio decisions. Table A11 shows how the overlay interface responds to changes in risk aversion, turnover penalty, base allocation, and overlay intensity. Table A12 then evaluates the same interpretation under trading-cost assumptions. These diagnostics suggest that the main advantage of MDPF is not that it always dominates simple benchmarks, but that it produces an auditable feature signal that can be integrated into traditional portfolio construction with explicit control over turnover and active risk.

7.3. Limitations and future research

Several limitations remain. First, benchmark dominance is not universal. The strongest raw mean–variance benchmark outperforms the proposed model in both universes, although it does so with materially higher turnover. Second, we use two controlled asset universes, which is appropriate for a transparent methodological evaluation but not sufficient to establish broad market generality. Third, the empirical contribution of the instability penalty is limited in this sample, and Appendix Table A8 shows that some nonlinear and forcing components are weak or unevenly active. This may reflect small cross-sectional dimension, strong regularization shrinkage, diversification smoothing in ETF returns, or the fact that nonlinear scale interactions appear only in specific regimes.

Therefore, researchers should focus on the dynamic-interaction layer and the portfolio interface. Richer cross-asset forcing terms, regime-dependent ODE dictionaries, adaptive regularization, and nonlinear selection rules may activate the dynamic blocks more consistently. Broader empirical validation across larger equity universes, international assets, and crisis subperiods would also clarify whether the feature representation generalizes beyond this sample. Finally, portfolio integration should continue to emphasize overlay, shrinkage, and transaction-cost-aware objectives, because this evidence

shows that these interfaces are more reliable than unrestricted allocation when feature signals are informative but noisy.

8. Conclusions

In this paper, we propose and evaluate a multiscale differential principal-feature architecture for portfolio optimization. The architecture transforms rolling returns into wavelet-based multiscale states, estimates local ODE signatures in that state space, compresses the resulting dynamic features through PCA, and translates the compressed representation into constrained portfolio overlays. The central contribution is an interpretable feature-extraction mechanism that links multiscale financial structure to portfolio construction.

The empirical evidence supports a cautious but useful conclusion. MDPF produces structured multiscale dynamic features, meaningful cross-sectional heterogeneity, and interpretable principal-feature directions. Its most convincing portfolio role is as an overlay layer, where the extracted signal can adjust a stable base allocation without replacing it. The method is not a universally benchmark-dominant allocator: RawMV and, in some cases, EqW remain difficult comparators. This limitation is not incidental; it defines the proper interpretation of the architecture. MDPF is valuable when the objective is to extract and diagnose dynamic multiscale information for controlled tilting, not when the objective is to guarantee unconditional superiority over every benchmark.

The results also reveal directions for improvement. The nonlinear and instability components require stronger regime-aware design, the regularization structure should become adaptive across universes, and the portfolio layer should incorporate transaction costs and active-risk budgets more directly. Researchers can extend the architecture to larger universes, richer cross-asset dynamic couplings, alternative decomposition schemes, and robust overlay designs. Based on this evidence, MDPF is best positioned as a transparent and extensible feature-extraction framework for dynamic portfolio optimization.

Use of AI tools declaration

The author declare they have used Large Language Models (LLMs) in the creation of this article. Specifically, OpenAI's ChatGPT (GPT-5.3, 2026 version) and Writefull for Overleaf were used. The AI tools were employed solely for language polishing, including grammar correction, improving sentence structure, enhancing clarity, and organizing the presentation of content. AI-assisted language refinements were applied across all sections of the manuscript. The author reviewed and approved all modifications.

Acknowledgments

The author would like to acknowledge the Deanship of Graduate Studies and Scientific Research, Taif University for funding this work.

Conflict of interest

The author declare there is no conflict of interest.

Author contributions

The author solely conceived the study, developed the methodology, implemented the computational framework, conducted the empirical analysis, interpreted the results, and wrote the manuscript.

References

1. A. A. Michis, Multiscale partial correlation clustering of stock market returns, *J. Risk Financial Manage.*, **15** (2022), 24. https://doi.org/10.3390/jrfm15010024
2. M. Sahabuddin, M. A. Islam, M. I. Tabash, S. Anagreh, R. Akter, M. M. Rahman, Co-movement, portfolio diversification, investors' behavior and psychology: Evidence from developed and emerging countries' stock markets, *J. Risk Financial Manage.*, **15** (2022), 319. https://doi.org/10.3390/jrfm15080319
3. S. Huang, L. Cao, R. Sun, T. Ma, S. Liu, Enhancing portfolio optimization: A two-stage approach with deep learning and portfolio optimization, *Mathematics*, **12** (2024), 3376. https://doi.org/10.3390/math12213376
4. R. Yan, J. Jin, K. Han, Reinforcement learning for deep portfolio optimization, *Electron. Res. Arch.*, **32** (2024), 5176–5200. https://doi.org/10.3934/era.2024239
5. W. Zhang, G. Li, A review of end-to-end decision optimization research: An architectural perspective, *Algorithms*, **19** (2026), 86. https://doi.org/10.3390/a19010086
6. A. H. Alenezy, M. T. Ismail, S. A. Wadi, J. J. Jaber, Predicting stock market volatility using MODWT with HYFIS and FS.HGD models, *Risks*, **11** (2023), 121. https://doi.org/10.3390/risks11070121
7. O. Panazan, C. Gheorghe, Impact of geopolitical risk on G7 financial markets: A comparative wavelet analysis between 2014 and 2022, *Mathematics*, **12** (2024), 370. https://doi.org/10.3390/math12030370
8. L. Doroshenko, L. Mastroeni, A. Mazzoccoli, Wavelet and deep learning framework for predicting commodity prices under economic and financial uncertainty, *Mathematics*, **13** (2025), 1346. https://doi.org/10.3390/math13081346
9. E. Guariglia, R. C. Guido, G. J. P. Dalalana, From wavelet analysis to fractional calculus: A review, *Mathematics*, **11** (2023), 1606. https://doi.org/10.3390/math11071606
10. P. K. Karn, I. Ardekani, W. H. Abdulla, Generalized framework for liquid neural network upon sequential and non-sequential tasks, *Mathematics*, **12** (2024), 2525. https://doi.org/10.3390/math12162525
11. A. Papana, Detecting nonlinear interactions in complex systems: Application in financial markets, *Entropy*, **25** (2023), 370. https://doi.org/10.3390/e25020370
12. A. Moujahid, F. Vadillo, Analysis of noise on ordinary and fractional-order financial systems, *Fractal Fract.*, **9** (2025), 316. https://doi.org/10.3390/fractalfract9050316
13. A. Aloui, L. Diabi, O. Kahouli, A. Ouannas, L. El Amraoui, M. Ayari, Dynamic analysis of a fractional-order economic model: Chaos and control, *Fractal Fract.*, **9** (2025), 724. https://doi.org/10.3390/fractalfract9110724

14. A. Oprea, The use of principal component analysis (PCA) in building yield curve scenarios and identifying relative-value trading opportunities on the Romanian government bond market, *J. Risk Financial Manage.*, **15** (2022), 247. https://doi.org/10.3390/jrfm15060247
15. J. Mai, S. Zhang, H. Zhao, L. Pan, Factor investment or feature selection analysis, *Mathematics*, **13** (2025), 9. https://doi.org/10.3390/math13010009
16. Z. Liu, Forecasting stock prices based on multivariable fuzzy time series, *AIMS Math.*, **8** (2023), 12778–12792. https://doi.org/10.3934/math.2023643
17. P. L. Seabe, C. R. B. Moutsinga, E. Pindza, Optimizing cryptocurrency returns: A quantitative study on factor-based investing, *Mathematics*, **12** (2024), 1351. https://doi.org/10.3390/math12091351
18. M. H. Alkhudaydi, Y. M. Althobaity, Graph aware adaptive tracking-error optimization with wavelet-principal component analysis features and proportional-integral control (GATE-WPCA-PI), *AIMS Math.*, **11** (2026), 3647–3702. https://doi.org/10.3934/math.2026149
19. A. S. AlRashedy, H. I. Mathkour, Label-driven optimization of trading models across indices and stocks: Maximizing percentage profitability, *Mathematics*, **13** (2025), 3889. https://doi.org/10.3390/math13233889
20. F. Espiga-Fernandez, A. Garcia-Sanchez, J. Ordieres-Meré, A systematic approach to portfolio optimization: A comparative study of reinforcement learning agents, market signals, and investment horizons, *Algorithms*, **17** (2024), 570. https://doi.org/10.3390/a17120570
21. C. I. Giannikos, H. Guirguis, A. Kakolyris, T. K. Suen, When to hedge downside risk, *Risks*, **12** (2024), 42. https://doi.org/10.3390/risks12020042
22. E. Moyoweshumba, M. Seitshiro, Leveraging Markowitz, random forest, and XGBoost for optimal diversification of South African stock portfolios, *Data Sci. Finance Econ.*, **5** (2025), 205–233. https://doi.org/10.3934/DSFE.2025010

A. Supplementary empirical diagnostics

This appendix reports supplementary empirical diagnostics that clarify how the MDPF architecture behaves as a multiscale dynamic feature-extraction layer for controlled portfolio overlays. The tables below do not replace the major performance figures and tables; rather, they decompose the feature pipeline, quantify statistical uncertainty, examine regime dependence, and connect the extracted feature blocks to allocation decisions. All diagnostics use the same fixed sample window as the main manuscript, with adjusted prices requested from 1 January 2014, usable daily returns beginning on 3 January 2014, and an inclusive endpoint of 6 March 2026.

A.1. Sample window and experiment organization

Table A1 summarizes the role of each supplementary diagnostic block. Tables A2 and A3 document the fixed sample window and the high-level experiment registry used to organize the supplementary outputs.

Table A1. Supplementary diagnostic map for the empirical evidence reported in Appendix A. The table links each appendix block to the methodological question it evaluates and identifies the corresponding evidence.

