My Paper Title

Your Name Here
Your Affiliation

***Abstract*—Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.**

***Index Terms*—ETF selection, portfolio optimization, ensemble methods, mean reversion, momentum, alpha generation**

I. Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

II. Related Work

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

III. Methodology

This section presents our systematic approach for ETF selection in a core-satellite portfolio framework. We develop a quantitative signal-based methodology that identifies satellite ETFs with the highest probability of generating positive alpha relative to a global market benchmark.

A. Portfolio Framework

We adopt a core-satellite portfolio structure that combines the stability of passive global diversification with the alpha-generation potential of tactical satellite positions. The core allocation consists of a global market-cap weighted index (iShares MSCI ACWI), representing approximately 3,000 stocks across developed and emerging markets. This provides broad exposure while serving as our benchmark for measuring excess returns.

The satellite allocation dynamically selects from a filtered universe of equity ETFs based on predictive signals. We allocate up to 40% to satellites, distributed equally among a maximum of four selected ETFs. When no satellites demonstrate sufficient alpha potential, the strategy defaults to 100% core allocation, ensuring we never force suboptimal positions.

B. ETF Universe Construction

Starting from a catalog of over 7,500 ETFs, we apply a series of practical filters to construct our investment universe. We restrict to Irish-domiciled ETFs for tax efficiency, require accumulating distribution policies to avoid dividend drag, and impose a maximum total expense ratio of 0.30%. To ensure adequate liquidity and data availability, we require a minimum fund size of 100 million EUR and at least five years of price history. These filters yield approximately 500 eligible ETFs that form our selection universe.

C. Signal Base Computation

The foundation of our selection methodology lies in computing a comprehensive set of 167 signal bases from raw price data. We organize these into 27 categories, each capturing different aspects of ETF behavior relative to the core benchmark. Most signals use day-based window parameters (21d, 63d, 126d, 252d) corresponding approximately to 1, 3, 6, and 12 months respectively, with trend indicators preserving industry-standard windows (e.g., 50-day, 200-day moving averages).

1) Return and Alpha Signals: The most fundamental signals derive from daily returns $r_{i,t} = (P_{i,t} - P_{i,t-1})/P_{i,t-1}$ and excess returns (alpha) $\alpha_{i,t} = r_{i,t} - r_{c,t}$, where $r_{c,t}$ denotes the core benchmark return. We also compute price ratios $P_{i,t}/P_{c,t}$ and their logarithmic transformations to capture relative valuation dynamics.

2) Relative Strength: To measure momentum relative to the benchmark, we compute multi-horizon relative strength as the ratio of ETF momentum to core momentum:

$$\mathrm{RS}_{i,t}^{(n)} = \frac{(P_{i,t} - P_{i,t-n})/P_{i,t-n}}{(P_{c,t} - P_{c,t-n})/P_{c,t-n}} \tag{1}$$

We evaluate lookback periods of 21, 63, 126, and 252 trading days, corresponding approximately to 1, 3, 6, and 12 months. Additionally, we compute beta-adjusted relative strength by normalizing RS by the ETF's rolling beta:

$$\mathrm{BetaAdjRS}_{i,t} = \frac{\mathrm{RS}_{i,t}^{(252)}}{|\beta_{i,t}| + \delta} \tag{2}$$

where $\delta \in \{0.3, 0.5, 1.0\}$ is a damping parameter to prevent extreme values for low-beta ETFs.

3) Risk Metrics: Understanding risk characteristics is essential for satellite selection. We compute rolling beta as the covariance of ETF returns with core returns divided by core variance, using 21, 63, and 126-day windows. Idiosyncratic return adjusts alpha by residual volatility. Rolling volatility (annualized standard deviation of returns) captures absolute risk at 21 and 63-day horizons, while relative volatility normalizes this against the core's volatility. Drawdown, measured as the percentage decline from the running maximum price, quantifies downside exposure. We also compute relative drawdown as the difference between ETF and core drawdowns.

4) Risk-Adjusted Performance: Raw returns can be misleading without risk context. We compute rolling Sharpe ratios (return per unit volatility), Sortino ratios (return per unit downside deviation), and Information ratios (alpha per unit tracking error) over 126-day windows. These metrics identify ETFs delivering efficient risk-adjusted returns rather than merely high absolute returns.

