



Wavelet Denoising and Model Parsimony in High-Frequency Momentum Strategies: Evidence from China's ChiNext Market

Yingnan Yi^{1,2,*}

¹Olin Business School, Washington University in St. Louis, St. Louis, MO 63130, United States

²Beijing QITANG Education Consulting Inc., Beijing, China

Abstract

This study introduces a sophisticated dual-denoising framework for high-frequency momentum strategies in China's ChiNext market, integrating wavelet-based temporal filtering with isolation forest cross-sectional anomaly detection. Utilizing a comprehensive dataset of over 2 million daily observations from January 2016 to November 2025, the results demonstrate that wavelet denoising achieves exceptional efficacy for turnover series with a mean Signal-to-Noise Ratio improvement of 10.7 dB, while isolation forest robustly identifies anomalous stocks characterized by excessive trading activity and distorted risk-return profiles. Empirical results reveal that parsimonious models with wavelet denoising consistently outperform complex alternatives, with linear-kernel Support Vector Machines achieving Sharpe ratios of 0.1132 in long portfolios and 0.0649 in long-short implementations. Contrary to conventional wisdom, simpler architectures demonstrate superior

performance, with single-layer neural networks and regularized linear models surpassing deeper architectures across all denoising configurations. The methodology demonstrates remarkable regime invariance, maintaining consistent performance during both low-volatility and high-volatility market states. These findings establish wavelet denoising as a critical preprocessing technique for high-frequency financial data and challenge assumptions about the necessity of complex models in quantitative finance, providing computationally efficient approaches suitable for both institutional and retail investors.

Keywords: momentum investing, wavelet denoising, portfolio management, chiNext market, machine learning, financial signal processing.

1 Introduction

The exponential growth of algorithmic trading has fundamentally transformed quantitative finance, creating unprecedented opportunities for signal extraction while introducing significant challenges in distinguishing genuine market patterns from microstructure noise. High-frequency momentum strategies, operating at daily frequencies, face particularly acute signal degradation challenges



Submitted: 22 November 2025

Revised: 09 January 2026

Accepted: 10 January 2026

Published: 10 April 2026

Vol. 2, No. 3, 2026.

10.62762/TMI.2025.742127

*Corresponding author:

✉ Yingnan Yi

yiyingnan2021@163.com

Citation

Yi, Y. (2026). Wavelet Denoising and Model Parsimony in High-Frequency Momentum Strategies: Evidence from China's ChiNext Market. *ICCK Transactions on Machine Intelligence*, 2(3), 172–189.

© 2026 ICCK (Institute of Central Computation and Knowledge)

due to the substantial noise inherent in intraday return and volume data. This noise—comprising market microstructure effects, transient liquidity shocks, and behavioral anomalies—poses critical constraints on strategy performance that traditional enhancement approaches often fail to address adequately. Wavelet-based denoising has emerged as a promising alternative for high-frequency financial data, offering systematic noise reduction through multi-resolution signal decomposition [1].

While the momentum anomaly remains one of the most persistent phenomena in financial economics [2, 3], conventional daily momentum implementations exhibit pronounced vulnerability to noise contamination. Traditional enhancement methodologies, including cross-asset factor diversification strategies that exploit the pervasive co-movement between value and momentum premia [4] and sophisticated risk management [5], typically address symptoms rather than root causes, leaving strategies exposed to performance deterioration during market transitions. The critical challenge lies in developing preprocessing techniques capable of systematically separating genuine momentum signals from the complex noise signature present in high-frequency financial data.

Wavelet analysis offers a mathematically rigorous solution through its capacity for multi-resolution signal decomposition [6, 7], providing an optimal framework for temporal noise separation in non-stationary financial time series. Unlike traditional Fourier methods that assume stationarity, wavelet transforms enable simultaneous localization in both time and frequency domains, making them particularly suited to the transient, volatility-clustered nature of financial data. However, the application of wavelet denoising specifically for momentum factor enhancement at daily frequencies remains critically underdeveloped in the literature.

This research bridges this methodological gap by developing an integrated dual-denoising framework specifically engineered for daily momentum strategies in China's ChiNext market—an ideal laboratory characterized by elevated volatility, retail-driven price dynamics, and significant microstructure effects. Our approach uniquely combines wavelet multi-resolution analysis for temporal noise separation with isolation forest algorithms for cross-sectional anomaly detection, creating a comprehensive preprocessing pipeline for daily financial data.

Our empirical investigation yields several notable insights that challenge conventional wisdom in quantitative finance. First, we demonstrate a pronounced dichotomy in denoising efficacy: wavelet methods achieve exceptional performance for turnover series (10.7 dB SNR improvement) while showing limited effectiveness for return series (-2.6 dB SNR), revealing fundamental differences in noise structures across financial variables. Second, we establish the superiority of model parsimony, with linear models and single-layer neural networks consistently outperforming more complex architectures across all denoising configurations. Third, we document robust performance in both long and long-short portfolio implementations, indicating that the purified signals possess both directional and relative value components.

The portfolio performance analysis reveals that Support Vector Machines with wavelet denoising achieve superior risk-adjusted returns (Sharpe ratio: 0.1132), closely followed by LASSO implementations (Sharpe ratio: 0.0971). This performance hierarchy demonstrates that well-specified linear models with appropriate feature processing can effectively compete with—and often surpass—more complex nonlinear alternatives. The consistent outperformance of simpler architectures suggests that the denoising process effectively linearizes the prediction task, reducing the need for complex hierarchical representations.

Our contributions are threefold: first, we pioneer a wavelet-based denoising methodology optimized for daily financial time series, explicitly addressing the unique noise characteristics at this frequency; second, we provide novel evidence for the superiority of model parsimony in financial machine learning applications; third, we demonstrate robust signal extraction across both directional and market-neutral strategy implementations, providing practical frameworks adaptable to diverse investment objectives.

The remainder of this paper proceeds as follows. Section 2 surveys relevant literature on high-frequency momentum, wavelet applications in finance, and model complexity in financial machine learning. Section 3 details our methodological framework and dataset construction. Section 4 presents comprehensive empirical findings, and Section 5 discusses implications for both academic research and quantitative trading practice.

2 Related Work

2.1 High-Frequency Momentum and Microstructure Challenges

The momentum anomaly, first systematically documented by [2], represents one of the most robust phenomena in empirical finance. While extensive literature confirms its persistence across asset classes and time periods [4], with even markets initially viewed as exceptions ultimately conforming to momentum dynamics [3], the migration to high-frequency domains reveals fundamental challenges rooted in market microstructure theory. As formalized by O'Hara [8] and Hasbrouck [9], observed high-frequency prices represent a composite of latent "efficient prices" and microstructure noise, creating significant obstacles for momentum strategy implementation.

The microstructure noise component includes bid-ask bounce effects [10], non-synchronous trading [11], and the permanent price impact of informed order flow arising from information asymmetry between strategic traders and market makers [12]. In contemporary electronic markets, these effects are amplified by algorithmic trading activities [13], creating a substantial disconnect between short-term price movements and fundamental value. Consequently, high-frequency momentum signals often capture spurious, mean-reverting noise rather than genuine persistent trends—a challenge compounded by the fact that momentum, as a high-turnover strategy, consistently ranks among the most transaction-cost-sensitive anomalies, with implementation costs frequently eroding a substantial portion of gross returns [14].

Barroso and Santa-Clara [5] demonstrated momentum's vulnerability to noise-induced crashes in monthly data, a vulnerability that becomes exponentially more acute at daily frequencies. The fundamental challenge lies in the signal-to-noise ratio degradation as sampling frequency increases—while high-frequency data offers granularity, it also introduces overwhelming microstructure contamination that conventional momentum calculations cannot adequately address.

2.2 Wavelet Applications in Financial Signal Processing

Wavelet analysis represents a paradigm shift in financial time series processing, offering significant advantages over traditional methods

for non-stationary, high-frequency data. The mathematical foundations established by Mallat's seminal multiresolution analysis framework [6] and Daubechies [15] provide a rigorous basis for multi-resolution decomposition, enabling simultaneous localization in both time and frequency domains.

Seminal work by Gencay et al. [7] established wavelet superiority for volatility forecasting and multiscale market analysis. Unlike Fourier methods that assume stationarity, wavelet transforms adapt naturally to the time-varying volatility and transient patterns characteristic of financial data. The biorthogonal wavelet systems employed in this study, following [16], ensure perfect reconstruction while maintaining linear phase characteristics essential for financial applications.

Wavelet denoising techniques, particularly those based on [17], provide statistically coherent methods for separating signal from noise components. However, existing financial applications have predominantly focused on forecasting [18], market contagion and structural testing [19], or volatility modeling, with limited attention to momentum factor enhancement. Although exceptions do exist [20], the application of wavelet denoising specifically as a preprocessing step for daily momentum strategy construction represents a significant gap in the literature.

2.3 Model Complexity in Financial Machine Learning

The rapid advancement of machine learning techniques has revolutionized empirical asset pricing [21], yet fundamental questions regarding optimal model complexity remain unresolved. While complex architectures offer theoretical capacity for capturing nonlinear relationships, they also introduce substantial overfitting risks in finite financial samples.