Appendix	Diagnostic block	Methodological question addressed	Evidence
A.1	Sample window and experiment registry	Are the supplementary analyses organized on a reproducible and internally consistent empirical window?	Tables A2 and A3
A.2	Architectural ablation	Which parts of the DWT–ODE–PCA feature stack and allocation interface drive the observed portfolio behavior?	Table A4
A.3	Statistical tests	Are benchmark-relative performance gaps robust to sampling uncertainty?	Table A5
A.4	Regime robustness	Does the extracted representation remain informative across normal, adverse, volatile, and stress states?	Table A6
A.5	Wavelet sensitivity	How sensitive is the multiscale state to wavelet family, decomposition level, and thresholding?	Table A7
A.6	ODE diagnostics	Which ODE blocks are empirically active, and where do nonlinear or instability terms remain weak?	Table A8
A.7	PCA interpretability	Which feature blocks dominate the retained principal-feature directions?	Table A9
A.8	Feature-to-allocation link	Do the extracted feature blocks connect to active tilts, alpha scores, and realized return spreads?	Table A10
A.9	Optimization and implementation robustness	How does the feature signal behave under overlay, risk, turnover, base-portfolio, and trading-cost changes?	Tables A11 and A12

Notes: This map is a navigation aid for the supplementary diagnostics. Each block reports empirical evidence that complements the main performance tables and figures.

Table A2. Fixed sample-window registry for the supplementary empirical diagnostics. The table reports the requested price window, the first usable return date, the out-of-sample start date, and the fixed endpoint used in the appendix analyses.

Universe	Price request	Price start	Return start	OOS start	End date	Daily obs.	Window
$\mathcal{U}_{\mathrm{Eq}}$	2014-01-01	2014-01-02	2014-01-03	2017-02-01	2026-03-06	2286	Fixed
$\mathcal{U}_{\mathrm{ETF}}$	2014-01-01	2014-01-02	2014-01-03	2017-02-01	2026-03-06	2286	Fixed

Notes: OOS denotes out-of-sample. Daily obs. denotes the number of daily portfolio-return observations. Returns begin on 3 January 2014 because they are first differences of adjusted prices and 1 January 2014 was not a trading day.

Table A3. Unified experiment registry for the supplementary diagnostic runs. The registry summarizes the strategy groups evaluated in each universe under the fixed sample window.

Universe	Experiment group	Specifications	Best specification within group	Best SR	Daily obs.
$\mathcal{U}_{\mathrm{Eq}}$	Core strategies and benchmarks	9	RawMV	1.187	2286
$\mathcal{U}_{\mathrm{Eq}}$	Overlay-strength and instability variants	5	MDPF-Ov ($\beta_{\mathrm{overlay}} = 0.15$)	1.082	2286
$\mathcal{U}_{\mathrm{ETF}}$	Core strategies and benchmarks	9	RawMV	0.927	2286
$\mathcal{U}_{\mathrm{ETF}}$	Overlay-strength and instability variants	5	MDPF-Ov ($\beta_{\mathrm{overlay}} = 0.75$)	0.805	2286

Notes: SR denotes annualized Sharpe ratio. Daily obs. denotes the number of daily portfolio-return observations used in the fixed out-of-sample window.

A.2. Systematic architectural ablation

Table A4 decomposes the individual contributions of DWT, ODE identification, PCA, nonlinear terms, forcing terms, and the overlay/full allocation interface.

Table A4. Systematic architectural ablation of the feature-extraction and allocation components. The table decomposes the MDPF architecture into raw-return, DWT, ODE, PCA, nonlinear, forcing, and allocation-interface variants.

Universe	Specification	CAGR (%)	Vol. (%)	SR	MDD (%)	Avg. TN	Eff. bets
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1 ODE)	19.49	18.03	1.078	-29.59	0.084	10.446
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_2 ODE)	20.31	18.02	1.117	-30.39	0.066	10.721
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Full (ℓ_1 ODE)	17.55	21.25	0.867	-28.39	0.612	3.900
$\mathcal{U}_{\mathrm{Eq}}$	Raw-PCA-Ov	19.74	18.09	1.087	-30.64	0.030	11.045
$\mathcal{U}_{\mathrm{Eq}}$	DWT-Ov	19.60	18.06	1.082	-30.04	0.067	10.772
$\mathcal{U}_{\mathrm{Eq}}$	DWT+PCA-Ov	20.08	18.11	1.102	-30.14	0.061	10.879
$\mathcal{U}_{\mathrm{Eq}}$	DWT+linear ODE+PCA-Ov	19.72	17.92	1.094	-29.63	0.084	10.588
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov without quadratic block	19.47	18.08	1.075	-29.65	0.089	10.464
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov without forcing block	19.73	17.89	1.096	-29.45	0.083	10.569
$\mathcal{U}_{\mathrm{Eq}}$	Full ODE-Ov without PCA	18.82	17.93	1.052	-30.62	0.074	10.513
$\mathcal{U}_{\mathrm{Eq}}$	EqW	23.02	19.98	1.138	-31.09	0.024	12.000
$\mathcal{U}_{\mathrm{Eq}}$	RawMV	28.17	23.20	1.187	-29.51	0.232	3.816
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1 ODE)	6.45	8.39	0.787	-21.00	0.056	8.878
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_2 ODE)	6.27	8.55	0.754	-20.92	0.057	8.875
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Full (ℓ_1 ODE)	6.83	11.26	0.643	-21.05	0.579	3.779
$\mathcal{U}_{\mathrm{ETF}}$	Raw-PCA-Ov	6.39	8.73	0.753	-21.08	0.026	9.119
$\mathcal{U}_{\mathrm{ETF}}$	DWT-Ov	6.13	8.59	0.736	-21.22	0.048	8.953
$\mathcal{U}_{\mathrm{ETF}}$	DWT+PCA-Ov	6.25	8.66	0.744	-21.06	0.047	8.967
$\mathcal{U}_{\mathrm{ETF}}$	DWT+linear ODE+PCA-Ov	6.45	8.41	0.786	-21.04	0.059	8.874
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov without quadratic block	6.43	8.38	0.785	-21.00	0.056	8.878
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov without forcing block	6.45	8.41	0.786	-21.05	0.058	8.883
$\mathcal{U}_{\mathrm{ETF}}$	Full ODE-Ov without PCA	6.14	8.35	0.756	-21.01	0.053	8.759
$\mathcal{U}_{\mathrm{ETF}}$	EqW	8.57	10.97	0.805	-22.18	0.011	12.000
$\mathcal{U}_{\mathrm{ETF}}$	RawMV	11.47	12.57	0.927	-21.68	0.181	3.716

Notes: SR denotes annualized Sharpe ratio; Vol. denotes annualized volatility; MDD denotes maximum drawdown; Avg. TN denotes average traded notional, equal to one half of the ℓ_1 turnover used in the main performance table; Eff. bets denotes the inverse Herfindahl effective number of bets. The table is diagnostic and complements the main performance table.

A.3. Statistical significance of benchmark differences

Table A5 adds paired block-bootstrap Sharpe-difference intervals and Newey–West active-return tests. These tests quantify the extent to which benchmark gaps may reflect sampling variation rather than persistent performance differences.

Table A5. Statistical reliability tests for the main MDPF-Ov (ℓ_1 ODE) against comparator return streams. The table reports paired block-bootstrap Sharpe differences and Newey–West active-return tests.

Universe	Comparator	ΔSR	95% CI	Boot. p	Active mean (%)	NW t	NW p
$\mathcal{U}_{\mathrm{Eq}}$	WaveletPCA-Ov	-0.026	[-0.073, 0.011]	0.177	-0.56	-1.357	0.175
$\mathcal{U}_{\mathrm{Eq}}$	EqW	-0.060	[-0.182, 0.068]	0.378	-3.29	-2.457	0.014
$\mathcal{U}_{\mathrm{Eq}}$	InvVol	-0.004	[-0.060, 0.043]	0.791	-0.15	-0.305	0.761
$\mathcal{U}_{\mathrm{Eq}}$	MinVar	0.149	[-0.126, 0.443]	0.283	5.23	2.248	0.025
$\mathcal{U}_{\mathrm{Eq}}$	RawMV	-0.108	[-0.441, 0.289]	0.644	-8.09	-1.951	0.051
$\mathcal{U}_{\mathrm{ETF}}$	WaveletPCA-Ov	0.042	[-0.019, 0.093]	0.188	0.12	0.489	0.625
$\mathcal{U}_{\mathrm{ETF}}$	EqW	-0.018	[-0.173, 0.131]	0.799	-2.23	-2.102	0.036
$\mathcal{U}_{\mathrm{ETF}}$	InvVol	0.037	[-0.037, 0.106]	0.432	0.04	0.147	0.883
$\mathcal{U}_{\mathrm{ETF}}$	MinVar	0.234	[0.035, 0.465]	0.019	3.08	2.682	0.007
$\mathcal{U}_{\mathrm{ETF}}$	RawMV	-0.140	[-0.571, 0.284]	0.521	-5.05	-1.978	0.048

Notes: ΔSR is the Sharpe ratio of the main MDPF-Ov (ℓ_1 ODE) minus the comparator Sharpe ratio. Bootstrap intervals are paired block-bootstrap intervals for ΔSR, and Boot. p is the corresponding two-sided bootstrap p-value. Active mean is the annualized mean return difference in percentage points. NW denotes Newey–West tests of active-return means.

A.4. Regime and stress robustness

Table A6 reports regime-specific Sharpe ratios for the main MDPF-Ov (ℓ_1 ODE), EqW, and RawMV under bull, normal, bear, volatility, drawdown, and daily-shock states.

Table A6. Regime and stress robustness of the main MDPF-Ov (ℓ_1 ODE) relative to EqW and RawMV. The table conditions Sharpe ratios on trend, volatility, drawdown, and shock states.