5) Trend Indicators: Price trends relative to moving averages provide insight into momentum regimes. We measure the percentage deviation of price from its 50-day and 200-day moving averages, as well as the "golden cross" ratio of the 50-day to 200-day moving average. These signals help distinguish between trending and mean-reverting market conditions.

6) Mean Reversion Signals: A key innovation in our framework is the comprehensive treatment of mean reversion signals, which identify potentially oversold conditions that may precede recovery. We compute price z-scores at 63, 126, and 252-day windows (deviation from rolling mean normalized by rolling standard deviation), inverted such that oversold conditions yield positive values. The Bollinger Band position measures where price sits within its volatility bands. Classical oscillators including the Relative Strength Index and Stochastic oscillator are inverted to generate buy signals when assets appear oversold.

We also compute alpha z-scores, RS z-scores at 126 and 252-day windows, distance from 20 and 100-day moving averages, and the price ratio z-score. Perhaps most importantly, we introduce the alpha drawdown reversion signal:

$$\text{AlphaDD}_{i,t} = -\left(\sum_{s=t-126}^{t} \alpha_{i,s} - \max_{u \le t}\left(\sum_{s=u-126}^{u} \alpha_{i,s}\right)\right) \tag{3}$$

This captures how far an ETF's cumulative alpha has fallen from its peak, identifying assets that have significantly underperformed recently and may be poised for mean reversion.

7) Higher Moments: To capture tail risk characteristics, we compute 63-day rolling skewness and kurtosis of returns. Negative skewness indicates asymmetric downside risk, while high kurtosis signals fat tails and potential for extreme moves.

8) Correlation and Dispersion: We measure each ETF's 63-day rolling correlation with the core benchmark, as well as an approximation of average pairwise correlation with the broader universe (computed efficiently as correlation with the equal-weighted market average). Low correlation signals diversification potential, while high correlation indicates the ETF moves in lockstep with the market.

9) Momentum Dynamics: Beyond static momentum, we capture acceleration and deceleration in relative strength. The momentum acceleration signals measure the 21-day and 63-day change in 6-month or 12-month relative strength, identifying ETFs whose outperformance is increasing or decreasing.

10) Signal Disagreement Features: We compute disagreement signals between short-term and long-term momentum. The bullish disagreement signal $(\text{RS}_{12m} - \text{RS}_{1m})$ identifies long-term winners experiencing short-term pullbacks—potential buy opportunities. Conversely, bearish disagreement $(\text{RS}_{1m} - \text{RS}_{12m})$ flags short-term strength in long-term laggards. We also measure trend-momentum divergence by comparing price-vs-MA z-scores with RS z-scores.

11) Regime-Adaptive Signals: Market conditions affect which signals work best. We compute a volatility regime indicator (core volatility z-score relative to its 252-day history) and a trend regime indicator (core price vs 200-day MA). Two composite signals boost base signals based on regime:

$$\text{TrendBoostedMom}_{i,t} = \text{RS}_{i,t}^{(126)} \times (1 + \max(0, \text{TrendRegime}_t)) \tag{4}$$

$$\text{VolBoostedReversion}_{i,t} = \text{AlphaDD}_{i,t} \times (1 + \min(2, \max(0, \text{VolRegime}_t \tag{5}$$

These signals amplify momentum in trending markets and mean reversion in high-volatility regimes.

D. Causal Smoothing Filters

Raw signals contain substantial noise that can lead to spurious rankings. To address this while preserving causality (avoiding look-ahead bias), we apply 27 smoothing filters to each signal base. All filters are strictly causal—they use only past data points at each time step. We employ both standard windows (21 and 63 days) and specialized adaptive filters to capture different temporal dynamics.

Raw (no filter): The unfiltered signal serves as a baseline (1 filter).

Exponential Moving Average (EMA): We apply EMAs with 21 and 63-day spans, implemented using IIR filtering for computational efficiency (2 filters):

$$\text{EMA}_t = \alpha \cdot x_t + (1-\alpha) \cdot \text{EMA}_{t-1}, \quad \alpha = \frac{2}{n+1} \tag{6}$$

Double EMA (DEMA): Reduced-lag smoothing computed as $\text{DEMA} = 2 \cdot \text{EMA} - \text{EMA}(\text{EMA})$ with 21 and 63-day spans (2 filters).