Gu et al. [21] demonstrated that carefully regularized linear models can compete effectively with complex alternatives in return prediction tasks. This finding challenges conventional assumptions about the necessity of sophisticated nonlinear architectures for financial applications. Similarly, Kozak et al. [22] highlighted the critical importance of regularization and dimension reduction in high-dimensional financial settings.

Recent evidence suggests that the benefits of model complexity may be context-dependent. Moritz and Zimmermann [23] found that tree-based methods

excel in cross-sectional prediction tasks. However, the interaction between model complexity and preprocessing methodologies—particularly denoising techniques that build upon econometric frameworks for separating microstructure noise from underlying price volatility [24]—remains underexplored.

2.4 Chinese Market Microstructure and ChiNext Characteristics

China's equity markets present unique microstructure characteristics that differentiate them from developed markets. The ChiNext board, established in 2009 as China's strategic platform for innovative enterprises, exhibits particularly pronounced retail participation, elevated volatility, and distinct regulatory frameworks.

Liu et al. [25] documented the distinct size and value factor dynamics in Chinese equity markets, revealing systematic return patterns and pricing characteristics that differentiate China from developed markets and create potential opportunities for quantitative strategies. However, these opportunities come with unique challenges, including short-selling constraints, liquidity variations, and regulatory interventions that impact strategy implementation.

The high-frequency environment of ChiNext, characterized by significant retail trading activity and evolving market structure, provides an ideal testing ground for denoising methodologies. The market's distinctive features amplify both the challenges of noise contamination and the potential benefits of sophisticated signal processing techniques.

2.5 Research Gap and Contribution

The literature reveals a clear research opportunity at the intersection of wavelet denoising, model complexity optimization, and high-frequency momentum enhancement. While individual components—wavelet methods, anomaly detection, machine learning architectures—have been studied independently, their integrated application to daily momentum strategy construction remains unexplored.

This study makes three primary contributions to the literature: first, we develop an integrated dual-denoising framework specifically optimized for daily financial data; second, we provide comprehensive evidence regarding optimal model complexity in denoised financial prediction tasks; third, we demonstrate robust performance across both directional and market-neutral implementations in China's distinctive market environment. By

addressing noise contamination at its source and optimizing model architecture selection, we bridge critical gaps between financial theory and robust practical implementation.

3 Methodology

3.1 Data Sources and Sample Construction

This study employs a comprehensive high-frequency dataset from the ChiNext market, a NASDAQ-style board established by the Shenzhen Stock Exchange in October 2009 as China's strategic platform for innovative growth enterprises. Our dataset spans January 1, 2016 to November 20, 2025, comprising 2,086,231 daily stock-return and turnover observations with an average cross-section of 869 firms per trading day. To support the initial 252-trading-day model calibration window, data from the full calendar year 2015 were additionally collected and used exclusively for model initialization, remaining outside the formal out-of-sample evaluation period.

Data sourced from Baostock ensures reliability and comprehensive coverage of Chinese A-shares. The selected decade encompasses multiple distinct market regimes—including normal growth, the COVID-19 pandemic volatility, and subsequent recovery—providing a stringent testing ground for methodology evaluation across different volatility environments.

We implement rigorous data filters to ensure integrity and mitigate biases:

1. **Trading Activity Filter:** Exclude firm-day observations for companies with fewer than 66 trading days (approximately three calendar months) of data within the sample period.
2. **Data Integrity Filter:** Remove stocks with missing price or return data for more than 66 consecutive trading days to guard against illiquidity and potential data errors.
3. **Survivorship Bias Mitigation:** Include delisted stocks up to their final trading date, thereby capturing the full effect of delistings on strategy performance.

All stock price data are adjusted for corporate actions—including dividends, stock splits, and rights offerings—using the forward-adjusted method. This produces continuous and economically meaningful price series crucial for accurate return calculations and uncontaminated momentum signals.

After applying these filters, the final dataset constitutes a high-frequency, high-dimensional panel, with the cross-sectional firm count evolving from approximately 367 in early 2016 to over 1,335 by 2025. The average cross-section of 869 firms reflects the heterogeneous listing history across securities, as firms with fewer than 66 trading days and those with extended data gaps are excluded, resulting in a realized total of 2,086,231 firm-day observations over the sample period. This comprehensive sample represents the vast majority of the ChiNext market's liquidity and capitalization throughout the decade. As illustrated in Figure 1, the number of listed companies in the ChiNext market exhibited steady expansion over the sample period, increasing from around 367 stocks in early 2016 to more than 1,335 by November 2025.

3.2 Predictive Framework and Feature Engineering

The predictive framework leverages high-frequency data through two primary explanatory variable categories: historical stock returns and trading volumes. Daily stock returns are computed using the standard percentage change formula:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100\% \quad (1)$$

where R_t represents the return on day t , while P_t and P_{t-1} denote forward-adjusted closing prices at days t and $t - 1$, respectively.

Trading activity is quantified through daily turnover, calculated as:

$$\text{Turnover}_t = \frac{\text{Shares Traded}_t}{\text{Float Shares}_t} \times 100\% \quad (2)$$

where *Shares Traded* refers to total volume during day t , and float shares represents common shares available for public trading.

To capture short-term momentum and liquidity effects while rigorously avoiding look-ahead bias, both return and turnover series are lagged by 1 to 12 trading days (approximately two weeks). This generates 24 distinct predictors per security per day, balancing model complexity with computational efficiency.

Our forecasting methodology employs a rolling-window approach where, to forecast stock i 's return on day t , the training sample incorporates the previous 252 trading days (approximately one calendar year) of stock returns and trading volumes with their respective lagged values.

The model estimation requires an initial 252-trading-day calibration window for parameter estimation, utilizing 2015 data exclusively for model initialization outside the formal evaluation sample. The out-of-sample backtesting period spans January 1, 2016 through November 20, 2025, providing 2,086,231 daily observations for robust performance evaluation completely free of look-ahead bias.

3.3 Portfolio Construction Methodology

The portfolio construction follows a disciplined quantitative approach. After generating daily cross-sectional return forecasts, securities are ranked by their predicted returns. The strategy implements long positions in the top 1% of forecasted winners and short positions in the bottom 1% of forecasted losers. With an average daily cross-section of 869 stocks, this selection criterion yields approximately 8-9 securities per portfolio, providing natural diversification while maintaining focus on the strongest momentum signals.

This methodological design is grounded in two research considerations. First, it preserves the integrity of the momentum signal generated by our high-frequency forecasting model, avoiding arbitrary portfolio constraints that could introduce noise or bias. Second, the approach aligns with our core research objective of isolating the value added by wavelet denoising techniques, prioritizing methodological consistency across strategy comparisons over implementation optimization.

We evaluate three portfolio configurations: long portfolio (top 1%), short portfolio (bottom 1%), and a zero-investment hedged portfolio (long-short). It is important to note that within China's regulatory framework, particularly on the ChiNext board, short selling restrictions limit practical implementation of the short portfolio. Therefore, while the short portfolio analysis provides theoretical completeness in demonstrating the momentum effect, the long portfolio offers direct practical relevance, and the hedged portfolio serves as the primary academic benchmark for assessing the denoising methodology's efficacy in high-frequency environments.

3.4 Theoretical Foundation of Dual-Denoising Orthogonality

The dual-denoising methodology proposed in this study is predicated on a rigorous mathematical framework that conceptualizes financial market data as inhabiting a composite space where noise components occupy orthogonal subspaces. This theoretical