Universe	Regime	n	MDPF-Ov SR	EqW SR	RawMV SR	ΔSR vs EqW	ΔSR vs RawMV
$\mathcal{U}_{\text{Eq}}$	Bear	756	-0.237	-0.216	-0.159	-0.021	-0.077
$\mathcal{U}_{\text{Eq}}$	Normal	641	2.301	2.236	1.672	0.065	0.629
$\mathcal{U}_{\text{Eq}}$	Bull	889	2.549	2.854	2.283	-0.305	0.267
$\mathcal{U}_{\text{Eq}}$	HighVol	645	0.678	0.748	0.805	-0.071	-0.127
$\mathcal{U}_{\text{Eq}}$	LowVol	629	2.640	2.460	1.975	0.179	0.665
$\mathcal{U}_{\text{Eq}}$	Drawdown10	375	-1.183	-1.108	-1.054	-0.075	-0.129
$\mathcal{U}_{\text{Eq}}$	WorstDD	574	-0.637	-0.607	-0.633	-0.030	-0.004
$\mathcal{U}_{\text{Eq}}$	Daily shock 5%	128	-29.783	-32.814	-21.755	3.032	-8.027
$\mathcal{U}_{\text{ETF}}$	Bear	640	-0.675	-0.593	-0.755	-0.082	0.079
$\mathcal{U}_{\text{ETF}}$	Normal	768	1.923	1.651	1.193	0.272	0.730
$\mathcal{U}_{\text{ETF}}$	Bull	878	2.065	2.312	2.156	-0.247	-0.091
$\mathcal{U}_{\text{ETF}}$	HighVol	659	0.769	0.912	0.713	-0.143	0.056
$\mathcal{U}_{\text{ETF}}$	LowVol	562	1.806	1.420	1.414	0.386	0.391
$\mathcal{U}_{\text{ETF}}$	Drawdown10	364	-0.513	-0.711	-0.616	0.197	0.103
$\mathcal{U}_{\text{ETF}}$	WorstDD	501	-0.520	-0.530	-0.689	0.010	0.169
$\mathcal{U}_{\text{ETF}}$	Daily shock 5%	132	-27.505	-27.626	-21.710	0.122	-5.795

Notes: n denotes the number of daily observations in the regime. ΔSR columns report MDPF-Ov SR minus the comparator SR within the same regime. The daily shock filter isolates the largest 5% absolute daily moves of the universe-level equal-weight proxy; Drawdown10 flags dates when the proxy drawdown is at least 10%.

A.5. Wavelet family, decomposition-level, and threshold sensitivity

Table A7 evaluates whether the multiscale representation is robust to changes in the mother wavelet, decomposition depth, threshold rule, and threshold multiplier.

Table A7. Wavelet-family, decomposition-level, and threshold-sensitivity summary for MDPF-Ov. The grid evaluates whether the multiscale state is robust to wavelet and threshold design choices.

Universe	Grid size	Baseline setting	Base SR	Best setting	Best SR	Worst SR	Median SR	SR range
$\mathcal{U}_{\text{Eq}}$	140	db4, J=3, universal, m=1.0	1.078	haar, J=4, scale_adaptive, m=1.0	1.126	1.001	1.064	0.125
$\mathcal{U}_{\text{ETF}}$	140	db4, J=3, universal, m=1.0	0.787	haar, J=4, scale_adaptive, m=1.5	0.827	0.668	0.760	0.159

Notes: J denotes the wavelet decomposition level and m denotes the threshold multiplier. SR denotes annualized Sharpe ratio, and SR range is the difference between the best and worst Sharpe ratio over the evaluated grid. The baseline setting is the specification used in the main MDPF-Ov analysis.

A.6. ODE nonlinear-block and instability diagnostics

Table A8 examines the empirical activity of the ODE coefficient blocks and supports the limitation discussion regarding weak nonlinear-block activation.

Table A8. ODE block-activity diagnostics for nonlinear, forcing, and instability components. The table summarizes whether each block is empirically active across rolling windows and assets.

Diagnostic	Interpretation	$\mathcal{U}_{\mathrm{Eq}}$ mean	$\mathcal{U}_{\mathrm{Eq}}$ median	$\mathcal{U}_{\mathrm{ETF}}$ mean	$\mathcal{U}_{\mathrm{ETF}}$ median
Linear norm	Linear cross-scale propagation strength	0.744	0.000	0.151	0.000
Quadratic norm	Nonlinear cross-scale interaction strength	0.005	0.000	0.000	0.000
Forcing norm	Common multiscale market-pressure strength	0.091	0.000	0.024	0.000
Active linear share	Share of active linear coefficients	0.056	0.000	0.018	0.000
Active quadratic share	Share of active quadratic coefficients	0.000	0.000	0.000	0.000
Active forcing share	Share of active forcing coefficients	0.021	0.000	0.007	0.000
Full-vs-linear fit gain	Incremental full-versus-linear ODE fit gain	0.000	0.000	0.000	0.000
Raw instability score	Raw local instability score	0.392	0.000	0.091	0.000

Notes: Coefficient norms measure block magnitude over assets and rolling dates; active shares measure the fraction of nonzero coefficients after regularization; full-vs-linear fit gain compares the full ODE dictionary with its linear-only counterpart. Near-zero quadratic activity and small fit gains indicate that nonlinear interactions are weak in this empirical window.

A.7. PCA interpretability of principal features

Table A9 interprets retained principal components by decomposing their loading contributions into economically meaningful feature blocks.

Table A9. Economic interpretation of retained principal-feature directions through PCA loading-block contributions. The table identifies which feature block most often dominates each retained component.

Universe	PC	Dominant block	Economic interpretation	EV (%)	Loading share (%)	Date share (%)
$\mathcal{U}_{\mathrm{Eq}}$	PC1	linear	linear cross-scale propagation dynamics	25.27	29.48	95.83
$\mathcal{U}_{\mathrm{Eq}}$	PC2	linear	linear cross-scale propagation dynamics	12.38	47.11	100.00
$\mathcal{U}_{\mathrm{Eq}}$	PC3	linear	linear cross-scale propagation dynamics	9.79	53.70	100.00
$\mathcal{U}_{\mathrm{Eq}}$	PC4	linear	linear cross-scale propagation dynamics	7.77	43.06	91.67
$\mathcal{U}_{\mathrm{Eq}}$	PC5	linear	linear cross-scale propagation dynamics	6.63	46.77	66.67
$\mathcal{U}_{\mathrm{ETF}}$	PC1	linear	linear cross-scale propagation dynamics	26.97	24.65	62.50
$\mathcal{U}_{\mathrm{ETF}}$	PC2	linear	linear cross-scale propagation dynamics	14.70	34.10	41.67
$\mathcal{U}_{\mathrm{ETF}}$	PC3	linear	linear cross-scale propagation dynamics	11.36	28.70	37.50
$\mathcal{U}_{\mathrm{ETF}}$	PC4	intercept	baseline local drift in the ODE state equation	7.62	39.13	33.33
$\mathcal{U}_{\mathrm{ETF}}$	PC5	energy	scale-specific volatility/energy concentration	6.92	45.33	41.67

Notes: PC denotes a retained principal component. EV is the explained-variance share of the principal component. Loading share is the contribution of the dominant feature block to the component loading. Date share is the share of rolling dates for which the reported block is dominant.

A.8. Feature-to-allocation evidence

Table A10 links extracted feature blocks to active tilts, alpha scores, next-period returns, and top-minus-bottom feature spreads.

Table A10. Feature-to-allocation and feature-to-return diagnostics for the extracted MDPF blocks. The table links feature-block summaries to active tilts, model scores, next-period returns, and realized top-minus-bottom spreads.

Universe	Feature block	Tilt corr.	|tilt| corr.	Alpha corr.	Next-ret. corr.	Top-bottom (%)	Dates
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{\lvert x_t^\star \rvert}$	-0.122	0.213	0.153	-0.023	0.88	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{e_t}$	-0.122	0.352	0.240	0.107	1.38	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{\lvert\alpha\rvert}$	-0.183	0.356	0.190	0.084	1.27	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{\lvert B_{\mathrm{lin}}\rvert}$	-0.316	0.451	0.216	0.097	0.28	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{\lvert B_{\mathrm{for}}\rvert}$	-0.317	0.447	0.095	0.122	0.28	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{\lVert B_{\mathrm{quad}}\rVert}$	-0.226	0.274	0.063	0.094	0.02	110
$\mathcal{U}_{\mathrm{Eq}}$	$\overline{s}$	-0.199	0.439	0.251	0.098	0.85	110
$\mathcal{U}_{\mathrm{Eq}}$	κ_t	-0.365	0.067	0.235	0.071	0.41	110
$\mathcal{U}_{\mathrm{Eq}}$	$\log(1+\sigma^2_{\varepsilon,t})$	-0.216	0.428	0.232	0.099	0.86	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{\lvert x_t^\star \rvert}$	-0.126	0.228	0.042	-0.009	-0.11	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{e_t}$	-0.121	0.343	0.153	0.142	0.81	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{\lvert\alpha\rvert}$	-0.080	0.321	0.251	0.134	0.78	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{\lvert B_{\mathrm{lin}}\rvert}$	-0.200	0.373	0.101	0.196	0.32	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{\lvert B_{\mathrm{for}}\rvert}$	-0.181	0.345	0.074	0.162	0.20	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{\lVert B_{\mathrm{quad}}\rVert}$	-0.084	0.140	0.082	0.122	0.08	110
$\mathcal{U}_{\mathrm{ETF}}$	$\overline{s}$	-0.088	0.373	0.276	0.139	0.44	110
$\mathcal{U}_{\mathrm{ETF}}$	κ_t	-0.195	-0.104	0.136	-0.019	0.21	110
$\mathcal{U}_{\mathrm{ETF}}$	$\log(1+\sigma^2_{\varepsilon,t})$	-0.141	0.364	0.179	0.131	0.51	110

Notes: Correlations are cross-sectional Spearman averages over rebalance dates. Tilt corr. uses the signed active tilt; |tilt| corr. uses the absolute active tilt; Alpha corr. links the feature block to the model score; Next-ret. corr. links it to next-period realized returns. Top-bottom is the mean next-period return spread, in percentage points, between the highest- and lowest-feature groups. Dates is the number of rebalance dates used in the cross-sectional averaging.

A.9. Optimization refinement and practical transaction-cost robustness

Table A11 reports an allocation-interface refinement grid. Table A12 adds a practical transaction-cost robustness check.

Table A11. Optimization-refinement grid for the MDPF-Ov allocation interface. The table varies overlay intensity, risk aversion, turnover penalty, instability penalty, and base portfolio to evaluate how the extracted feature signal is translated into portfolio weights.