Triple EMA (TEMA): Minimal-lag smoothing using $\text{TEMA} = 3 \cdot \text{EMA} - 3 \cdot \text{EMA}^2 + \text{EMA}^3$ with 21 and 63-day spans (2 filters).

Zero-Lag EMA (ZLEMA): Momentum-adjusted EMA that compensates for lag by using price momentum, with 21 and 63-day spans (2 filters).

Hull Moving Average: The Hull MA combines weighted moving averages to reduce lag while maintaining smoothness, using 21 and 63-day periods (2 filters):

$$\text{Hull}_t = \text{WMA}\left(2 \cdot \text{WMA}(x, n/2) - \text{WMA}(x, n), \sqrt{n}\right) \tag{7}$$

Triangular MA (TRIMA): Double-smoothed SMA providing center-weighted averaging, with 21 and 63-day windows (2 filters).

Gaussian MA: Bell-curve weighted averaging with smooth frequency rolloff, using 21 and 63-day windows (2 filters).

Butterworth Low-Pass Filter: Second-order Butterworth filters with cutoff periods of 21 and 63 days provide excellent noise reduction with a smooth frequency response (2 filters). We use SciPy's *lfilter* function for causal (one-directional) filtering.

Kalman Filter: An adaptive smoothing filter that adjusts based on signal-to-noise ratio (2 filters). We use two configurations: "fast" (process variance 10^{-4}) for responsive tracking and "slow" (process variance 10^{-5}) for stronger smoothing.

Savitzky-Golay Filter: Polynomial smoothing that preserves local maxima and minima better than moving averages, implemented with 21 and 63-day windows using cubic polynomials (2 filters). We ensure causality by setting the filter origin to use only past points.

Median Filter: Outlier-robust filtering that preserves edges, with 21 and 63-day windows (2 filters).

Kaufman Adaptive MA (KAMA): Adapts smoothing based on market efficiency ratio, distinguishing trending from choppy markets, using a 21-day period (1 filter).

Regime-Switching Filter: Automatically switches between fast (10-day) and slow (50-day) smoothing based on market volatility regime (1 filter).

Combining 167 signal bases with 27 filters yields 4,509 signal variants. Each filter-signal combination captures different temporal dynamics, allowing the selection algorithm to identify which combinations provide the strongest predictive power.

E. Indicator Transformations

Each filtered signal variant undergoes 25 indicator transformations that capture different aspects of the signal's behavior. These indicators form the final predictive features used for ETF selection.

Level: The raw filtered signal value.

Momentum indicators: We compute the change in signal value over 5, 21, 63, and 126-day horizons, capturing short to medium-term signal dynamics.

Z-score indicators: Signal deviation from its rolling mean, normalized by rolling standard deviation, computed at 21, 63, and 126-day windows. These identify when signals are at extreme values relative to recent history.

Cross-sectional indicators: At each time point, we compute the cross-sectional rank (ETF's percentile rank among all ETFs, yielding a uniform distribution in $[0, 1]$) and cross-sectional z-score (deviation from cross-sectional mean divided by cross-sectional standard deviation). These ensure selection is based on relative positioning rather than absolute magnitudes.

Trend indicators: Ratios of short to long-term moving averages of the signal—MA(21)/MA(63), MA(63)/MA(252), and a longer-term variant—identify whether the signal is trending up or down.

Mean reversion indicators: Distance from 63, 126, and 252-day rolling means, identifying when signals have deviated significantly and may revert.

Velocity and acceleration: First and second derivatives of the signal (normalized by rolling standard deviation), capturing the rate and change in rate of signal movement.

Drawdown: Distance from the 63-day rolling maximum, measuring how far the signal has fallen from recent peaks.

Range position: Where the signal sits within its 21-day high-low range, normalized to $[0, 1]$.

Volatility ratios: Ratio of short-term (5-day) to longer-term (21, 63, 126-day) signal volatility, identifying periods of unusual signal movement.

Percentile: The signal's position within its 252-day rolling distribution.

Relative volatility: Signal volatility normalized by core benchmark signal volatility.