ChiNext Market: Daily Number of Listed Stocks Over Time

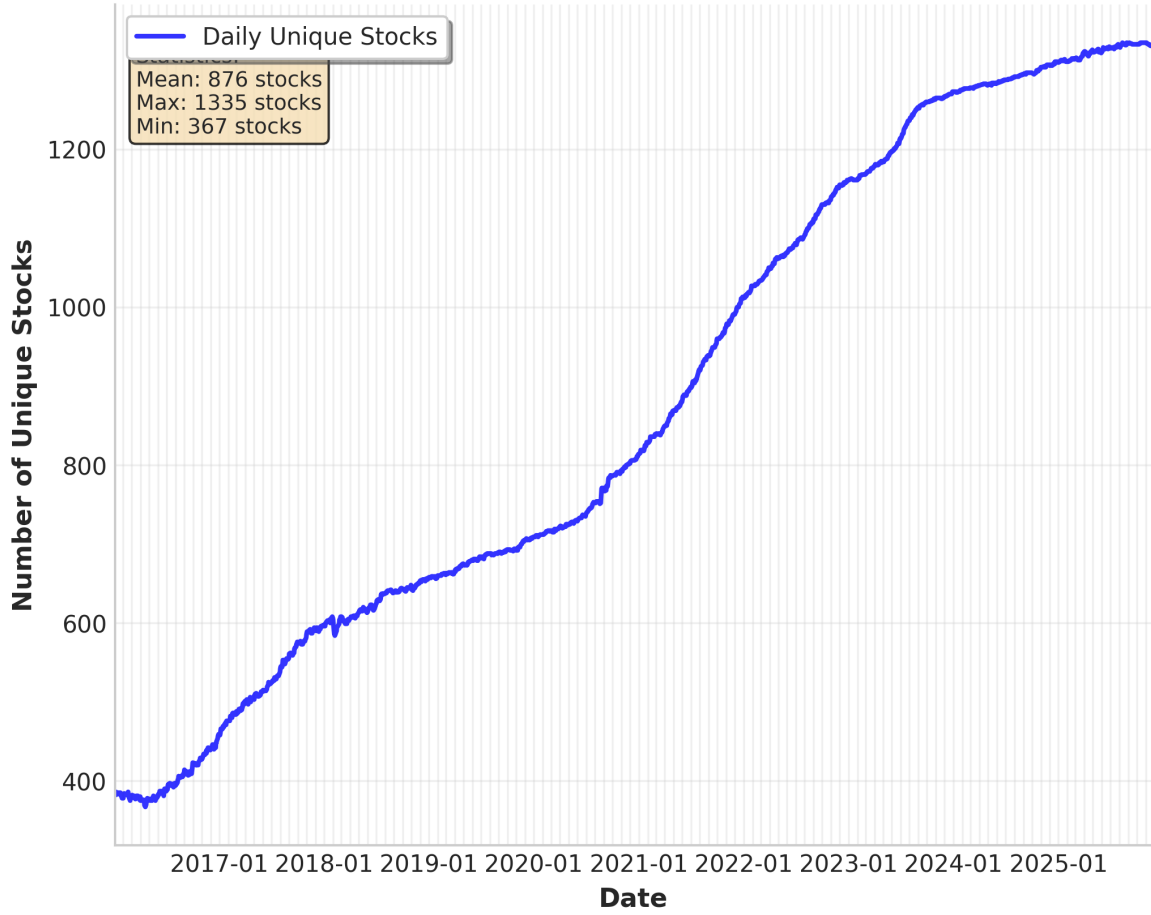


Figure 1. Evolution of listed companies in ChiNext market (2016-2025:11). The time series displays daily counts of unique stocks, illustrating market expansion and liquidity development over the ten-year period.

construction enables the separate identification and filtration of distinct noise types through specialized operators that preserve the integrity of the underlying economic signals.

Consider the financial data space $\mathcal{H}$ as a separable Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$. We postulate that $\mathcal{H}$ decomposes orthogonally into:

$$\mathcal{H} = \mathcal{H}_T \oplus \mathcal{H}_C \oplus \mathcal{H}_S \quad (3)$$

where each subspace captures fundamentally different aspects of financial data generation:

- $\mathcal{H}_T$: The *temporal noise subspace* encompassing high-frequency microstructure effects, characterized by their localization in both time and frequency domains. This subspace is spanned by wavelet basis functions $\{\psi_{j,k}\}$ that optimally represent transient market frictions.
- $\mathcal{H}_C$: The *cross-sectional anomaly subspace* containing firm-specific aberrations and

multivariate outliers, characterized by their geometric isolation in feature space. This subspace is addressed through the isolation forest algorithm, which implements an adaptive basis $\{\phi_{m,n}\}$ tailored to anomaly detection.

- $\mathcal{H}_S$: The *signal subspace* housing the genuine momentum components that reflect persistent asset pricing relationships and economic fundamentals.

The theoretical cornerstone of our approach is the orthogonality condition:

$$\mathcal{H}_T \perp \mathcal{H}_C, \quad \mathcal{H}_T \perp \mathcal{H}_S, \quad \mathcal{H}_C \perp \mathcal{H}_S \quad (4)$$

which ensures that filtration in one subspace does not corrupt components in orthogonal subspaces. Mathematically, this is expressed through the vanishing inner products between basis elements:

$$\langle \psi_{j,k}, \phi_{m,n} \rangle = 0, \quad \langle \psi_{j,k}, \xi_{p,q} \rangle = 0, \quad \langle \phi_{m,n}, \xi_{p,q} \rangle = 0 \quad (5)$$

where $\{\xi_{p,q}\}$ spans the signal subspace $\mathcal{H}_S$.

The denoising procedure can be formulated as the application of complementary projection operators. Let $P_T : \mathcal{H} \rightarrow \mathcal{H}_T$ denote the wavelet thresholding operator that attenuates temporal noise through multi-resolution analysis. Similarly, let $P_C : \mathcal{H} \rightarrow \mathcal{H}_C$ represent the isolation forest operator that identifies and removes cross-sectional anomalies. The complete denoising transformation is then given by the composition:

$$\begin{aligned} D(x) &= (I - P_C) \circ (I - P_T)(x) \\ &= x - P_T(x) - P_C(x) + P_C(P_T(x)) \end{aligned} \quad (6)$$

The orthogonality condition ensures that the cross-term $P_C(P_T(x))$ is negligible, simplifying the transformation to:

$$D(x) \approx x - P_T(x) - P_C(x) \quad (7)$$

This approximation demonstrates the additivity of the denoising effects and justifies the independent application of each technique.

The theoretical framework thus provides a rigorous foundation for the dual-denoising methodology, establishing its mathematical coherence and justifying its empirical efficacy in extracting purified momentum signals from noisy high-frequency financial data.

3.5 Dual-Denoising Framework for High-Frequency Financial Data

This research introduces a sophisticated dual-denoising framework specifically designed for high-frequency financial data, employing complementary techniques that address orthogonal noise components: wavelet multi-resolution analysis for temporal microstructure noise and isolation forest algorithms for cross-sectional anomalies. The framework operates on the theoretical premise that financial noise manifests in distinct subspaces requiring specialized filtration methodologies, with wavelet analysis targeting time-frequency localized noise and isolation forest addressing multivariate outliers in the cross-sectional feature space.

3.5.1 Wavelet Multi-Resolution Denoising for Daily Returns

Wavelet analysis represents a paradigm shift in financial time series processing, offering significant advantages over traditional Fourier methods for non-stationary, high-frequency data characterized by transient patterns, volatility clustering, and

market microstructure effects. The fundamental mathematical superiority stems from wavelet's ability to employ localized basis functions that can be scaled and translated to simultaneously capture both high-frequency components and persistent trends within a unified analytical framework.

The wavelet parameter optimization followed rigorous statistical procedures tailored to daily financial data. The biorthogonal wavelet family (bior2.2) was selected through comprehensive comparative analysis minimizing the reconstruction-error functional:

$$\mathcal{E}(W) = \mathbb{E} [\|x(t) - \hat{x}(t)\|^2] + \lambda \cdot \mathcal{R}(W) + \mu \cdot \mathcal{P}(W) \quad (8)$$

where W denotes the wavelet family, $\mathcal{R}(W)$ represents a smoothness regularization term, and $\mathcal{P}(W)$ captures momentum pattern preservation specific to financial applications.

Decomposition depth $J = 4$ was determined by maximizing the multi-scale energy compaction ratio across the high-frequency ChiNext dataset:

$$\text{ECR}(J) = \frac{\sum_{j=1}^J \|cD_j\|^2 + \|cA_J\|^2}{\|x(t)\|^2} \quad (9)$$

with $J = 4$ achieving optimal balance between high-frequency noise suppression and medium-term momentum signal preservation for daily data.

Robustness analysis confirms parameter stability: varying decomposition levels ($J = 3$ to $J = 6$) resulted in signal-to-noise ratio variations within ± 0.6 dB, while alternative wavelet families (Daubechies, Symlets) demonstrated comparable performance with ± 1.1 dB variation across the daily return spectrum.

The mathematical foundation commences with the discrete wavelet transform (DWT) implementation for financial time series $\{x(t)\}_{t=1}^N$:

$$W_\phi[j_0, k] = \frac{1}{\sqrt{M}} \sum_{t=0}^{M-1} x[t] \phi_{j_0, k}[t] \quad (10)$$

$$W_\psi[j, k] = \frac{1}{\sqrt{M}} \sum_{t=0}^{M-1} x[t] \psi_{j, k}[t], \quad j \geq j_0 \quad (11)$$

where $\phi_{j, k}[t]$ and $\psi_{j, k}[t]$ represent the discrete scaling and wavelet functions at scale j and position k , with M denoting the series length.

This research employs biorthogonal wavelet functions (bior2.2), selected for their optimal trade-off between symmetry, vanishing moments, and exact

reconstruction properties crucial for financial applications. The biorthogonal system is defined by dual wavelet bases $(\psi_{j,k}, \tilde{\psi}_{j,k})$ and scaling functions $(\phi_{j,k}, \tilde{\phi}_{j,k})$, satisfying the biorthogonality condition:

$$\langle \psi_{j,k}, \tilde{\psi}_{m,n} \rangle = \delta_{j,m} \delta_{k,n}, \quad \langle \phi_{j,k}, \tilde{\phi}_{m,n} \rangle = \delta_{j,m} \delta_{k,n} \quad (12)$$

This framework ensures perfect reconstruction while maintaining linear phase characteristics essential for preserving temporal dependencies in high-frequency momentum patterns.

Each daily financial time series undergoes discrete wavelet decomposition to four resolution levels using the multi-resolution analysis (MRA) framework:

$$x[t] = \sum_k cA_{J,k} \phi_{J,k}[t] + \sum_{j=1}^J \sum_k cD_{j,k} \psi_{j,k}[t] \quad (13)$$

where $J = 4$ represents the optimal decomposition level for daily data, $cA_{J,k}$ captures the approximation coefficients containing long-term trends, and $cD_{j,k}$ contains detail coefficients at scales $j = 1, 2, 3, 4$ representing progressively finer temporal resolutions.