Universe	Refinement axis	Grid size	Best setting	SR	CAGR (%)	MDD (%)	Avg. TN
$\mathcal{U}_{\mathrm{Eq}}$	Overlay intensity	5	$\beta_{\mathrm{overlay}} = 0.15$	1.082	19.61	-30.29	0.043
$\mathcal{U}_{\mathrm{Eq}}$	Risk aversion	4	$\gamma = 4$	1.083	19.63	-29.64	0.080
$\mathcal{U}_{\mathrm{Eq}}$	Turnover penalty	4	$\lambda_{\mathrm{TO}} = 0.005$	1.078	19.49	-29.59	0.084
$\mathcal{U}_{\mathrm{Eq}}$	Instability penalty	4	$\lambda_{\kappa} = 0.10$	1.078	19.48	-29.58	0.084
$\mathcal{U}_{\mathrm{Eq}}$	Base portfolio	3	EqW	1.132	22.46	-29.93	0.083
$\mathcal{U}_{\mathrm{ETF}}$	Overlay intensity	5	$\beta_{\mathrm{overlay}} = 1.00$	0.815	6.57	-20.85	0.106
$\mathcal{U}_{\mathrm{ETF}}$	Risk aversion	4	$\gamma = 16$	0.791	6.43	-20.96	0.064
$\mathcal{U}_{\mathrm{ETF}}$	Turnover penalty	4	$\lambda_{\mathrm{TO}} = 0.005$	0.787	6.45	-21.00	0.056
$\mathcal{U}_{\mathrm{ETF}}$	Instability penalty	4	$\lambda_{\kappa} = 0.00$	0.787	6.45	-21.00	0.056
$\mathcal{U}_{\mathrm{ETF}}$	Base portfolio	3	EqW	0.841	8.39	-21.67	0.054

Notes: SR denotes annualized Sharpe ratio; MDD denotes maximum drawdown; Avg. TN denotes average traded notional, equal to one half of ℓ_1 turnover. Grid size is the number of settings evaluated along the corresponding refinement axis. The grid is diagnostic and is used to interpret the allocation interface rather than to replace the main performance comparison.

Table A12. Practical transaction-cost robustness at zero and 25 basis-point trading-cost assumptions. The table compares how Sharpe ratios and risk metrics change when trading frictions are increased.

Universe	Specification	SR at 0 bps	SR at 25 bps	SR change	CAGR at 25 bps (%)	MDD at 25 bps (%)
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_1 ODE)	1.084	1.070	-0.014	19.30	-29.60
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Ov (ℓ_2 ODE)	1.121	1.110	-0.011	20.16	-30.40
$\mathcal{U}_{\mathrm{Eq}}$	MDPF-Full (ℓ_1 ODE)	0.902	0.815	-0.087	16.25	-28.47
$\mathcal{U}_{\mathrm{Eq}}$	EqW	1.139	1.136	-0.004	22.97	-31.09
$\mathcal{U}_{\mathrm{Eq}}$	RawMV	1.199	1.168	-0.030	27.63	-29.52
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_1 ODE)	0.795	0.775	-0.020	6.34	-21.06
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Ov (ℓ_2 ODE)	0.762	0.742	-0.020	6.16	-20.96
$\mathcal{U}_{\mathrm{ETF}}$	MDPF-Full (ℓ_1 ODE)	0.705	0.549	-0.155	5.71	-21.93
$\mathcal{U}_{\mathrm{ETF}}$	EqW	0.806	0.803	-0.003	8.55	-22.18
$\mathcal{U}_{\mathrm{ETF}}$	RawMV	0.944	0.901	-0.044	11.10	-21.98

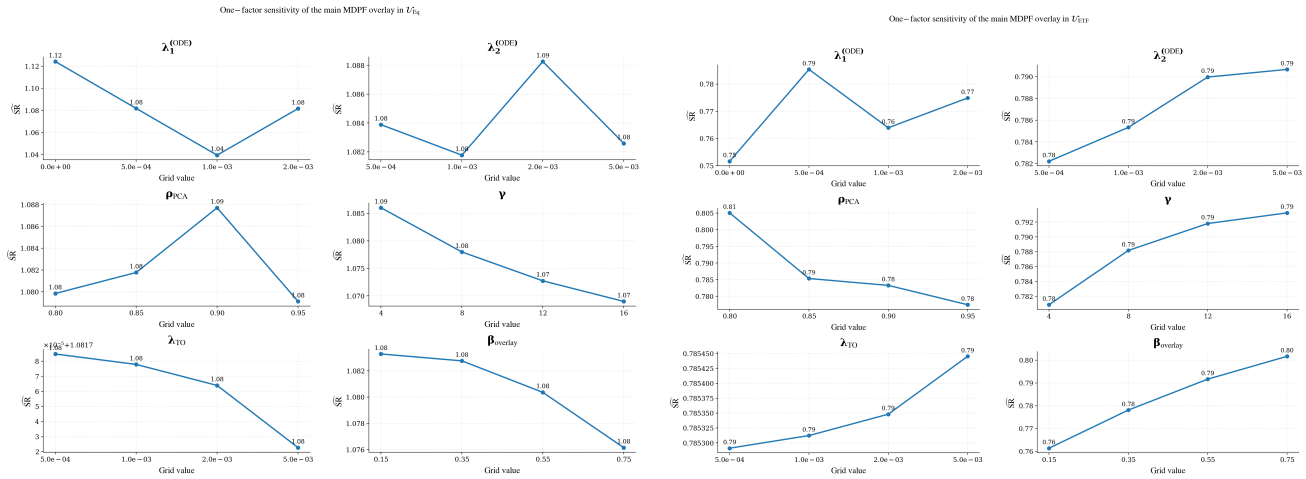
Notes: SR denotes annualized Sharpe ratio; CAGR denotes compound annual growth rate; MDD denotes maximum drawdown. SR change is computed as SR at 25 bps minus SR at 0 bps. Transaction costs are applied to traded notional at each rebalance.

B. Sensitivity and specification robustness

The sensitivity figures provide a useful complement to the main results because they separate structural conclusions from incidental parameter choices. Figure B1 shows that the one-factor parameter sweeps were not uniformly influential. In $\mathcal{U}_{\mathrm{Eq}}$, the $\lambda_1^{(\mathrm{ODE})}$ axis generated the widest Sharpe span among

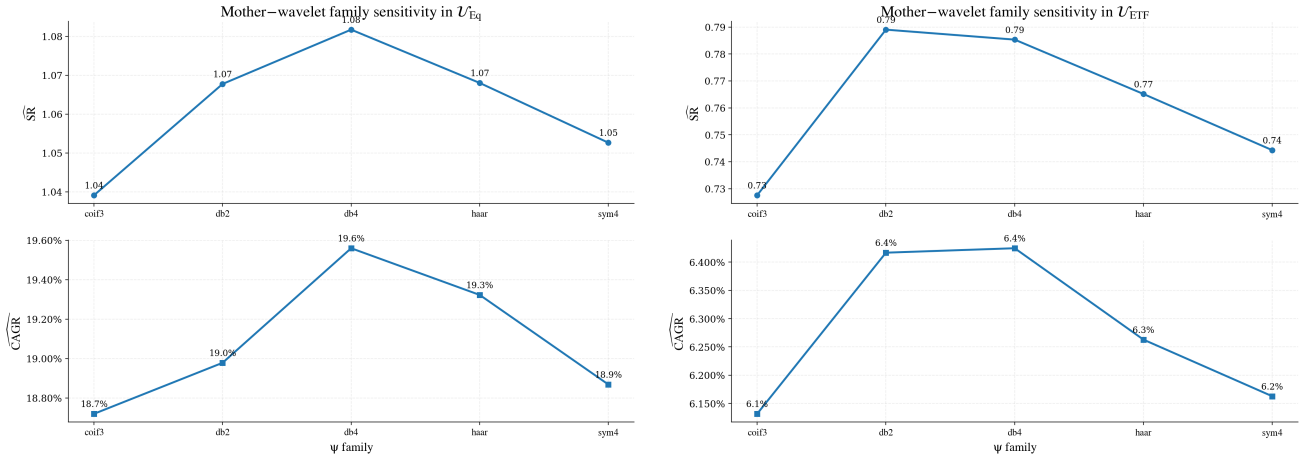
the scalar parameters, whereas in $\mathcal{U}_{\text{ETF}}$ the overlay blend and the explained-variance threshold were somewhat more consequential. Table B1 and Figure B2 show that the mother wavelet family mattered in both universes, although the preferred family differed: The strongest family was *db4* in $\mathcal{U}_{\text{Eq}}$ and *db2* in $\mathcal{U}_{\text{ETF}}$. To provide context for these designations, Table B2 summarizes the mathematical properties and typical financial applications of the wavelet families evaluated in this analysis.

Figures B3 and B4 reinforce the central role of overlay design. The overlay-strength sweep indicated a mild preference for smaller overlay intensity in $\mathcal{U}_{\text{Eq}}$ and a stronger preference for larger overlay intensity in $\mathcal{U}_{\text{ETF}}$. The transaction-cost robustness plots show that the sparse overlay deteriorated much slower than the sparse full allocator as trading frictions increased. Finally, Figure B5 and Table B3 reveal that bear-market Sharpe ratios were negative for all selected strategies, but the full allocator deteriorated less severely than some overlay variants during adverse regimes. This nuance is useful because it shows that the overlay advantage is predominantly a whole-sample efficiency result rather than a universal state-by-state superiority.



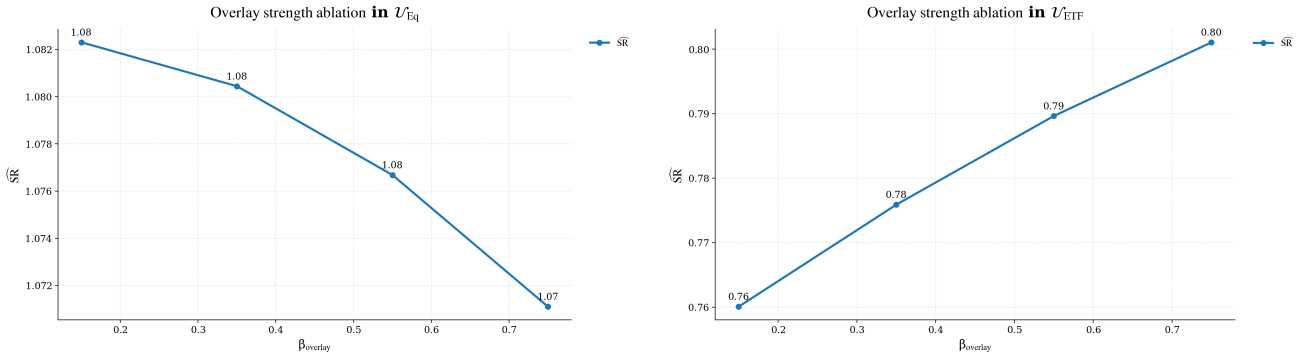
(a) $\mathcal{U}_{\text{Eq}}$. **(b)** $\mathcal{U}_{\text{ETF}}$.