Combining 4,509 signal variants with 25 indicators yields 112,725 total features. This comprehensive feature space allows systematic identification of which signal-filter-indicator combinations have genuine predictive power.

F. Cross-Sectional Normalization

Beyond the cross-sectional indicators described above, we apply normalization to ensure comparability across features. Signal values vary dramatically in scale and distribution across different signal types and market regimes. The cross-sectional rank divides each ETF's rank by the count of available ETFs, yielding a uniform distribution between 0 and 1. The cross-sectional z-score subtracts the cross-sectional mean and divides by cross-sectional standard deviation. These transformations ensure that selection decisions are based on relative positioning within the current universe rather than absolute signal magnitudes.

G. Feature Discovery and Selection

Given the large feature space (112,725 features), systematic evaluation is essential. We employ GPU-accelerated computation using CuPy to evaluate all features efficiently. For each feature, we compute predictive metrics on a rolling monthly basis:

Top-k advantage: The average monthly alpha of the top-4 ranked ETFs minus the average alpha across all ETFs. This measures the feature's ability to identify outperformers.

Hit rate: The percentage of months where the top-4 selection outperforms the universe average. This captures consistency of the signal.

Quintile spread: The return difference between the top 20% and bottom 20% of ranked ETFs, indicating overall discriminative power.

Features are ranked by their top-k advantage, with hit rate used as a secondary criterion. The best-performing features are then analyzed to understand which signal-filter-indicator combinations provide genuine predictive power, as opposed to those that may be artifacts of noise or overfitting.

H. Walk-Forward Backtesting

To obtain realistic performance estimates, we employ strict walk-forward backtesting. On each monthly rebalancing date, signals are computed using only data available up to that point. The selection algorithm ranks ETFs and chooses satellites based solely on historical information. Performance is then measured over the subsequent month on data unseen during signal computation.

This methodology prevents look-ahead bias and data snooping, ensuring that reported performance reflects what an investor could have actually achieved in real-time. We use a 252-day (one year) lookback for signal computation, monthly rebalancing on the last trading day of each month, and select $k = 4$ satellite ETFs per rebalancing.

All filters are strictly causal—the Savitzky-Golay filter, for example, uses the *origin* parameter to ensure only past points contribute to each output. This attention to causality is critical; non-causal filters can show artificially inflated performance by implicitly using future information.

I. Performance Evaluation

We evaluate feature quality through multiple complementary metrics:

- **Top-k advantage:** Monthly alpha of selected ETFs vs universe average
- **Hit rate:** Percentage of months with positive selection alpha
- **Quintile spread:** Return spread between top and bottom quintiles

We also analyze performance by filter type and indicator type to identify systematic patterns. For example, we compare the average top-k advantage of Hull-filtered signals versus raw signals, or level indicators versus momentum indicators.

Critically, our primary optimization objective is maximizing the percentage of months with positive rolling alpha rather than maximizing cumulative alpha. This emphasis on consistency over magnitude reflects the practical reality that investors find steady outperformance more valuable than volatile returns that average to the same level.

IV. Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

V. Discussion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

VI. Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Appendix

Appendix: Signal Framework Details

This appendix provides complete documentation of the signal framework used for ETF selection. The framework consists of three stages: signal base computation, smoothing filter application, and indicator transformation.

Signal Bases

Tables I and II list all 95 signal bases computed from raw price data. Each signal captures a different aspect of ETF behavior relative to the core benchmark (ACWI). Signals are organized into categories: returns and alpha, relative strength, risk metrics, risk-adjusted performance, trend indicators, mean reversion signals, higher moments, correlation, momentum dynamics, signal disagreement, regime-adaptive signals, and seasonality.

Smoothing Filters

Table III describes the 14 causal smoothing filter types (24 configurations with different parameters) applied to each signal base. All filters are strictly causal—they only use past data at each time step, ensuring no look-ahead bias in the backtest. The filters range from simple exponential moving averages to adaptive filters (KAMA, regime-switching) and sophisticated signal processing methods (Kalman, Butterworth, Savitzky-Golay).

Indicator Transformations

Table IV documents the 25 indicator transformations applied to filtered signals. These indicators capture different temporal and cross-sectional aspects of signal behavior, including momentum at multiple horizons, z-scores, trend ratios, mean reversion, velocity, acceleration, and volatility ratios.