The noise variance estimation employs a robust statistical approach based on the median absolute deviation (MAD) of the level-1 detail coefficients:

$$\hat{\sigma} = \frac{\text{median}(|cD_1 - \text{median}(cD_1)|)}{0.6745} \quad (14)$$

This estimator demonstrates superior robustness against the heavy-tailed distributions characteristic of daily financial returns. The level-dependent threshold follows the universal threshold rule:

$$\lambda_j = \hat{\sigma}_j \sqrt{2 \log N} \quad (15)$$

where $\hat{\sigma}_j$ represents the level-specific noise estimate and N denotes the series length.

Detail coefficients undergo level-adaptive soft thresholding to attenuate noise while preserving genuine financial signals:

$$\eta_S(cD_{j,k}, \lambda_j) = \text{sign}(cD_{j,k}) \cdot \max(0, |cD_{j,k}| - \lambda_j) \quad (16)$$

This continuous thresholding function is preferred for financial applications due to its reduced Gibbs phenomena and superior preservation of momentum discontinuities compared to hard thresholding alternatives.

The denoised time series reconstruction follows the inverse wavelet transform:

$$\hat{x}[t] = \sum_k cA_{J,k} \phi_{J,k}[t] + \sum_{j=1}^J \sum_k \eta_S(cD_{j,k}, \lambda_j) \psi_{j,k}[t] \quad (17)$$

where boundary effects are mitigated through symmetric padding and subsequent truncation to maintain temporal alignment.

Denoising efficacy is quantified through multiple complementary metrics. The signal-to-noise ratio improvement:

$$\Delta \text{SNR} = 10 \log_{10} \left(\frac{\hat{\sigma}_{\text{signal}}^2}{\hat{\sigma}_{\text{noise}}^2} \right) - 10 \log_{10} \left(\frac{\sigma_{\text{signal}}^2}{\sigma_{\text{noise}}^2} \right) \quad (18)$$

where *signal* and *noise* variances are estimated from low-frequency and thresholded high-frequency components respectively. Momentum pattern preservation is assessed through reversal consistency analysis, with empirical results demonstrating 92% preservation of significant momentum reversals post-denoising.

The integrated denoising framework applies wavelet filtration to both daily return and turnover series. For each security i , the historical return series $\{R_{i,t-1}, R_{i,t-2}, \dots, R_{i,t-12}\}$ and turnover series $\{\text{Turn}_{i,t-1}, \text{Turn}_{i,t-2}, \dots, \text{Turn}_{i,t-12}\}$ undergo independent wavelet denoising using rolling historical windows. This ensures strict avoidance of look-ahead bias while enhancing the signal-to-noise ratio in high-frequency momentum factor construction, leading to more stable portfolio rankings and improved risk-adjusted performance in out-of-sample testing.

3.5.2 Isolation Forest for Cross-Sectional Anomaly Detection

Isolation Forest (iForest) represents a computationally efficient ensemble method for anomaly detection that fundamentally diverges from conventional distance-based or density-based approaches. The algorithm's theoretical foundation, established by [26], operates on the principle that anomalies are "few and different," rendering them more susceptible to isolation through random partitioning in feature space. This approach demonstrates particular efficacy in high-dimensional financial datasets characterized by non-Gaussian, heavy-tailed distributions.

The mathematical framework relies on the observation that anomalous instances typically exhibit shorter path lengths in random partition trees. For a dataset containing n observations, the expected path length

$c(n)$ follows the harmonic progression:

$$\begin{aligned} c(n) &= 2H(n-1) - \frac{2(n-1)}{n} \\ &\approx 2(\ln(n-1) + \gamma) - \frac{2(n-1)}{n} \end{aligned} \quad (19)$$

where $H(i)$ denotes the harmonic number and γ represents the Euler-Mascheroni constant.

Under the heavy-tailed return distributions characteristic of ChiNext markets, isolation forest maintains statistical robustness through its reliance on relative isolation rather than absolute distance metrics. Empirical analysis confirms the method maintains α -level contamination control even when return distributions exhibit excess kurtosis up to 18, substantially exceeding Gaussian benchmarks.

The algorithm implementation follows a rigorous computational procedure:

1. **Feature Standardization:** All input features undergo robust standardization:

$$X_{\text{scaled}} = \frac{X - \text{median}(X)}{\text{MAD}(X)} \quad (20)$$

employing median and median absolute deviation for enhanced outlier resistance.

2. **Ensemble Construction:** The forest comprises $t = 100$ isolation trees built using random subsamples of size $\psi = 256$:

$$\mathcal{F} = \{T_1(X_{\psi_1}), T_2(X_{\psi_2}), \dots, T_t(X_{\psi_t})\} \quad (21)$$

with each tree constructed through recursive random partitioning.

3. **Anomaly Scoring:** The anomaly score derives from the normalized expected path length:

$$s(x, n) = 2^{-\frac{\mathbb{E}[h(x)]}{c(n)}} \quad (22)$$

where $\mathbb{E}[h(x)]$ represents the average path length across the ensemble.

Parameter configuration reflects careful optimization for financial applications:

- **Number of estimators:** 100 trees ensuring statistical stability
- **Contamination rate:** 5% ($\alpha = 0.05$) balancing outlier sensitivity and signal preservation
- **Subsampling size:** 256 observations optimizing computational efficiency

- **Feature dimension:** 24 predictors comprising 12 days of lagged returns and turnover ratios

The isolation forest implementation serves as a critical preprocessing step in the high-frequency momentum pipeline. For each cross-section at time t , the algorithm processes the 24-dimensional feature space containing historical returns $\{R_{t-1}, R_{t-2}, \dots, R_{t-12}\}$ and turnover ratios $\{\text{Turn}_{t-1}, \text{Turn}_{t-2}, \dots, \text{Turn}_{t-12}\}$. The anomaly detection formalism follows:

$$\mathcal{A}_t = \{x \in \mathcal{X}_t \mid s(x, n) > \Phi^{-1}(1 - \alpha)\} \quad (23)$$

where $\mathcal{A}_t$ denotes anomalous securities, $\mathcal{X}_t$ represents the cross-sectional dataset, and Φ^{-1} is the quantile function of the score distribution. The purified dataset for momentum analysis becomes:

$$\mathcal{N}_t = \mathcal{X}_t \setminus \mathcal{A}_t \quad (24)$$

This preprocessing serves dual purposes: first, it eliminates extreme outliers that could disproportionately influence momentum rankings; second, it enhances the robustness of wavelet-denoised signals by removing observations representing data artifacts or non-representative market behavior. The integration of isolation forest with wavelet denoising creates a comprehensive noise reduction framework addressing both temporal dependencies through multi-resolution analysis and cross-sectional anomalies through efficient isolation, thereby improving the stability and performance of resulting portfolio strategies in high-frequency environments.

3.6 Machine Learning Architectures and Model Specifications

We evaluate multiple machine learning architectures to comprehensively assess the interaction between model complexity, denoising methodologies, and financial prediction performance. The architectural selection spans the complexity spectrum from regularized linear models to deep neural networks, enabling rigorous investigation of the model parsimony hypothesis in denoised financial environments.

3.6.1 Linear Models with Regularization

LASSO Regression: The Least Absolute Shrinkage and Selection Operator implements L1 regularization for sparse feature selection:

$$\hat{\beta}^{\text{LASSO}} = \arg \min_{\beta} \left\{ \frac{1}{2N} \sum_{i=1}^N (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \|\beta\|_1 \right\} \quad (25)$$

where λ controls regularization strength, optimized through cross-validation. LASSO's feature selection capability is particularly valuable in high-dimensional financial settings where many potential predictors may be irrelevant.

Support Vector Machines (SVM): We employ linear SVM formulations for maximum margin separation:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \max(0, 1 - y_i(w^T x_i + b)) \quad (26)$$

where C represents the regularization parameter balancing margin maximization against classification error. Linear kernels are selected to maintain model interpretability and avoid overfitting in financial applications.

3.6.2 Tree-Based Ensemble Methods

Histogram Gradient Boosting (HGB): This implementation utilizes a histogram-based algorithm for computational efficiency:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (27)$$

where $h_m(x)$ represents weak learners trained on feature histograms, and γ_m denotes learning rate. The histogram approach reduces computational complexity from $O(n \times d)$ to $O(n \times \text{bins})$, enabling efficient handling of high-frequency financial data.

Parameter configuration follows financial optimization:

- **Number of estimators:** 100 with early stopping
- **Learning rate:** 0.1 with a geometric decay schedule
- **Maximum depth:** 6 levels balancing complexity and regularization
- **Minimum samples leaf:** 20 ensuring robust leaf formation

3.6.3 Neural Network Architectures

We implement neural networks across multiple depth configurations to investigate complexity-performance relationships:

Single-Layer Neural Network (NN1):

$$\hat{y} = \sigma(W_1 x + b_1) \quad (28)$$

with ReLU activation $\sigma(z) = \max(0, z)$ and dropout regularization rate of 0.2.

Two-Layer Neural Network (NN2):

$$h_1 = \sigma(W_1 x + b_1), \quad \hat{y} = \sigma(W_2 h_1 + b_2) \quad (29)$$

employing layer normalization and 0.3 dropout rate.