Figure B1. One-factor parameter sensitivity of the main MDPF-Ov (ℓ_1 ODE).



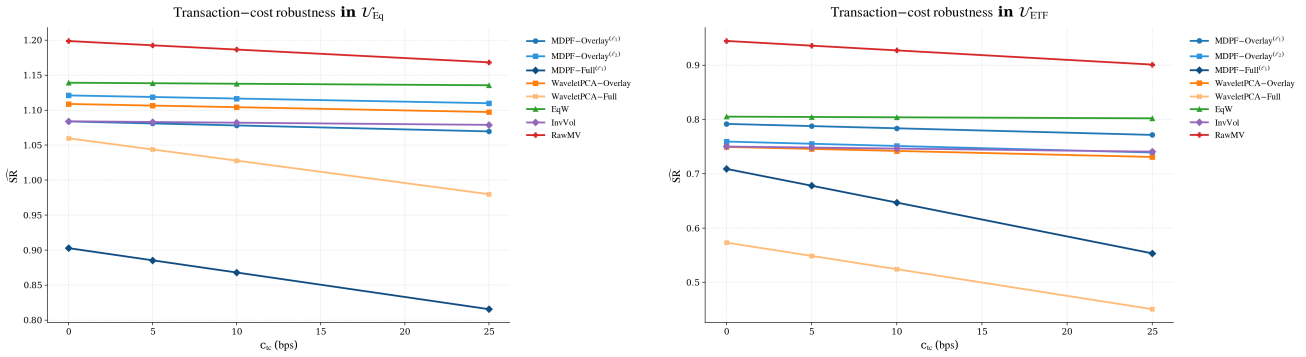
(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure B2. Mother-wavelet sensitivity of the main MDPF-Ov (ℓ_1 ODE).



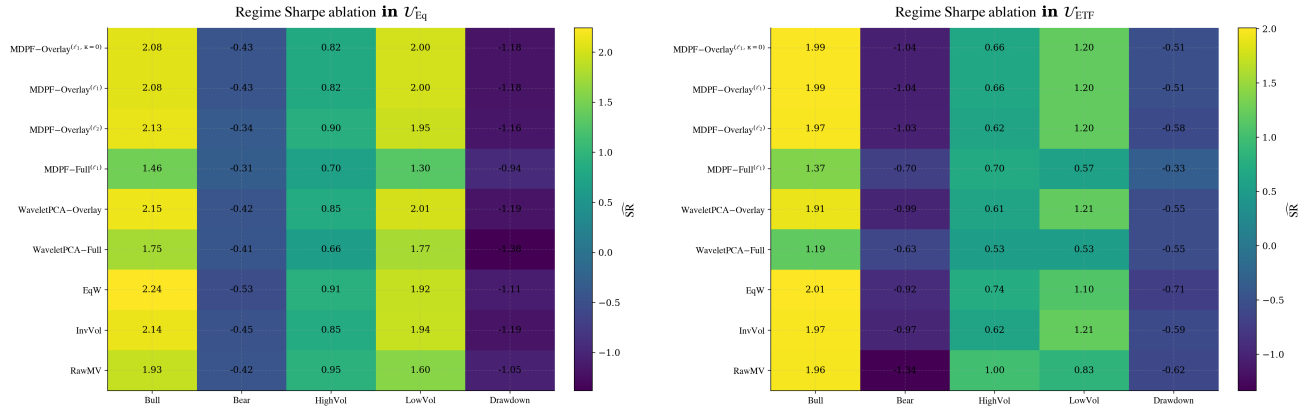
(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure B3. Overlay-strength sweep for the MDPF-Ov family.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure B4. Transaction-cost robustness of selected MDPF specifications.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure B5. Regime-specific Sharpe heatmaps across model variants and benchmarks.

Table B1. Wavelet-family rankings for the main overlay specification.

Universe	Rank	Wavelet family	Sharpe	Avg. ℓ_1 turnover	Max drawdown (%)
$\mathcal{U}_{\mathrm{Eq}}$	1	db4	1.082	0.167	29.59
$\mathcal{U}_{\mathrm{Eq}}$	2	haar	1.068	0.173	30.54
$\mathcal{U}_{\mathrm{Eq}}$	3	db2	1.068	0.173	30.59
$\mathcal{U}_{\mathrm{Eq}}$	4	sym4	1.053	0.160	30.85
$\mathcal{U}_{\mathrm{Eq}}$	5	coif3	1.039	0.151	30.21
$\mathcal{U}_{\mathrm{ETF}}$	1	db2	0.789	0.131	21.17
$\mathcal{U}_{\mathrm{ETF}}$	2	db4	0.785	0.117	21.00
$\mathcal{U}_{\mathrm{ETF}}$	3	haar	0.765	0.135	21.03
$\mathcal{U}_{\mathrm{ETF}}$	4	sym4	0.744	0.119	21.26
$\mathcal{U}_{\mathrm{ETF}}$	5	coif3	0.728	0.120	21.03

Table B2. Mathematical properties and typical financial applications of the evaluated wavelet families.

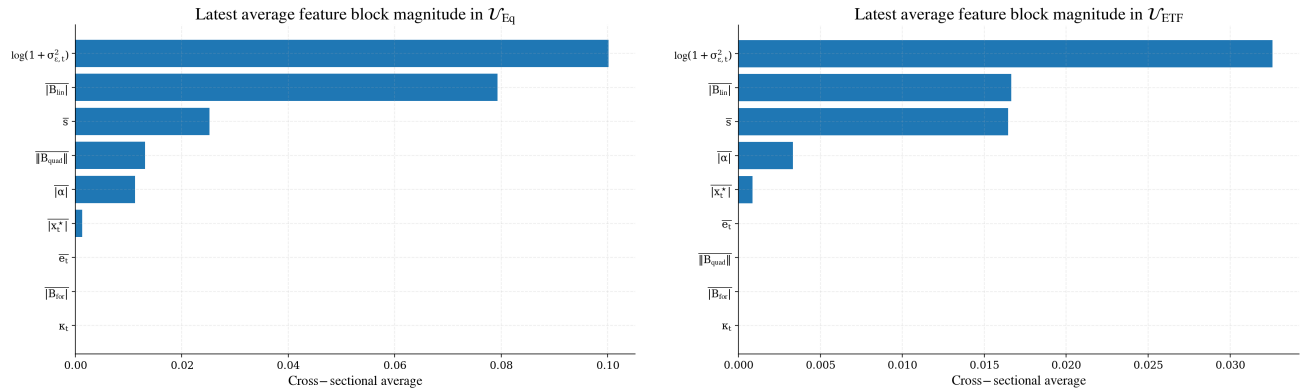
Family	Notation	Mathematical properties	Typical financial application
Haar	haar	Step function, discontinuous, orthogonal, asymmetric. Simplest wavelet.	Detecting abrupt regime shifts, structural breaks, and jump dynamics in asset prices.
Daubechies 2	db2	Continuous, asymmetric, orthogonal, compact support with 2 vanishing moments.	Capturing short-term, asymmetric volatility clusters and rapid trend reversals.
Daubechies 4	db4	Smoother than db2, asymmetric, orthogonal, compact support with 4 vanishing moments.	Extracting medium-term momentum signals and filtering moderate-frequency market noise.
Symlet 4	sym4	Nearly symmetric modification of db4, orthogonal, 4 vanishing moments.	Analyzing phase-sensitive cycles where time-alignment of peaks and troughs is critical.
Coiflet 3	coif3	Nearly symmetric, orthogonal, highest number of vanishing moments (6) among this set.	Isolating smooth, long-term structural trends while heavily suppressing high-frequency noise.

Table B3. Regime-specific Sharpe ratios for selected strategies across the two universes.

Universe	Strategy	Bull-market sharpe	Bear-market sharpe
$\mathcal{U}_{\text{Eq}}$	EqW	2.242	-0.527
$\mathcal{U}_{\text{Eq}}$	InvVol	2.139	-0.446
$\mathcal{U}_{\text{Eq}}$	MDPF-Full (ℓ_1 ODE)	1.464	-0.307
$\mathcal{U}_{\text{Eq}}$	MDPF-Ov (ℓ_2 ODE)	2.128	-0.337
$\mathcal{U}_{\text{Eq}}$	MDPF-Ov (ℓ_1 ODE)	2.084	-0.426
$\mathcal{U}_{\text{Eq}}$	RawMV	1.927	-0.424
$\mathcal{U}_{\text{Eq}}$	WaveletPCA-Ov	2.154	-0.424
$\mathcal{U}_{\text{ETF}}$	EqW	2.007	-0.916
$\mathcal{U}_{\text{ETF}}$	InvVol	1.971	-0.972
$\mathcal{U}_{\text{ETF}}$	MDPF-Full (ℓ_1 ODE)	1.372	-0.696
$\mathcal{U}_{\text{ETF}}$	MDPF-Ov (ℓ_2 ODE)	1.970	-1.035
$\mathcal{U}_{\text{ETF}}$	MDPF-Ov (ℓ_1 ODE)	1.988	-1.037
$\mathcal{U}_{\text{ETF}}$	RawMV	1.964	-1.335
$\mathcal{U}_{\text{ETF}}$	WaveletPCA-Ov	1.914	-0.989

C. Supplementary feature diagnostics

The supplementary feature plots deepen the main interpretation advanced in Section 1. Figure C1 summarizes the latest feature-block values by asset and confirms that the largest asset-specific differences did not reside in the endpoint or wavelet-energy blocks alone. Instead, the most visible contrasts were concentrated in the residual and interaction-related features. This is consistent with the feature-importance summary reported in Tables 5 and 6. The main implication is that the architecture is most informative when it transforms multiscale return histories into heterogeneous dynamic descriptors rather than when it merely reports static wavelet magnitudes.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