Feature Space Summary

Combining all components yields the complete feature space:

- **Signal Bases**: 95 fundamental signals
- **Smoothing Filters**: 24 causal filter configurations (including raw)
- **Indicators**: 25 transformations
- **Total Features**: $95 \times 24 \times 25 = 57{,}000$

TABLE I

SIGNAL BASES (PART 1): RETURNS, MOMENTUM, RISK & QUALITY

Signal	Windows (days)	Description
	Returns & Alpha	
Daily Return	–	$(P_t - P_{t-1})/P_{t-1}$
Alpha vs Core	–	$r_{ETF} - r_{Core}$
Alpha vs Universe	–	$r_{ETF} - \bar{r}_{univ}$
Price Ratio vs Core	–	P_{ETF}/P_{Core}
Price Ratio vs Universe	–	$r_{ETF}/\bar{r}_{univ}$
Log Ratio vs Core	–	$\ln(P_{ETF}/P_{Core})$
Log Ratio vs Universe	–	$\ln(r_{ETF}/\bar{r}_{univ})$
Cumulative Return	21, 63, 126, 252	$\sum(P_t/P_{t-1} - 1)$
Cumulative Alpha vs Core	21, 63, 126, 252	$\sum(r_{ETF} - r_{Core})$
Cumulative Alpha vs Universe	21, 63, 126, 252	$\sum(r_{ETF} - \bar{r}_{univ})$
	Momentum	
Momentum	21, 63, 126, 252	$(P_t - P_{t-n})/P_{t-n}$
Relative Strength vs Core	21, 63, 126, 252	$\text{Mom}_{ETF}/\text{Mom}_{Core}$
Relative Strength vs Universe	21, 63, 126, 252	$\text{Mom}_{ETF}/\text{Mom}_{univ}$
Skip-Month Momentum	t: 63, 126, 252; s: 21, 42, 63	t-day return, skip s days
52-Week High Proximity	–	$P_t/P_{52w,high}$
52-Week Low Proximity	–	$P_t/P_{52w,low}$
Rate of Change (ROC)	10, 20	$(P_t - P_{t-n})/P_{t-n} \times 100$
	Risk Metrics	
Beta	21, 63, 126, 252	Cov/Var with core
Volatility	21, 63, 126, 252	$\sigma \times \sqrt{252}$
Downside Deviation	21, 63, 126, 252	$\sqrt{\text{E}[\min(r,0)^2]} \times \sqrt{252}$
Idiosyncratic Return	21, 63, 126, 252	$r_{ETF} - \beta_w \cdot r_{Core}$ (CAPM residual)
Relative Volatility	21, 63, 126, 252	$\sigma_{ETF}/\sigma_{Core}$
Drawdown	–	$(P - P_{max})/P_{max}$
Relative Drawdown	–	DD_{ETF} - DD_{Core}
Drawdown Duration	–	Days since last peak
Recovery Rate	21, 63, 126, 252	ΔDD / time underwater
CVaR (95%)	21, 63, 126, 252	$\text{E}[r \mid r < \text{VaR}_{0.05}]$
Ulcer Index	21, 63, 126, 252	$\sqrt{\text{mean}(\text{DD}^2)}$
	Risk-Adjusted Performance	
Sharpe Ratio	21, 63, 126, 252	$\bar{r}/\sigma \times \sqrt{252}$
Information Ratio	21, 63, 126, 252	$\bar{\alpha}/\sigma_\alpha \times \sqrt{252}$
Sortino Ratio	21, 63, 126, 252	$\bar{r}/\sigma_{down} \times \sqrt{252}$
Calmar Ratio	21, 63, 126, 252	$\bar{r}/\lvert\text{MaxDD}\rvert$
Treynor Ratio	21, 63, 126, 252	$(\bar{r} - r_f)/\beta$
Omega Ratio	21, 63, 126, 252	$\sum(r > 0)/\lvert\sum(r < 0)\rvert$
Gain-to-Pain Ratio	21, 63, 126, 252	$\sum r^+/\lvert\sum r^-\rvert$
Ulcer Performance Index	21, 63, 126, 252	$(\bar{r} - r_f)$/Ulcer
Recovery Factor	21, 63, 126, 252	Total return / $\lvert$MaxDD$\rvert$
	Win/Loss Analysis	
Win Rate	21, 63, 126, 252	$\#(r > 0)/\#(r \neq 0)$
Payoff Ratio	21, 63, 126, 252	$\bar{r}^+/\lvert\bar{r}^-\rvert$
Profit Factor	21, 63, 126, 252	$\sum r^+/\lvert\sum r^-\rvert$
Tail Ratio	21, 63, 126, 252	$\lvert P_{95}\rvert/\lvert P_5\rvert$ of returns
Stability of Returns	21, 63, 126, 252	R^2 of cumulative log returns
	Beta-Adjusted Relative Strength	
Beta-Adjusted RS	d: 0.3, 0.5, 1.0; β_w: 21, 63, 126, 252	$\text{RS}_{252}/(\lvert\beta\rvert + d)$
	Trend Indicators	
Price vs Moving Average	20, 50, 100, 200	$P/\text{MA}_n - 1$
Moving Average Crossover	20/50, 50/200	MA_s/MA_l - 1
MACD	12/26/9	EMA_{12} - EMA_{26}, signal = EMA_9
PPO (Percentage Price Osc)	12/26/9	(EMA_{12} - EMA_{26})/$\text{EMA}_{26} \times 100$
DPO (Detrended Price Osc)	20	$P - \text{MA}_{n/2+1\text{ ago}}$
TRIX	15	ROC of triple EMA
KST (Know Sure Thing)	10/15/20/30	Weighted sum of 4 ROC periods
Aroon Oscillator	25	Aroon_{up} - Aroon_{down}