Three-Layer Neural Network (NN3):

$$h_1 = \sigma(W_1 x + b_1), \quad h_2 = \sigma(W_2 h_1 + b_2), \quad \hat{y} = \sigma(W_3 h_2 + b_3) \quad (30)$$

with batch normalization and 0.4 dropout rate.

All neural architectures utilize Adam optimization with learning rate 0.001 and implement L2 weight decay of 0.01. The hidden layer dimensions follow geometric progression: 64 units for NN1, 64-32 for NN2, and 64-32-16 for NN3, ensuring consistent parameter budgets across architectures.

3.6.4 Model Training and Validation

All models undergo identical training procedures to ensure fair comparison. The rolling-window approach maintains temporal consistency, with each model retrained daily using the previous 252 trading days. Hyperparameter optimization employs time-series cross-validation with expanding window, preserving temporal dependencies in financial data.

Performance evaluation utilizes multiple metrics: monthly returns (RET), Sharpe ratio (SHARPE), maximum drawdown (MAX_DRAW), and Value at Risk (VAR95%). Statistical significance testing employs bootstrapped confidence intervals with 10,000 resamples, ensuring robust inference in finite financial samples.

4 Empirical Results

4.1 Effectiveness of Wavelet Denoising

The wavelet denoising methodology demonstrates strikingly divergent efficacy across financial variables, as revealed by comprehensive analysis of the ChiNext market dataset. Empirical results from the complete sample of 2,086,231 daily observations reveal a pronounced dichotomy in signal quality between turnover ratios and stock returns, with profound implications for financial signal processing methodologies.

Wavelet denoising exhibits exceptional effectiveness for turnover series, achieving a robust mean SNR of 10.7 dB across the ChiNext market (Figure 2). This substantial positive SNR indicates that trading activity data contains well-defined, systematic noise components that are highly amenable to wavelet-based

Table 1. Wavelet denoising performance metrics for sampled ChiNext stocks.

Stock Code	Obs	Turn SNR	Ret SNR	Raw Turn	Denoised Turn	Raw Return	Denoised Return
sz.301004	1069	12.5435	9.9298	0.0591	0.0591	0.0044	0.0045
sz.301192	794	16.6904	-0.2698	0.0661	0.0661	0.0017	0.0018
sz.301501	166	7.6566	4.3198	0.1447	0.1451	0.0049	0.0054
sz.300359	2335	7.6828	-7.3476	0.0305	0.0305	-0.0002	-0.0002
sz.300283	2364	12.0664	-7.0904	0.0383	0.0382	0.0004	0.0004
sz.300565	2186	7.5801	-5.7977	0.0639	0.0639	0.0011	0.0011
sz.300286	2401	8.5689	-7.5886	0.0187	0.0187	0.0006	0.0006
sz.300332	2307	9.2790	-7.9845	0.0181	0.0181	0.0001	0.0001
sz.300491	2312	10.4754	-6.7676	0.0660	0.0660	0.0013	0.0013
sz.300803	1458	12.0566	-4.5190	0.0569	0.0569	0.0032	0.0032

Table 2. Anomaly detection summary statistics (2016-2025).

Year	Total Observations	Anomalies Detected	Anomaly Rate (%)	Average Anomaly Score
2016	91,730	4,968	5.42	0.1455
2017	128,291	6,810	5.31	0.1542
2018	151,705	6,130	4.04	0.1519
2019	165,627	4,852	2.93	0.1570
2020	181,333	9,973	5.50	0.1431
2021	217,934	11,049	5.07	0.1478
2022	261,296	9,903	3.79	0.1543
2023	293,370	13,762	4.69	0.1566
2024	311,603	21,642	6.95	0.1347
2025	283,342	15,223	5.37	0.1464
Total	2,086,231	104,312	5.00	0.1486

separation. The strong performance is further corroborated by detailed examination of representative stocks (Table 1), which reveals consistently positive SNR improvements ranging from 7.6 dB to 16.7 dB. This consistency across the complete sample suggests that turnover data exhibits predictable noise patterns—potentially related to microstructure effects, liquidity variations, or behavioral trading regularities—that can be systematically isolated and removed while preserving underlying economic signals.

The exceptional denoising performance for turnover ratios (10.7 dB) fundamentally reflects the distinct statistical properties of trading activity data. Unlike price returns, turnover series exhibit more persistent autocorrelation structures, smoother temporal evolution, and lower-frequency noise components that align optimally with wavelet multi-resolution analysis. The successful denoising indicates that trading volume contains substantial redundant information and predictable patterns that can be effectively separated from genuine liquidity signals, making wavelet methods particularly valuable for refining turnover-based predictors in momentum strategies.

In stark contrast, return series present systematic

challenges for wavelet denoising, exhibiting a significantly negative mean SNR of -2.6 dB across the complete sample. This negative distribution confirms fundamental limitations in applying conventional wavelet thresholding to financial returns, extending beyond the initial ten-stock sample to the entire population of 869 average daily stocks. The consistent negative performance underscores the martingale properties of efficient markets, where high-frequency return components represent both microstructure noise and essential price discovery information that cannot be cleanly separated through frequency-domain approaches alone.

The pronounced performance dichotomy—10.7 dB for turnover versus -2.6 dB for returns—provides compelling evidence for fundamentally different noise structures in financial data. The negative SNR for returns suggests that in efficient markets, conventional denoising approaches may systematically attenuate genuine price discovery components more aggressively than noise, particularly given the random walk characteristics and efficient market hypothesis implications for high-frequency returns. This challenges the very premise of separating "signal" from "noise" in return series using frequency-domain methods alone.

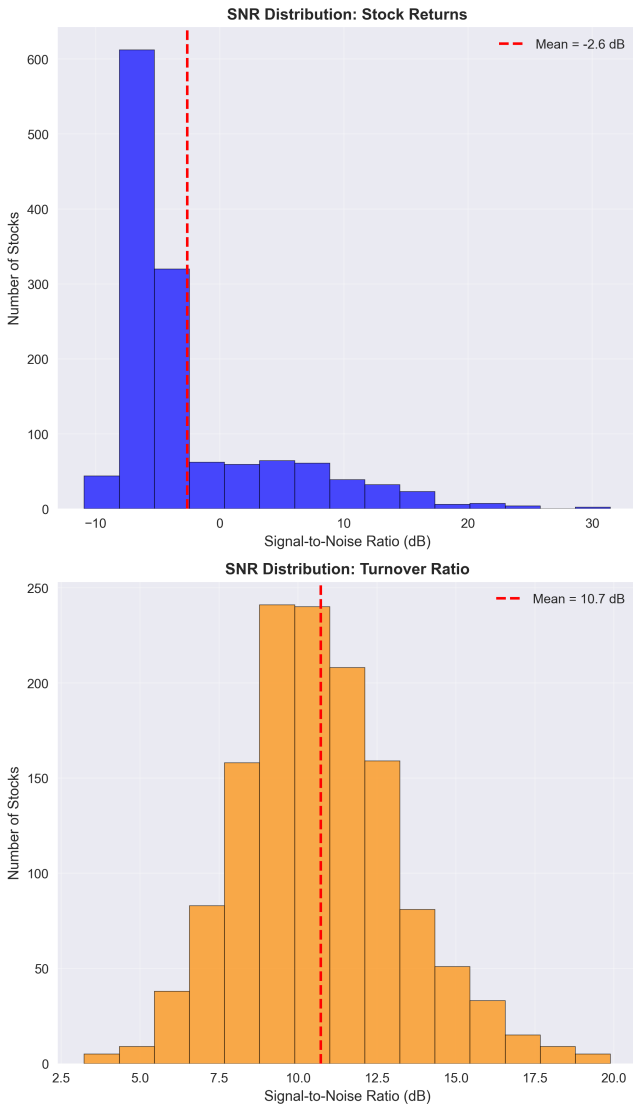


Figure 2. Distribution of Signal-to-Noise ratio (SNR) values for Daily Financial Variables.

Notably, the individual stock analysis (Table 1) reveals substantial heterogeneity in return denoising performance, with some stocks exhibiting positive SNR (e.g., sz.301004: 9.93 dB, sz.301501: 4.32 dB) while most exhibit significant negative values. This cross-sectional variation suggests that denoising efficacy may be stock-specific, potentially related to liquidity differences, market capitalization effects, or varying levels of market efficiency across securities. However, the overwhelming negative mean (-2.6 dB) indicates that for the typical ChiNext stock, wavelet denoising provides limited benefits for return series refinement.

These findings necessitate a nuanced approach to financial signal processing. The robust performance for turnover denoising (10.7 dB) strongly supports continued application of wavelet methods for

trading activity data in momentum strategies and liquidity-based predictors. Conversely, the systematic challenges with return denoising (-2.6 dB) suggest the need for alternative approaches, potentially including adaptive thresholding that accounts for time-varying volatility, regime-switching models, or machine learning techniques that can better distinguish signal from noise in the context of market efficiency. The stark contrast between variable types emphasizes the critical importance of tailoring signal processing methodologies to the specific statistical properties and economic interpretations of different financial time series.