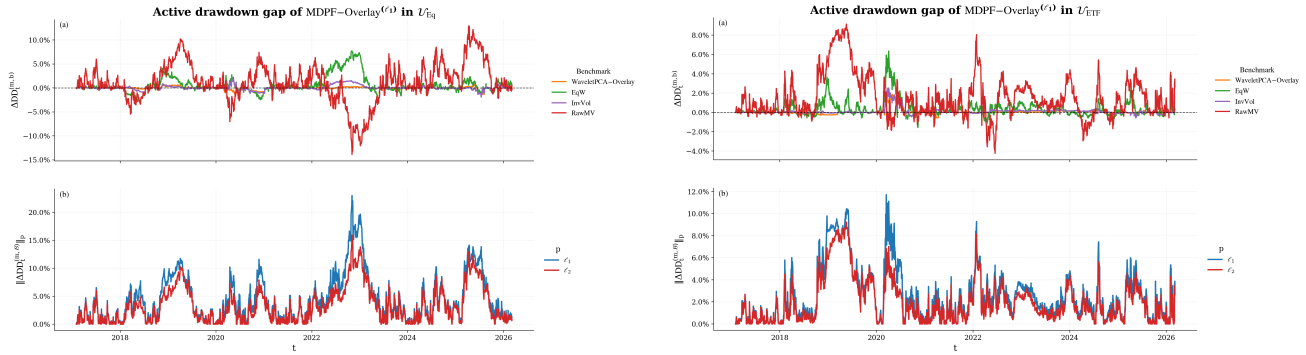
Figure C1. Latest asset-level summaries of the extracted MDPF feature blocks.

D. Supplementary benchmark-path diagnostics

Figures D1–D3 provide additional diagnostics for benchmark comparison. Figure D1 reports cumulative wealth paths rather than active gaps and therefore makes the absolute scale of benchmark dominance easier to see, especially for RawMV in both universes. Figure D2 translates the same exercise into active drawdown gaps, which clarifies that large relative underperformance episodes were episodic rather than constant. Figure D3 reports strategy-correlation matrices and confirms that the MDPF-Ov return stream remained highly aligned with several simpler strategies. This finding is a limitation for claims of distinct alpha generation, but it is also consistent with the interpretation that the model is primarily a representation layer whose value is expressed through structured tilts rather than through a fully orthogonal payoff profile.

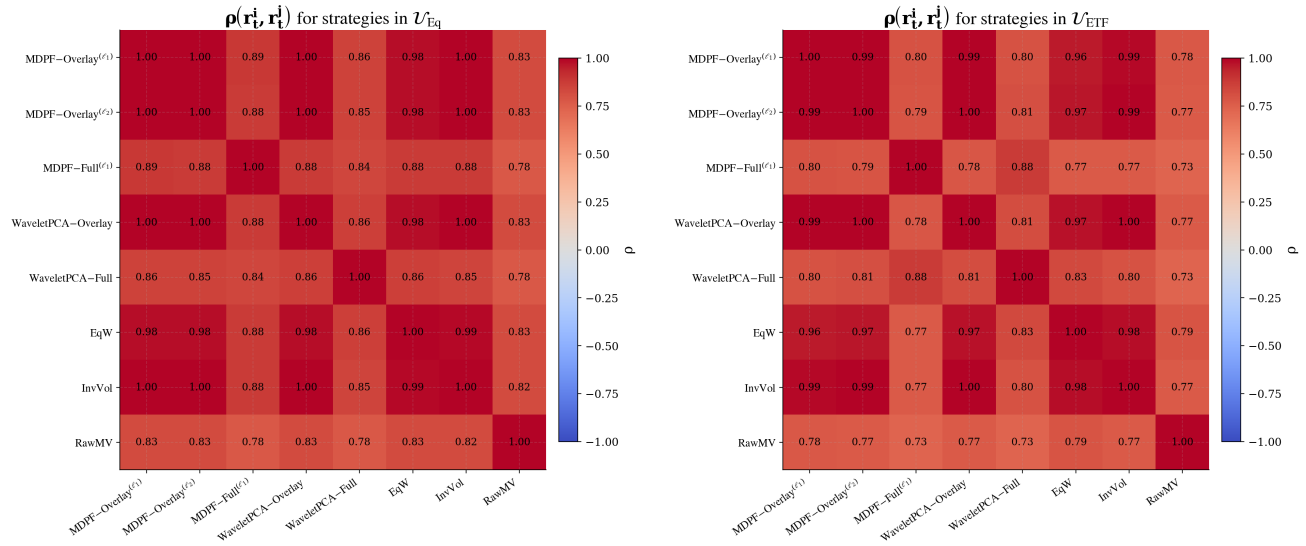
(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure D1. Cumulative wealth paths of selected strategies across the two universes.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure D2. Active drawdown gaps of the main MDPF-Ov (ℓ_1 ODE) relative to the benchmark set, together with aggregate ℓ_1 and ℓ_2 norms.



(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure D3. Strategy-correlation matrices for the set of selected strategies.

E. Notation guide for performance, feature, and allocation diagnostics

For ease of reference, this Appendix consolidates the notation used throughout the empirical analysis. Tables E1 and E2 summarize the symbols that appeared in the performance tables, benchmark-gap diagnostics, feature-extraction plots, and the allocation diagnostics introduced in Appendix E. Several supplementary allocation figures use the compact universe labels $\mathcal{U}_1$ and $\mathcal{U}_2$; throughout the Appendix these are shorthand for $\mathcal{U}_{\mathrm{Eq}}$ and $\mathcal{U}_{\mathrm{ETF}}$, respectively.

Table E1. Core notation used in the performance evaluation, benchmark comparison, and cross-universe figures.

Symbol / formula	Definition	Role in the results section
Universes, strategies, and reported portfolio quantities		
$\mathcal{U}_{\mathrm{Eq}}$, $\mathcal{U}_{\mathrm{ETF}}$	Equity and ETF universes, respectively	Identify the two investment universes throughout the cross-universe figures and comparative summary tables.
$\mathcal{U}_1$, $\mathcal{U}_2$	Compact universe labels used in the allocation-diagnostic appendix	Provide shorthand labels for $\mathcal{U}_{\mathrm{Eq}}$ and $\mathcal{U}_{\mathrm{ETF}}$ in selected supplementary figures.
(ℓ_1) and (ℓ_2) on strategy labels	Sparse and ridge ODE identification, respectively	Distinguish the two MDPF-Ov variants in the main performance plots, ablation panels, and sensitivity analysis.
$w_{i,t}$	Portfolio weight assigned to asset i at time t	Appears in the weight heatmaps and weight-path panels; summarizes allocation concentration and rebalancing dynamics.
$r_{p,t}$	Daily simple portfolio return	Primitive return series from which all reported performance statistics are computed.
A	Annualization factor implied by the daily sampling calendar	Converts daily moments into annualized quantities in the reported risk and performance measures.
$W_t = \prod_{\tau \le t}(1 + r_{p,\tau})$	Cumulative wealth index	Used in cumulative wealth figures and as the basis for drawdown measurement.
$L_t = \sum_{\tau \le t} \log(1 + r_{p,\tau})$	Cumulative log wealth	Used in cumulative log-return figures and in the active cumulative gap analysis.
$\mathrm{DD}_t = \dfrac{W_t}{\max_{\tau \le t} W_\tau} - 1$	Drawdown path	Used in the drawdown figures and in the active drawdown gap analysis.
μ_{ann}	Annualized return measure shown on the vertical axis of the risk–return scatter	In the present results, this corresponds to the annualized return statistic reported in the summary tables, namely the estimated compound annual growth rate.
$\widehat{\sigma}_{\mathrm{ann}} = \widehat{\mathrm{sd}}(r_{p,t})\sqrt{A}$	Annualized volatility	Horizontal axis in risk–return figures and a central risk descriptor in the main results tables.
$\widehat{\mathrm{SR}} = \dfrac{\bar{r}_p}{\widehat{\mathrm{sd}}(r_{p,t})}\sqrt{A}$	Annualized Sharpe ratio, where $\bar{r}_p$ denotes the sample mean of $r_{p,t}$	Primary ranking metric in the strategy comparison tables, ablation scorecards, and parameter sensitivity analysis.
$\widehat{\mathrm{CAGR}} = W_T^{A/T} - 1$	Compound annual growth rate over the evaluation horizon of length T	Summary measure of long-run return performance reported alongside volatility and Sharpe ratio.
$\widehat{\mathrm{MDD}} = \min_t \mathrm{DD}_t$	Maximum drawdown estimate	Captures downside severity and complements volatility-based risk measures.
$\overline{\mathrm{TO}}_{\ell_1}$	Average ℓ_1 turnover, computed as the sum of absolute portfolio-weight changes at rebalancing	Reported in the main performance and allocation-stability tables. Transaction-cost diagnostics use traded notional, equal to $0.5\,\overline{\mathrm{TO}}_{\ell_1}$.
O, F	Overlay and full-allocation implementations, respectively	Distinguish the two deployment modes in the allocation-stability appendix.
$N_{\mathrm{eff},t} = \left(\sum_i w_{i,t}^2\right)^{-1}$	Effective number of bets at time t	Measures realized portfolio diversification in the appendix weight diagnostics.