TABLE II
SIGNAL BASES (PART 2): MEAN REVERSION, DYNAMICS & REGIME

Signal	Windows (days)	Description
	Mean Reversion – Z-Scores (inverted for buy signals)	
Price Z-Score	21, 63, 126, 252	$-(P-\bar{P})/\sigma$
Alpha Z-Score	21/252, 63/252, 126/252, 252/252	$-(\sum \alpha_n)/(\sigma_{252} \times \sqrt{n})$
RS Z-Score	21, 63, 126, 252	Inv. RS z-score
Distance from MA	20, 50, 100, 200	$-(P-\text{MA})/\text{MA}$
Price Ratio Z-Score	252	Inv. price ratio z
	Mean Reversion – Technical Oscillators	
Bollinger Reversion	20, 2σ	$1-(P-\text{lower})/(\text{upper}-\text{lower})$
RSI Reversion	14	$50-\text{RSI}$
Stochastic Reversion	14	$50-K$ (close-only)
Williams %R	14	$(H_n-C)/(H_n-L_n) \times -100$
TSI (True Strength Index)	25/13	Double-smoothed price change ratio
CCI (Commodity Channel Index)	20	$(P-\text{SMA})/(0.015 \times \text{MAD})$
Ultimate Oscillator	7/14/28	Weighted 3-timeframe buying pressure
Donchian Position	20	$(P-L_n)/(H_n-L_n)$
	Mean Reversion – Drawdown-Based	
Drawdown Reversion	–	$-$drawdown
Alpha Drawdown Reversion	21, 63, 126, 252	$-(\sum \alpha_n - \max(\sum \alpha_n))$
Sector Rotation Reversion	21, 63, 126, 252	$-\text{RS}_n$ (contrarian)
	Higher Moments & Complexity	
Skewness	21, 63, 126	Rolling skewness
Kurtosis	21, 63, 126	Rolling excess kurtosis
Return Autocorrelation	21, 63, 126, 252	$\text{Corr}(r_t, r_{t-1})$ over window
Hurst Exponent	63, 126, 252	R/S analysis trend persistence
Return Entropy	63, 126	$-\sum p_i \log p_i$ of return bins
	Correlation & Capture Ratios	
Core Correlation	21, 63, 126, 252	$\text{Corr}(r_{ETF}, r_{Core})$
Diversification Benefit	63	$1-$ core correlation
Market Correlation	63	Avg pairwise corr
Return Dispersion	63	Cross-sectional σ of returns
Crowding/Herding	63	% of ETFs with same momentum sign
Up Capture Ratio	21, 63, 126, 252	$\bar{r}^{up}_{ETF}/\bar{r}^{up}_{Core}$ when core >0
Down Capture Ratio	21, 63, 126, 252	$\bar{r}^{down}_{ETF}/\bar{r}^{down}_{Core}$ when core <0
Capture Ratio Spread	21, 63, 126, 252	Up Capture $-$ Down Capture
	Momentum Dynamics	
Momentum Acceleration	n: 126, 252; h: 21, 63	h-day Δ in RS_n
	Signal Disagreement	
Bullish Disagreement	21, 252	RS_{252} - RS_{21}
Bearish Disagreement	21, 252	RS_{21} - RS_{252}
Timeframe Disagreement	21/63, 63/126, 126/252, 63/252	RS_l - RS_s (adjacent pairs)
Trend-Momentum Divergence	50, 63	PriceMA_z - RS_z
	Regime-Adaptive Signals	
Volatility Regime	63/252	Core vol z-score
Trend Regime	200	Core vs MA_{200}
Drawdown Regime	–	Core drawdown state
Dispersion Regime	63	Cross-sectional return spread
Trend-Boosted Momentum	126, 200	$\text{RS}_{126}\times$ trend
Volatility-Boosted Reversion	63, 252	AlphaDD $\times$ vol
	Seasonality	
Month-of-Year Effect	–	Binary: 1 if favorable month (Nov–Apr)
Month Sine/Cosine	–	$\sin/\cos(2\pi \cdot \text{month}/12)$