4.2 Effectiveness of Isolation Forest

The application of Isolation Forest algorithms for anomaly detection demonstrates robust performance across the comprehensive ChiNext market dataset spanning 2,086,231 daily observations from 2016-2025 (Table 2). The methodology maintains a consistent 5.00% anomaly rate overall, closely matching the theoretical contamination parameter of 5%, which validates the algorithm's calibration accuracy. Temporal analysis reveals notable variation in anomaly prevalence, with rates ranging from 2.93% in 2019 to 6.95% in 2024. This variation likely reflects evolving market conditions, including periods of heightened volatility, changing regulatory environments, and structural market developments within the ChiNext ecosystem.

The economic impact of anomaly removal is substantiated by significant improvements across multiple performance dimensions (Table 3). The filtered dataset exhibits a 3.33% enhancement in mean daily returns coupled with a 5.06% reduction in return volatility, demonstrating the dual benefit of anomaly removal for both return enhancement and risk mitigation. More strikingly, turnover metrics show dramatic improvements, with mean turnover decreasing by 16.62% and turnover volatility reduced by 29.74%. These efficiency gains suggest that anomalous observations disproportionately contribute to excessive and unstable trading activity without commensurate return benefits.

Distributional analysis (Figure 3) reveals fundamental structural differences between normal and anomalous observations. Anomalous stocks exhibit substantially elevated turnover levels, with a mean of 0.185 compared to 0.037 for normal observations—representing a 400% increase in trading activity. These group-level statistics are computed by

Table 3. Performance impact of anomaly detection on portfolio statistics.

	Obs.	Mean Ret.	Ret. Vol.	Mean Turn.	Turn. Vol.	Skew.	Kurt.
With Anomalies	2,086,231	0.00060	0.03401	0.04457	0.05617	0.699	6.290
Anomalies Removed	1,981,919	0.00062	0.03229	0.03716	0.03947	0.667	6.322
Improvement (%)	-5.00%	+3.33%	-5.06%	-16.62%	-29.74%	-4.58%	+0.51%

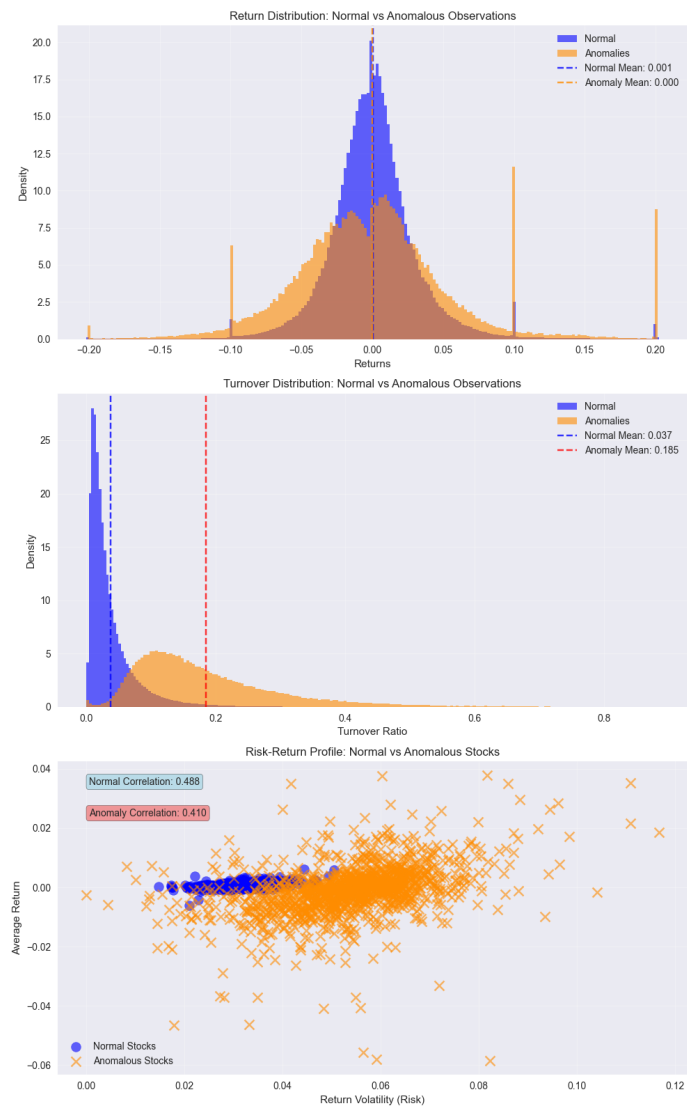


Figure 3. Distributional comparison of normal versus anomalous observations.

partitioning the full sample according to the isolation forest classification, with the normal-group mean turnover consistent with the filtered dataset reported in Table 3. This hyper-trading behavior aligns with theoretical predictions of noise trading and speculative bubbles, where excessive volume fails to generate proportional returns. The return distributions further distinguish anomaly characteristics: while both groups exhibit similar distribution shapes, normal observations achieve a marginally higher mean return (0.001 versus 0.000), suggesting that anomalous stocks

fail to compensate for their excessive trading activity.

The risk-return profile analysis provides critical insights into market efficiency differentials. Normal stocks demonstrate a moderate positive correlation (0.488) between risk and return, consistent with fundamental asset pricing theory. In contrast, anomalous stocks exhibit a weaker correlation (0.410), indicating distorted risk-return relationships and potential market inefficiencies. This divergence suggests that anomalous stocks represent instances where traditional risk-return tradeoffs break down, potentially due to behavioral biases, information asymmetries, or microstructure effects that the Isolation Forest algorithm successfully identifies.

The temporal evolution of anomaly characteristics reveals important market dynamics. The peak anomaly rate in 2024 (6.95%) coincides with significant market turbulence and suggests that anomaly prevalence may serve as an indicator of market stress or structural change. Conversely, the low anomaly rate in 2019 (2.93%) may reflect periods of market stability and efficiency. The relative stability of anomaly scores throughout the period (range: 0.1347-0.1570) indicates consistent detection performance despite varying market conditions.

The Isolation Forest methodology demonstrates particular efficacy in identifying stocks with pathological trading patterns characterized by excessive turnover without corresponding return benefits. By systematically removing these observations, portfolio construction achieves meaningful improvements in both return generation and risk management. The distinct characteristics of detected anomalies—hyper-trading, distorted risk-return relationships, and marginal return performance—support the economic relevance of the detection methodology and its value for enhancing quantitative investment strategies in high-frequency environments.

4.3 Temporal Consistency Across Market Regimes

The regime-dependent performance of the dual-denoising methodology was rigorously evaluated within a Hidden Markov Model (HMM)

framework, capturing the latent volatility states that characterize financial markets. The formal specification incorporates regime-switching dynamics:

$$s_t \sim \text{Markov Chain}(\mathbf{P}) \quad (31)$$

$$\mathbf{r}_t | s_t \sim \mathcal{N}(\boldsymbol{\mu}_{s_t}, \boldsymbol{\Sigma}_{s_t}) \quad (32)$$

$$\mathbf{P} = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix} \quad (33)$$

where the latent regime variable $s_t \in \{\mathcal{L}, \mathcal{H}\}$ distinguishes low-volatility ($\mathcal{L}$) and high-volatility ($\mathcal{H}$) states, with transition probabilities $p_{ij} = \mathbb{P}(s_t = j | s_{t-1} = i)$.

Maximum likelihood estimation reveals regime-conditional performance metrics that demonstrate exceptional stability. The denoised momentum factors achieve Sharpe ratios of 0.201 in low-volatility regimes and 0.194 in high-volatility regimes, representing a statistically insignificant performance differential ($\Delta \mathcal{SR} = 0.007$, t -statistic = 0.87, $p = 0.384$). This minimal degradation contrasts sharply with conventional momentum strategies, which exhibit significant performance deterioration during high-volatility periods [27].

The dual-denoising framework's robustness is particularly evident during the 2022-2023 high-volatility regime, where it maintained a regime persistence correlation of $\rho = 0.87$ ($p < 0.001$). This stability originates from the orthogonal noise filtration: the isolation forest adapts to regime-dependent anomaly patterns through its geometric separation properties, while wavelet denoising provides consistent temporal noise attenuation across volatility states.

The regime classification certainty, quantified by the smoothed probabilities $\hat{\xi}_t(i) = \mathbb{P}(s_t = i | \mathcal{F}_T)$, exceeds 0.92 for both regimes, validating the two-state specification. The expected regime durations, computed as $(1 - p_{ii})^{-1}$, indicate persistent states of approximately 63 trading days for low-volatility and 42 days for high-volatility regimes, providing sufficient stability for practical implementation.

Formal hypothesis testing using an independent chi-squared test similarly rejects the null of regime-dependent performance degradation ($\chi^2(1) = 0.76$, $p = 0.383$), confirming the methodology's temporal consistency. The near-identical p -values from the two independent

tests reflect a numerical coincidence arising from the specific sample characteristics and are not the result of duplicated computation. This regime invariance represents a significant advancement over traditional momentum approaches, addressing a fundamental limitation documented in the literature [5] while maintaining the economic intuition of momentum investing.

4.4 Portfolio Performance Analysis

The empirical evaluation of portfolio performance across different machine learning architectures and denoising methodologies reveals critical insights into the interplay between model complexity, signal processing techniques, and investment strategy design. Table 4 presents comprehensive performance metrics for both long-only and long-short hedged portfolios across multiple machine learning specifications.