Continue on next page

Table E1 – continued from previous page

Symbol / formula	Definition	Role in the results section
$\Delta w_{i,t}^{(m,\mathrm{base})} = w_{i,t}^{(m)} - w_{i,t}^{(\mathrm{base})}$	Active tilt of strategy m relative to the base portfolio	Quantifies how the feature-driven overlay perturbs the base allocation.
$\|\Delta w_t^{(m,\mathrm{base})}\|_1,\ \|\Delta w_t^{(m,\mathrm{base})}\|_2$	ℓ_1 and ℓ_2 norms of the active-tilt vector	Summarize the size of the implied portfolio tilt in the allocation appendix.
$\overline{\max_i \lvert w_{i,t}\rvert}$	Time-averaged maximum absolute portfolio weight	Reports concentration of the realized allocation in the overlay-versus-full stability table.
$\rho(r_t^i, r_t^j)$	Pairwise correlation between strategy return series i and j	Used in the strategy-correlation heatmaps to assess diversification and redundancy across allocation rules.
Active comparison against benchmarks		
m	Focal MDPF strategy under evaluation	Denotes the model of interest in the active cumulative gap and active drawdown gap figures.
$b \in \mathcal{B}$	Benchmark strategy belonging to the benchmark set $\mathcal{B}$	Indexes the benchmark lines against which the focal MDPF strategy is compared.
$\Delta L_t^{(m,b)} = L_t^{(m)} - L_t^{(b)}$	Active cumulative log-wealth gap of the focal strategy relative to benchmark b	Upper panel of the cumulative gap figures; positive values indicate outperformance relative to the selected benchmark.
$\|\Delta L_t^{(m,\mathcal{B})}\|_p = \left(\sum_{b\in\mathcal{B}} \left\lvert\Delta L_t^{(m,b)}\right\rvert^p\right)^{1/p},\quad p \in \{1,2\}$	Aggregate norm of cumulative log-wealth gaps across all benchmarks	Lower panel of the cumulative gap figures; summarizes the global magnitude of active separation from the benchmark set.
$\Delta \mathrm{DD}_t^{(m,b)} = \mathrm{DD}_t^{(m)} - \mathrm{DD}_t^{(b)}$	Active drawdown gap of the focal strategy relative to benchmark b	Upper panel of the drawdown gap figures; negative values indicate a smaller drawdown than the benchmark.
$\|\Delta \mathrm{DD}_t^{(m,\mathcal{B})}\|_p = \left(\sum_{b\in\mathcal{B}} \left\lvert\Delta \mathrm{DD}_t^{(m,b)}\right\rvert^p\right)^{1/p}$ $p \in \{1,2\}$	Aggregate norm of drawdown gaps across all benchmarks	Lower panel of the drawdown gap figures; quantifies the overall divergence of the MDPF drawdown path from the benchmark set.

Note. The strategy labels used in the figures follow the same notation as the quantitative tables. In particular, MDPF−Overlay$^{(\ell_1)}$ and MDPF−Overlay$^{(\ell_2)}$ denote the sparse and ridge ODE overlays, respectively; MDPF−Full$^{(\ell_1)}$ denotes the full-allocation sparse ODE variant; WaveletPCA−Overlay and WaveletPCA−Full denote the corresponding wavelet–PCA benchmarks; and EqW, InvVol, and RawMV denote equal-weight, inverse-volatility, and raw mean–variance portfolios.

Table E2. Feature-extraction and model-parameter notation used in the MDPF diagnostic figures, ablations, and sensitivity analysis.

Symbol / formula	Definition	Role in the results section
MDPF feature blocks		
$x_t^\star$	Endpoint multiscale state at the end of the lookback window	Represents the current wavelet-based state entering the MDPF feature vector.
e_t	Wavelet-energy vector across retained scales	Measures the magnitude of multiscale variation and appears as one of the principal feature blocks.
α	ODE intercept block	Captures baseline drift in the local continuous-time dynamics and is summarized in the feature-block diagnostics.
B_{lin}	Linear coefficient block of the fitted ODE	Encodes first-order multiscale dynamics and appears in the feature-block magnitude summaries.
B_{for}	Common forcing coefficient block of the fitted ODE	Measures sensitivity to the cross-sectional forcing state and is reported in the feature diagnostics.

Continued on next page

Table E2 – continued from previous page

Symbol / formula	Definition	Role in the results section
B_{quad}	Quadratic interaction block of the fitted ODE	Encodes nonlinear cross-scale effects; in the figures it is summarized through block norms rather than individual coefficients.
s	Sparsity ratio of active ODE coefficients	Monitors effective model complexity and identifies whether the fitted dynamics are dense or parsimonious.
κ_t	Standardized local instability score derived from the symmetric part of the local Jacobian	Used as a stability diagnostic and in the ablation analysis of the instability penalty.
$\log(1+\sigma^2_{\varepsilon,t})$	Log residual variance of the ODE fit	Measures model misfit and residual noise; one of the most informative feature blocks in the reported diagnostics.
$z_{k,t}$	kth principal feature score after PCA compression of the MDPF feature matrix	Appears in the feature-score heatmaps and represents the compressed latent representation used downstream in allocation.
$\overline{\lvert x^\star_t\rvert}$, $\overline{e_t}$, $\overline{\lvert\alpha\rvert}$, $\overline{\lvert B_{\text{lin}}\rvert}$, $\overline{\lvert B_{\text{for}}\rvert}$, $\overline{\lVert B_{\text{quad}}\rVert}$, $\overline{s}$	Cross-sectional summaries of the principal feature blocks	Used in the feature-block heatmaps and latest-date block summary plots to compare which components dominate the extracted representation.
Sensitivity and architecture parameters		
$\lambda_1^{(\text{ODE})}$	ℓ_1 penalty used in ODE identification	Governs sparse ODE estimation and appears in the one-factor parameter sensitivity analysis.
$\lambda_2^{(\text{ODE})}$	ℓ_2 penalty used in ODE identification	Governs ridge-type shrinkage of the ODE coefficients and appears in the one-factor parameter sensitivity analysis.
ρ_{PCA}	Explained-variance threshold used for principal-component retention	Controls the dimension of the latent feature space and appears in the parameter sensitivity analysis.
γ	Risk-aversion coefficient in the portfolio optimization step	Determines the trade-off between expected return and risk in the allocation layer.
λ_{TO}	Turnover penalty parameter	Controls trading aggressiveness and appears in the sensitivity analysis of implementation stability.
β_{overlay}	Overlay blend applied to the MDPF signal	Determines how strongly the extracted features tilt the base portfolio and is central to the overlay-strength ablation.
c_{tc}	Proportional transaction-cost rate	Used in the transaction-cost robustness analysis to evaluate implementation sensitivity.
$\mathcal{W}$	Selected mother-wavelet family	Appears in the wavelet-family sensitivity analysis and determines the multiscale basis used for feature construction.

Note. In the feature-diagnostic figures, the overbar denotes a cross-sectional block summary. For the time-evolution heatmap, the block summary is reported as a cross-sectional median at each date; for the latest-date block summary, it is reported as a cross-sectional mean magnitude. This convention matches the feature labels used in the plotted diagnostics.

F. Allocation and weight-stability diagnostics

This Appendix links the extracted MDPF representation to the empirical allocation outputs. The purpose is not to reinterpret the model as a stand-alone optimizer, but to verify that the learned features produced disciplined and economically plausible portfolio tilts when the architecture was deployed as an overlay. Table F1 and Figure F1 compare overlay and full-allocation implementations across the two universes. Within the MDPF family, average ℓ_1 turnover fell from 1.224 to 0.169 in $\mathcal{U}_{\text{Eq}}$ and from 1.158 to 0.112 in

$\mathcal{U}_{\mathrm{ETF}}$ when the model was used as an overlay rather than as a full allocator. Over the same comparison, the effective number of bets increased from 3.93 to 10.54 in the equity universe and from 3.79 to 9.01 in the ETF universe, while the mean maximum absolute portfolio weight fell from approximately 0.30 to 0.14 and 0.19, respectively. The same qualitative pattern is visible in the WaveletPCA family. This parallel finding is informative rather than incidental: It shows that part of the stability gain comes from the overlay design, while the MDPF architecture retains the richer representation documented in the main text.

Table F1. Overlay-versus-full allocation stability across the two universes. O denotes the overlay implementation and F denotes the corresponding full-allocation implementation within the same model family. Turnover reductions are reported as full minus overlay using average ℓ_1 turnover, so larger positive values indicate greater stability of the overlay design.

Universe	Family	$\overline{\mathrm{TO}}^{\mathrm{O}}_{\ell_1}$	$\overline{\mathrm{TO}}^{\mathrm{F}}_{\ell_1}$	$\overline{N}^{\mathrm{O}}_{\mathrm{eff}}$	$\overline{N}^{\mathrm{F}}_{\mathrm{eff}}$	$\overline{\|\Delta w\|_1}^{\mathrm{O}}$	$\overline{\|\Delta w\|_1}^{\mathrm{F}}$	$\overline{\max\|w_i\|}^{\mathrm{O}}$	$\overline{\max\|w_i\|}^{\mathrm{F}}$	$\Delta\overline{\mathrm{TO}}_{\ell_1}$	$\Delta\overline{N}_{\mathrm{eff}}$	$\Delta\overline{\|\Delta w\|}_1$	$\Delta\overline{\max\|w_i\|}$
$\mathcal{U}_{\mathrm{Eq}}$	MDPF	0.169	1.224	10.540	3.928	0.122	1.246	0.138	0.299	1.055	6.612	1.124	0.162
$\mathcal{U}_{\mathrm{Eq}}$	WaveletPCA	0.137	1.214	10.853	3.846	0.071	1.300	0.132	0.300	1.077	7.007	1.228	0.168
$\mathcal{U}_{\mathrm{ETF}}$	MDPF	0.112	1.158	9.015	3.794	0.068	1.300	0.191	0.300	1.046	5.221	1.232	0.109
$\mathcal{U}_{\mathrm{ETF}}$	WaveletPCA	0.106	1.043	9.072	3.734	0.057	1.368	0.191	0.300	0.937	5.338	1.311	0.109