Note: All window parameters use days: 21d $\approx$ 1 month, 63d $\approx$ 3 months, 126d $\approx$ 6 months, 252d $\approx$ 1 year. Total: **167 signal bases**. All signals validated against industry-standard libraries (ta, empyrical, quantstats).

TABLE III
CAUSAL SMOOTHING FILTERS APPLIED TO SIGNAL BASES

Filter	Parameters	Description
Raw	–	Unfiltered signal S_t passed through directly
Exponential MA	span: 21, 63	$S_t^{ema} = \alpha S_t + (1-\alpha) S_{t-1}^{ema}$, $\alpha = 2/(n+1)$
Double EMA	span: 21, 63	$2 \cdot \text{EMA} - \text{EMA(EMA)}$, reduced lag
Triple EMA	span: 21, 63	$3 \cdot \text{EMA} - 3 \cdot \text{EMA}^2 + \text{EMA}^3$, minimal lag
Zero-Lag EMA	span: 21, 63	EMA of $(2S_t - S_{t-\text{lag}})$, momentum-adjusted
Hull MA	period: 21, 63	$\text{WMA}(2 \cdot \text{WMA}_{n/2} - \text{WMA}_n, \sqrt{n})$, low lag
Triangular MA	period: 21, 63	SMA(SMA(S_t)), double-smoothed, center-weighted
Gaussian MA	window: 21, 63	Bell-curve weights $w_i \propto e^{-i^2/2\sigma^2}$, smooth rolloff
Kaufman Adaptive	period: 21; fast/slow: 2/30	Adapts α via efficiency ratio, trend-sensitive
Median Filter	window: 21, 63	Rolling median, outlier-robust, preserves edges
Regime-Switching	fast/slow: 10/50; thresh	Uses fast MA in high-vol, slow MA in low-vol
Butterworth	cutoff: 21, 63; order=2	Low-pass filter, attenuates $f > 1/n$ day^{-1}
Kalman Filter	Q: 10^{-4}, 10^{-5}; R=10^{-2}	Steady-state Kalman, adapts to signal noise
Savitzky-Golay	window: 21, 63; poly=3	Cubic polynomial fit, preserves local peaks

Note: All filters are strictly causal—they only use past data at each time step. 21 days $\approx$ 1 month, 63 days $\approx$ 3 months.

Each feature is named using the convention: `signal__filter__indicator`, for example `vol_boosted_reversion__hull_63d__xs_rank`.