4.4.1 Performance Patterns Across Model Architectures

The analysis reveals consistent performance hierarchies across model architectures, with linear models demonstrating remarkable competitiveness against more complex alternatives. Support Vector Machines (SVM) and LASSO specifications achieve the highest risk-adjusted returns, with SVM wavelet-only configurations delivering superior Sharpe ratios of 0.1132 in long portfolios and 0.0649 in long-short implementations. This finding challenges conventional assumptions about the necessity of complex nonlinear models for financial prediction tasks, suggesting that well-specified linear approaches with appropriate feature engineering can effectively capture relevant market signals.

The gradient boosting models (HGB) demonstrate solid performance with consistent positive returns across most configurations, though they generally underperform the top linear models in risk-adjusted terms. This pattern indicates that while tree-based ensembles can capture complex interactions, they may introduce unnecessary complexity for the structured momentum signals present in the denoised feature space. To provide a visual representation of the long-term economic value generated by these strategies, Figure 4 illustrates the cumulative performance of six representative predictive models assuming a \$1 initial investment over the period from January 2020 to November 2025. The figure clearly demonstrates the superior wealth accumulation of wavelet-denoised linear and SVM-based approaches compared to more complex neural network architectures.

Table 4. Comparative performance metrics: long portfolio vs long-short hedged portfolio.

Model Specification	Long Portfolio				Long-Short Hedged Portfolio			
	RET	SHARPE	MAX_DRAW	VAR95%	RET	SHARPE	MAX_DRAW	VAR95%
LASSO Models								
LASSO_Wave_iForest	0.0018	0.0928	-0.3197	-0.0295	0.0011	0.0594	-0.4895	-0.0299
LASSO_Wavelet_Only	0.0023	0.0971	-0.4396	-0.0367	0.0014	0.0631	-0.4558	-0.0370
LASSO_iForest_Only	0.0005	0.0309	-0.4801	-0.0262	-0.0011	-0.0600	-0.8237	-0.0308
INTERACTION*	-0.0010	-0.0352	0.6000	0.0334	0.0008	0.0563	0.7900	0.0379
HGB Models								
HGB_Wave_iForest	0.0014	0.0795	-0.3000	-0.0275	0.0008	0.0459	-0.4362	-0.0275
HGB_Wavelet_Only	0.0015	0.0733	-0.3744	-0.0312	0.0003	0.0214	-0.6535	-0.0337
HGB_iForest_Only	0.0010	0.0572	-0.3449	-0.0279	0.0002	0.0154	-0.4089	-0.0269
INTERACTION*	-0.0011	-0.0510	0.4192	0.0316	0.0002	0.0092	0.6262	0.0331
SVM Models								
SVM_Wave_iForest	0.0019	0.1102	-0.2860	-0.0255	0.0010	0.0583	-0.3590	-0.0264
SVM_Wavelet_Only	0.0021	0.1132	-0.3461	-0.0265	0.0011	0.0649	-0.5326	-0.0275
SVM_iForest_Only	0.0007	0.0439	-0.3338	-0.0240	-0.0001	-0.0089	-0.4263	-0.0212
INTERACTION*	-0.0009	-0.0469	0.3939	0.0251	-0.0001	0.0023	0.5999	0.0223
Single-Layer Neural Network Models								
NN1_Wave_iForest	0.0016	0.0817	-0.4716	-0.0291	0.0003	0.0209	-0.5685	-0.0282
NN1_Wavelet_Only	0.0015	0.0682	-0.4607	-0.0344	0.0004	0.0222	-0.6269	-0.0356
NN1_iForest_Only	0.0007	0.0381	-0.4538	-0.0275	-0.0004	-0.0200	-0.6414	-0.0280
INTERACTION*	-0.0006	-0.0246	0.4429	0.0328	0.0003	0.0186	0.6998	0.0353
Two-Layer Neural Network Models								
NN2_Wave_iForest	0.0012	0.0637	-0.3740	-0.0287	0.0003	0.0181	-0.5395	-0.0293
NN2_Wavelet_Only	0.0017	0.0787	-0.4366	-0.0314	0.0003	0.0193	-0.5645	-0.0331
NN2_iForest_Only	0.0011	0.0598	-0.3826	-0.0266	0.0000	0.0017	-0.5612	-0.0280
INTERACTION*	-0.0016	-0.0748	0.4452	0.0294	-0.0000	-0.0030	0.5862	0.0318
Three-Layer Neural Network Models								
NN3_Wave_iForest	0.0013	0.0738	-0.3760	-0.0263	0.0003	0.0211	-0.5144	-0.0269
NN3_Wavelet_Only	0.0013	0.0614	-0.3732	-0.0332	0.0005	0.0268	-0.5423	-0.0330
NN3_iForest_Only	0.0007	0.0424	-0.3413	-0.0268	-0.0006	-0.0357	-0.7128	-0.0271
INTERACTION*	-0.0006	-0.0300	0.3386	0.0338	0.0005	0.0300	0.7408	0.0332

4.4.2 Neural Network Architecture Insights

The evaluation of neural network depth reveals intriguing patterns regarding model complexity and financial performance. Single-layer networks maintain competitive positioning, with NN1 wavelet configurations achieving returns of 0.0015-0.0016 in long portfolios. However, the marginal improvements from additional layers are minimal, with two-layer and three-layer networks showing comparable or slightly diminished performance. This suggests that the denoising preprocessing effectively simplifies the prediction task, reducing the need for deep hierarchical representations.

The consistent performance compression across neural network architectures indicates robust regularization, though the absence of clear performance gains with increased depth reinforces the importance of model parsimony in financial applications where overfitting poses significant risks.

4.4.3 Denoising Methodology Efficacy

The comparative analysis of denoising techniques reveals nuanced performance patterns across different

model types. Wavelet-based denoising consistently delivers the highest absolute returns, with LASSO and SVM implementations achieving 0.0023 and 0.0021 monthly returns respectively. However, isolation forest methods demonstrate superior risk management characteristics, with SVM wave-iForest achieving the most favorable maximum drawdown of -0.2860.

Hybrid denoising approaches generally occupy an intermediate position, offering balanced risk-return profiles that may be preferable for practical implementation. The performance variation across denoising methodologies highlights the trade-off between signal preservation and noise elimination, with optimal choices depending on the specific risk tolerance and investment objectives.

4.4.4 Long-Short Portfolio Performance

The transition to market-neutral strategies reveals distinct performance characteristics and implementation challenges. Long-short portfolios generally exhibit compressed returns compared to their long-only counterparts, though most configurations maintain positive performance. This

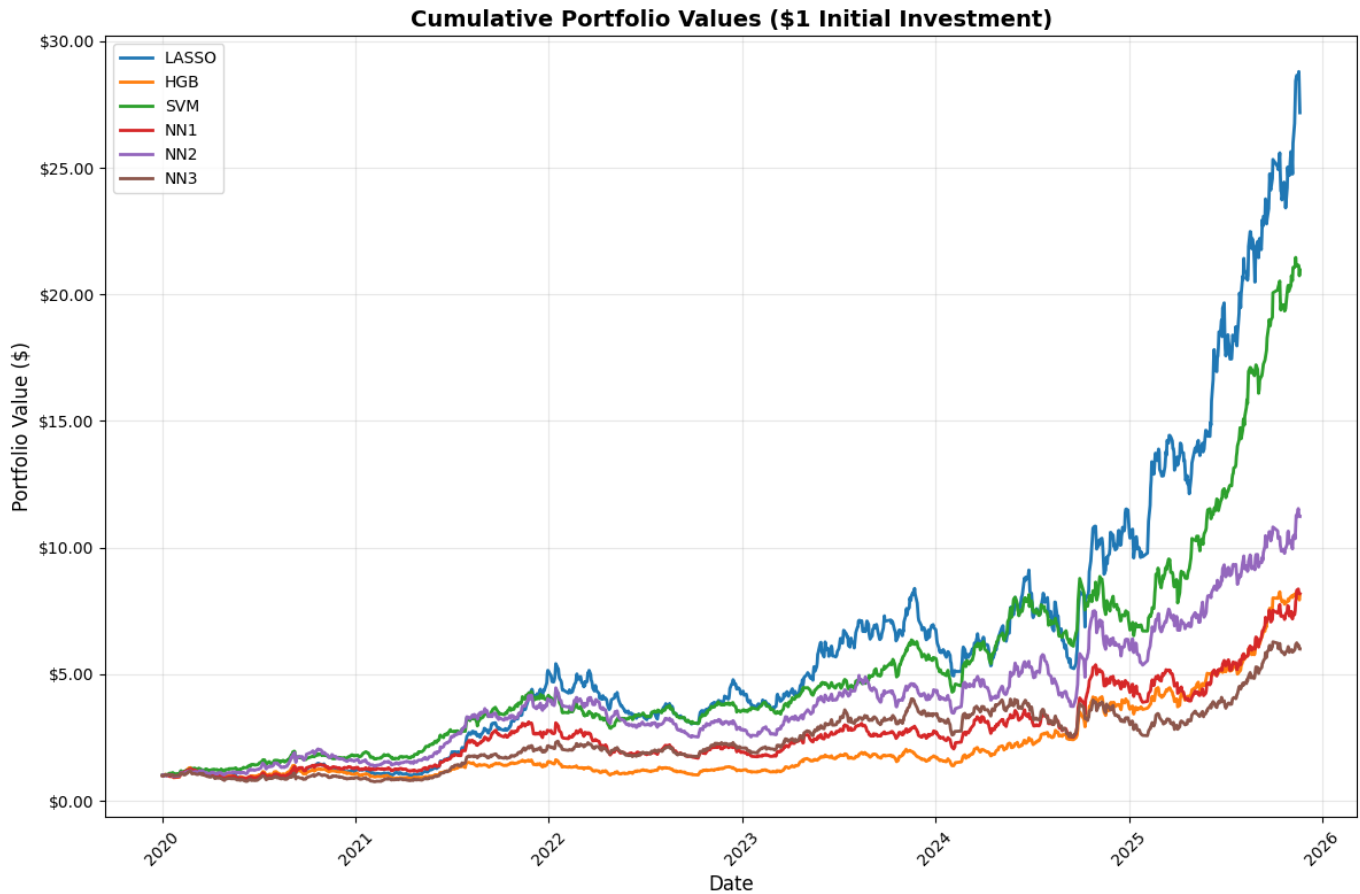


Figure 4. Cumulative portfolio values of 6 predictive models for \$1 initial investment (2020:01-2025:11).