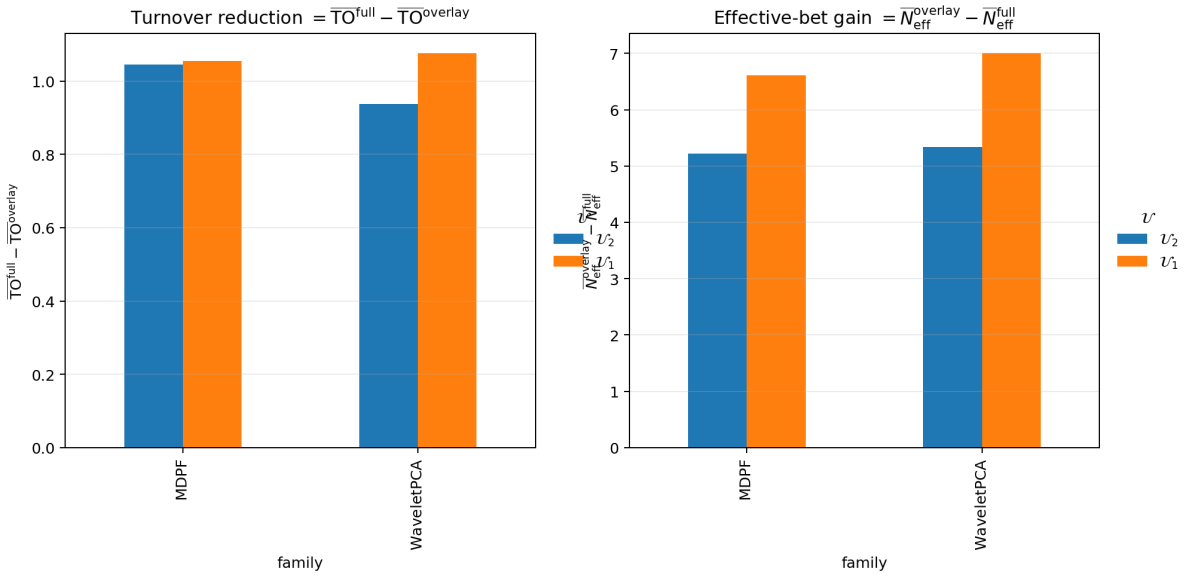
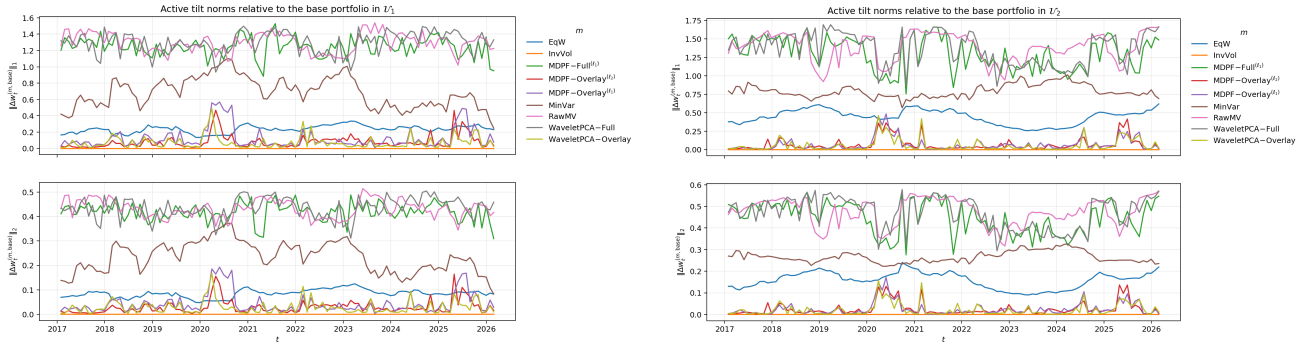


Figure F1. Cross-universe overlay-versus-full allocation stability. The left panel reports the reduction in average ℓ_1 turnover, $\Delta\overline{\mathrm{TO}}_{\ell_1} = \overline{\mathrm{TO}}^{\mathrm{F}}_{\ell_1} - \overline{\mathrm{TO}}^{\mathrm{O}}_{\ell_1}$, and the right panel reports the gain in effective diversification, $\Delta\overline{N}_{\mathrm{eff}} = \overline{N}^{\mathrm{O}}_{\mathrm{eff}} - \overline{N}^{\mathrm{F}}_{\mathrm{eff}}$. In the plotted legend, $\mathcal{U}_2$ and $\mathcal{U}_1$ correspond to $\mathcal{U}_{\mathrm{ETF}}$ and $\mathcal{U}_{\mathrm{Eq}}$, respectively.

Figure F2 examines the time variation in active tilts relative to the base portfolio. In both universes, the main MDPF-Ov (ℓ_1 ODE) maintained small and persistent active-tilt norms, whereas the full allocators and the more aggressive benchmark allocators deviate much more strongly from the base allocation. In $\mathcal{U}_{\mathrm{Eq}}$, $\|\Delta w_t^{(m,\mathrm{base})}\|_1$ for the sparse overlay spent most of the sample below about 0.2, while the sparse full allocator and the wavelet full allocator fluctuated mainly in the 1.2–1.5 range. The ETF

universe shows the same ordering, with even smaller overlay norms. These trajectories are important because they indicate that the extracted features act primarily as controlled tilts rather than as triggers for repeated wholesale reallocation.



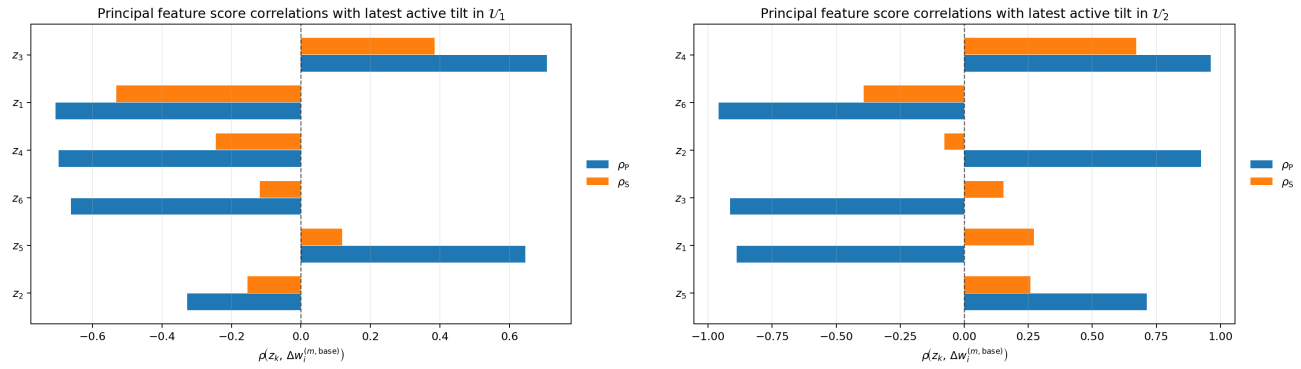
(a) $\mathcal{U}_{\mathrm{Eq}}$. **(b)** $\mathcal{U}_{\mathrm{ETF}}$.

Figure F2. Time-series of active-tilt norms relative to the base portfolio. The main MDPF-Ov (ℓ_1 ODE) remains close to the base allocation in both universes, while the full allocators and more aggressive benchmarks produce substantially larger deviations.

To connect the allocation outputs directly to the extracted representation, Table F2 and Figure F3 report cross-sectional correlations between the latest active tilt of MDPF-Ov (ℓ_1 ODE) and the learned features. In $\mathcal{U}_{\mathrm{Eq}}$, the strongest latent-score associations involved z_3 ($\rho_{\mathrm{P}} = 0.707$), z_1 (-0.707), and z_4 (-0.697), while the dominant block-level quantities were the intercept magnitude, log residual variance, and wavelet energy, each with absolute Pearson correlation above 0.71. In $\mathcal{U}_{\mathrm{ETF}}$, the alignment was materially stronger: z_4 and z_6 reached absolute Pearson correlations above 0.95, and $\overline{|B_{\mathrm{lin}}|}$ and κ_t reached -0.953. Because these statistics were computed across only twelve assets per universe and at a single evaluation date, they should be read as descriptive alignment diagnostics rather than as structural estimates. Even with that caution, the results showed that the allocation layer reacted systematically to the extracted representation. The features were not merely computed; they were economically active in the construction of the final portfolio tilts.

Table F2. Strongest contemporaneous associations between the latest active tilt of the main MDPF-Ov (ℓ_1 ODE) and the extracted representation. For each universe, the table reports the three largest absolute Pearson correlations among principal feature scores and among block-level feature summaries. The statistics are descriptive cross-sectional diagnostics over the twelve assets in each universe.

Universe	Type	Quantity	ρ_P	p_P	ρ_S	p_S
$\mathcal{U}_{Eq}$	Latent score	z_3	0.707	0.010	0.385	0.217
$\mathcal{U}_{Eq}$	Latent score	z_1	-0.707	0.010	-0.531	0.075
$\mathcal{U}_{Eq}$	Latent score	z_4	-0.697	0.012	-0.245	0.443
$\mathcal{U}_{Eq}$	Feature block	$\overline{\lvert\alpha\rvert}$	-0.723	0.008	-0.329	0.297
$\mathcal{U}_{Eq}$	Feature block	$\log(1+\sigma^2_{\varepsilon,t})$	-0.719	0.008	-0.364	0.245
$\mathcal{U}_{Eq}$	Feature block	$\overline{e_t}$	-0.712	0.009	-0.343	0.276
$\mathcal{U}_{ETF}$	Latent score	z_4	0.962	0.000	0.671	0.017
$\mathcal{U}_{ETF}$	Latent score	z_6	-0.960	0.000	-0.392	0.208
$\mathcal{U}_{ETF}$	Latent score	z_2	0.924	0.000	-0.077	0.812
$\mathcal{U}_{ETF}$	Feature block	$\overline{\lvert B_{lin}\rvert}$	-0.953	0.000	-0.480	0.114
$\mathcal{U}_{ETF}$	Feature block	κ_t	-0.953	0.000	-0.480	0.114
$\mathcal{U}_{ETF}$	Feature block	$\log(1+\sigma^2_{\varepsilon,t})$	-0.907	0.000	0.154	0.633



(a) $\mathcal{U}_{Eq}$. **(b)** $\mathcal{U}_{ETF}$.

Figure F3. Principal feature-score correlations with the latest active tilt of the main MDPF-Ov (ℓ_1 ODE). The bars report Pearson and Spearman correlations across assets and illustrate how the compressed latent representation is translated into the final cross-sectional portfolio tilt.

Taken together, the weight diagnostics strengthen the interpretation of MDPF as a feature-extraction methodology. The value of the architecture does not arise from more concentrated or more frequently traded portfolios. Instead, the extracted representation is useful because it generates diversified and comparatively stable tilts that remain tightly linked to the latent MDPF states. This result is consistent with the main-text conclusion that the architecture is most effective as an overlay layer rather than as a stand-alone allocator.

AIMS Press