The GPU-accelerated computation pipeline processes all 57,000 features efficiently using:

- CuPy for parallel computation on NVIDIA GPUs
- Streaming mode to limit memory usage
- Disk caching for rolling statistics to avoid redundant computation
- Vectorized operations for signal bases and indicators

TABLE IV
INDICATOR TRANSFORMATIONS APPLIED TO FILTERED SIGNALS

Indicator	**Parameters**	**Description**
	Level & Momentum	
Level	–	Raw filtered signal S_t passed through directly
Momentum	horizon: 5, 21, 63, 126	$(S_t - S_{t-h})/\lvert S_{t-h}\rvert$, rate of change
Momentum Acceleration	horizon: 5, 21, 63	$\Delta(\text{mom}_h)$, is momentum speeding up?
Velocity	norm: 5, 21, 63	$\Delta S_t/\sigma_w$, normalized first derivative
Acceleration	norm: 5, 21, 63	$\Delta^2 S_t/\sigma_w$, normalized second derivative
Curvature	norm: 5, 21, 63	$\Delta^2 S_t/\lvert\Delta S_t\rvert$, detects inflection points
	Statistical Normalization	
Z-Score	window: 21, 63, 126, 252	$(S_t - \bar{S}_w)/\sigma_w$, time-series standardization
Cross-Sectional Z-Score	–	$(S_i - \bar{S}_{xs})/\sigma_{xs}$, relative to peers
Cross-Sectional Rank	–	Sigmoid of cross-sectional z-score, bounded $[0, 1]$
Percentile	window: 63, 126, 252	$(S_t - \bar{S}_w)/(2\sigma_w)$, historical position
	Trend Indicators	
Trend Short	MA: 5/21	$(\bar{S}_5 - \bar{S}_{21})/\lvert\bar{S}_{21}\rvert$, short-term trend
Trend Medium	MA: 21/63	$(\bar{S}_{21} - \bar{S}_{63})/\lvert\bar{S}_{63}\rvert$, medium-term trend
Trend Long	MA: 63/126	$(\bar{S}_{63} - \bar{S}_{126})/\lvert\bar{S}_{126}\rvert$, long-term trend
Trend Extended	MA: 126/252	$(\bar{S}_{126} - \bar{S}_{252})/\lvert\bar{S}_{252}\rvert$, extended trend
Divergence	window: 21, 63	$S_t - \bar{S}_w$, signal vs smoothed version
	Mean Reversion	
Reversion	window: 21, 63, 126, 252	$(S_t - \bar{S}_w)/\lvert\bar{S}_w\rvert$, deviation from mean
Envelope	window: 21, 63	$(S_t - \bar{S}_w)/(k \cdot \sigma_w)$, Bollinger-style bands
	Breakout & Range	
Distance to High	window: 21, 63, 126	$(S_t - S_{max})/\sigma_w$, breakout detection
Distance to Low	window: 21, 63, 126	$(S_t - S_{min})/\sigma_w$, breakdown detection
Drawdown	window: 21, 63, 126	$(S_t - S_{max})/\lvert S_{max}\rvert$, decline from peak
Range Position	window: 21, 63, 126	$(S_t - S_{min})/(S_{max} - S_{min})$, position in range
Ratio to Peak	window: 21, 63, 126	S_t/S_{max}, percentage of peak value
	Volatility Metrics	
Volatility Ratio	short/long: 5/21, 21/63, 63/126	$\sigma_{short}/\sigma_{long}$, volatility regime indicator
Relative Volatility	short/long: 21/252, 63/252	$\sigma_{short}/\sigma_{long}$, short vs long-term vol
Signal-to-Noise	window: 21, 63, 126	$\lvert\bar{S}_w\rvert/\sigma_w$, signal clarity measure
Roughness	window: 21, 63, 126	$\sum\lvert\Delta S\rvert/\lvert S_T - S_0\rvert$, path noise vs net move
	Higher Moments	
Skewness	window: 21, 63, 126	Rolling asymmetry of signal distribution
Kurtosis	window: 21, 63, 126	Rolling peakedness, extreme event likelihood
	Regime Indicators	
Above Mean	window: 21, 63, 126	Binary: $S_t > \bar{S}_w$, simple regime indicator
	Cross-Sectional Dynamics	
Dispersion	–	Cross-sectional σ of signal, market disagreement
Convergence	window: 21, 63	$\Delta\sigma_{xs}$, is signal converging across ETFs?

Note: 95 signals $\times$ 25 filters $\times$ 76 indicators = **180,500 features**. All indicators are strictly causal.