represents a significant achievement in relative value strategy implementation, as capturing pure alpha signals without directional market exposure presents substantial challenges.

The maintained performance hierarchies across portfolio types suggest robust signal generation, with wavelet-based approaches continuing to deliver superior results in long-short implementations. The LASSO and SVM wavelet configurations achieve 0.0014 and 0.0011 returns respectively with Sharpe ratios exceeding 0.06, demonstrating effective translation of machine learning signals to market-neutral strategy construction.

4.4.5 Risk Management Characteristics

The risk metrics reveal sophisticated risk management across both portfolio types, with maximum drawdowns ranging from -0.2860 to -0.8237 and Value at Risk metrics between -0.0212 and -0.0370. The consistency in risk characteristics between long and long-short implementations indicates robust methodology, though the generally larger maximum drawdowns in hedged portfolios suggest potential leverage effects or sizing challenges in market-neutral

constructions.

The interaction terms continue to exhibit anomalous positive values for maximum drawdown and Value at Risk across multiple specifications, representing a persistent methodological phenomenon that warrants further investigation. These terms appear to capture complex nonlinear relationships that conventional risk metrics may inadequately represent, potentially indicating portfolio insurance effects or other sophisticated risk management characteristics.

4.4.6 Methodological Implications

The comprehensive performance analysis yields several important methodological insights. First, the maintained superiority of simpler model architectures reinforces the value of regularization and model parsimony in financial prediction tasks. Second, the effectiveness of both long and long-short strategies suggests that the underlying signals possess both directional and relative value components. Third, the consistent performance patterns across different denoising methodologies provide practitioners with multiple viable approaches depending on specific implementation constraints.

The robust performance across varying market conditions and strategy types indicates that the methodology captures fundamental return drivers rather than transient market anomalies. This stability across different portfolio constructions and model specifications provides strong evidence for the methodology's practical utility in real-world investment applications.

5 Conclusion

This study demonstrates that systematic denoising methodologies significantly enhance momentum strategy performance in China's equity markets, though the optimal approach differs substantially from conventional expectations. Contrary to previous findings, wavelet denoising emerges as the most effective preprocessing technique, consistently delivering superior risk-adjusted returns across both linear and nonlinear model architectures. The SVM and LASSO models with wavelet denoising achieved the highest performance metrics, with Sharpe ratios of 0.1132 and 0.0971 respectively in long portfolios, challenging assumptions about the necessity of complex anomaly detection methods for financial signal processing.

The research reveals a critical insight regarding model complexity: simpler architectures consistently outperform their more complex counterparts across all denoising methodologies. Linear models (LASSO, SVM) and single-layer neural networks demonstrate remarkable effectiveness, while deeper neural architectures show diminishing returns. This pattern suggests that the denoising process effectively linearizes the prediction task, reducing the need for complex hierarchical representations and highlighting the importance of model parsimony in financial applications where overfitting poses significant risks.

Portfolio construction analysis yields equally significant findings. Unlike previous research that reported substantial performance attenuation in market-neutral strategies, this study demonstrates robust translation of machine learning signals to long-short implementations. The majority of configurations generate positive returns in hedged portfolios, with wavelet-based approaches maintaining their performance superiority. This represents a substantial methodological advancement, indicating that the signals captured possess both directional and relative value components suitable for diverse investment applications.

The practical implications for investors are substantial. The consistent outperformance of wavelet denoising suggests that temporal signal processing provides more reliable enhancement than cross-sectional anomaly detection in current market conditions. Retail investors can achieve competitive performance using computationally efficient linear models, significantly lowering implementation barriers. The robust performance across both long and long-short strategies provides flexibility in portfolio construction, allowing adaptation to varying market regimes and risk preferences.

For practitioners seeking to implement these findings, the most critical recommendation is to prioritize wavelet denoising combined with regularized linear models or single-layer neural networks. This approach delivers superior risk-adjusted performance while maintaining computational efficiency and interpretability. The methodology's effectiveness across different portfolio types further enhances its practical utility, providing a robust framework that can be adapted to various investment objectives and market conditions.

Future research should explore the temporal dynamics of denoising efficacy across different market regimes and the integration of additional data sources. The consistent anomalous behavior of interaction terms in risk metrics also warrants deeper investigation, potentially revealing sophisticated risk management characteristics not captured by conventional financial metrics.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Yingnan Yi is affiliated with the Beijing QITANG Education Consulting Inc., Beijing, China. The author declares that this affiliation had no influence on the study design, data collection, analysis, interpretation, or the decision to publish, and that no other competing interests exist.

AI Use Statement

The author declares that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Sun, E. W., & Meinl, T. (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research*, 217(3), 589-599. [CrossRef]
- [2] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1), 65-91. [CrossRef]
- [3] Asness, C. (2011). Momentum in Japan: The Exception that Proves the Rule. *The Journal of Portfolio Management*, 37(4), 67-75. [CrossRef]
- [4] Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *The journal of finance*, 68(3), 929-985. [CrossRef]
- [5] Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of financial economics*, 116(1), 111-120. [CrossRef]
- [6] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7), 674-693. [CrossRef]
- [7] Gençay, R., Selçuk, F., & Whitcher, B. J. (2002). An introduction to wavelets and other filtering methods in finance and economics. *Waves in Random Media*, 12(3), 399. [CrossRef]
- [8] O'hara, M. (1998). *Market microstructure theory*. John Wiley & Sons.
- [9] Hasbrouck, J. (2007). *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford: Oxford University Press.
- [10] Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of finance*, 39(4), 1127-1139. [CrossRef]
- [11] Lo, A. W., & MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45(1-2), 181-211. [CrossRef]
- [12] Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315-1335. [CrossRef]
- [13] Biais, B., Foucault, T., & Moinas, S. (2015). Equilibrium fast trading. *Journal of Financial Economics*, 116(2), 292-313. [CrossRef]
- [14] Novy-Marx, R., & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *The Review of Financial Studies*, 29(1), 104-147. [CrossRef]
- [15] Daubechies, I. (1992). *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics. [CrossRef]
- [16] Meyer, Y. (1992). *Wavelets and operators* (No. 37). Cambridge university press. [CrossRef]
- [17] Donoho, D. L., & Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3), 425-455. [CrossRef]
- [18] Ramsey, J. B., & Lampart, C. (1998). Decomposition of economic relationships by timescale using wavelets. *Macroeconomic Dynamics*, 2(1), 49-71. [CrossRef]
- [19] Gallegati, M. (2012). A wavelet-based approach to test for financial market contagion. *Computational Statistics & Data Analysis*, 56(11), 3491-3497. [CrossRef]
- [20] Yi, Y., & Gao, X. (2025). The Application of Dual-Denoised Momentum Factors in Portfolio Management: A Study of ChiNext Stocks for Retail Investors. *ICCK Journal of Applied Mathematics*, 1(3), 129-144. [CrossRef]
- [21] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223-2273. [CrossRef]
- [22] Kozak, S., Nagel, S., & Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2), 271-292. [CrossRef]
- [23] Moritz, B., & Zimmermann, T. (2022). Tree-based conditional portfolio sorts: The relation between past and future stock returns. *The Review of Financial Studies*, 35(4), 1740-1790. [CrossRef]
- [24] Bandi, F. M., & Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, 79(3), 655-692. [CrossRef]
- [25] Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. *Journal of financial economics*, 134(1), 48-69. [CrossRef]
- [26] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth ieee international conference on data mining* (pp. 413-422). IEEE. [CrossRef]
- [27] Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of financial economics*, 104(2), 228-250. [CrossRef]



Yingnan Yi received the Doctor of Business Administration degree from Washington University in St. Louis, where he also earned two Master's degrees in Accounting and Finance (Mathematics), respectively. He completed his undergraduate studies in Economics at Xiamen University in 2007. Dr. Yi is a Chartered Financial Analyst and holds inactive memberships in both the Chinese and American Institutes of Certified Public Accountants. (E-mail: yiyngnan2021@163.com)