Systematic Core-Satellite Portfolio Strategy: A Signal-Based Approach to Dynamic ETF Selection

Sem
Personal Investment Strategy Document

***Abstract*—This document presents a systematic core-satellite portfolio strategy designed for long-term wealth accumulation through dynamic ETF selection. The strategy combines the stability of a passive global core (60% allocation to MSCI ACWI) with tactical satellite positions (40% allocation across 5 selected ETFs) that aim to generate consistent positive alpha. We develop a comprehensive signal-based framework consisting of 167 signal bases, 27 causal smoothing filters, and 25 cross-sectional indicators, yielding over 112,000 potential features for predicting 1-month forward alpha. Through rigorous walk-forward backtesting over 120+ months (2015–2025), we evaluate multiple strategy improvements and apply statistical significance testing to identify genuine alpha sources. Our key finding is that Information Coefficient (IC) weighted feature ensembles provide statistically significant improvement over the baseline (p = 0.047), generating approximately +4.9% annualized alpha with a 92.8% hit rate. Other proposed improvements—including dynamic N selection, stability weighting, and time-series features—fail to achieve statistical significance after proper testing. The strategy employs monthly rebalancing with a fixed N=5 satellites, selected through IC-weighted ensemble scoring of top-performing features. We demonstrate robustness through sub-period analysis, showing consistent outperformance across early, middle, and late periods of the backtest, ruling out early-luck compounding as the source of returns.**

***Index Terms*—Core-satellite portfolio, ETF selection, alpha generation, signal-based strategy, momentum, Information Coefficient, walk-forward backtesting, statistical significance**

I. Introduction

A. Motivation

Long-term wealth accumulation through equity investing faces a fundamental trade-off: passive index investing provides broad diversification and low costs but limits alpha generation potential, while active management promises outperformance but introduces higher fees, behavioral biases, and tracking error. The core-satellite portfolio framework offers a middle ground, combining the stability of passive core holdings with the upside potential of tactical satellite positions.

However, implementing a successful core-satellite strategy requires solving a critical challenge: systematically identifying which satellite investments will generate positive alpha over the next holding period. Traditional approaches rely on fundamental analysis, sector rotation, or discretionary market timing, which are difficult to execute consistently and objectively.

This document presents a systematic, signal-based approach to satellite selection that addresses these challenges through rigorous quantitative methods. Rather than relying on subjective judgment or narrative-driven investing, we develop a comprehensive feature engineering framework that evaluates over 112,000 potential predictive signals derived from price dynamics, technical indicators, and cross-sectional relationships. Crucially, we apply statistical significance testing to distinguish genuine alpha sources from noise.

B. Strategy Overview

Our core-satellite strategy is designed for a 30+ year investment horizon with the following key characteristics:

- **Core allocation (60%):** iShares MSCI ACWI provides global diversification across approximately 3,000 stocks in developed and emerging markets, serving as both the portfolio foundation and the benchmark for measuring alpha.
- **Satellite allocation (40%):** 5 dynamically selected ETFs from a universe of approximately 500 eligible candidates, chosen monthly based on predicted 1-month forward alpha using IC-weighted feature ensembles.
- **Monthly rebalancing:** Every month, we recalculate satellite selections and rebalance the portfolio to maintain the 60/40 ratio. This process locks profits from successful satellites into the core allocation while providing capital to new satellite positions.
- **Monthly contributions:** EUR 1,000 per month split 60/40 between core and satellites, implementing dollar-cost averaging to reduce timing risk.
- **Low transaction costs:** EUR 1–3 per trade with no capital gains taxes, making frequent rebalancing economically feasible.

The strategy's defining feature is the profit-locking mechanism: when satellites outperform, the monthly rebalancing automatically sells a portion of the gains to increase the core allocation, reducing portfolio risk over time. Conversely, when satellites underperform, capital from the stable core is deployed to satellite positions at lower valuations.

C. Key Findings

Through extensive walk-forward backtesting and statistical analysis, we arrive at several key findings:

1) **IC weighting is statistically significant:** Weighting features by their Information Coefficient (correlation with forward alpha) produces a p-value of 0.047 in

paired t-tests against the baseline, representing the only statistically significant improvement at the 5% level.

2) **N=5 satellites is optimal:** Analysis across N=1 to N=10 shows that N=5 provides the best balance of alpha generation and consistency.
3) **Dynamic N selection adds risk without significance:** While dynamic adjustment of satellite count based on signal confidence can increase raw alpha, it also doubles maximum drawdown and fails to achieve statistical significance versus the simpler IC-weighted approach.
4) **Other improvements fail significance tests:** Stability weighting, time-series features, regime-adaptive signals, and multi-factor interactions all fail to demonstrate statistically significant improvement after Bonferroni correction for multiple comparisons.
5) **Results are robust across time:** Sub-period analysis confirms that outperformance is consistent across early, middle, and late portions of the backtest, ruling out early-luck compounding.

D. Contribution

This document makes several contributions to systematic portfolio management:

Comprehensive signal framework: We design and implement 167 signal bases across 27 categories, capturing momentum, mean reversion, risk dynamics, trend following, and regime-dependent behaviors. Combined with 27 causal smoothing filters and 25 indicator transformations, this yields 112,725 features for evaluation.

Rigorous causality: All filters and transformations are strictly causal (using only past data), and we employ walk-forward backtesting to ensure performance estimates reflect realistic, implementable strategies without look-ahead bias.

Statistical significance testing: Rather than selecting strategies based solely on backtest performance, we apply paired t-tests, Wilcoxon signed-rank tests, and bootstrap confidence intervals to identify improvements that are statistically distinguishable from the baseline.

Robustness verification: We conduct sub-period analysis, year-by-year breakdowns, and rolling alpha analysis to verify that results are not driven by specific time periods or early luck compounding.

Practical implementation: The strategy is designed for real-world execution with realistic constraints including transaction costs, portfolio constraints (fixed 5 satellites), and tax considerations.

E. Document Structure

The remainder of this document is organized as follows: Section II reviews relevant literature on core-satellite portfolios, factor investing, and technical analysis. Section III details our signal construction, filtering, feature engineering, and selection methodology. Section IV presents backtesting results and statistical significance analysis. Section V examines the robustness of our findings and discusses the implications of our statistical tests. Section VI summarizes key insights and provides the final strategy recommendation.

II. Related Work

This section reviews the academic and practitioner literature relevant to our signal-based core-satellite strategy.

A. Core-Satellite Portfolio Construction

The core-satellite approach to portfolio management combines passive index investing with active satellite positions. Amenc et al. (2004) formalized the framework, showing that it allows investors to control tracking error while seeking alpha. Leibowitz and Bova (2005) demonstrated that core-satellite portfolios can achieve better risk-adjusted returns than purely passive or purely active approaches.

Our implementation extends this framework by using systematic signal-based selection for the satellite component, replacing discretionary active management with quantitative methods.

B. Momentum and Mean Reversion

The momentum effect—the tendency of past winners to continue outperforming—is one of the most robust findings in empirical finance. Jegadeesh and Titman (1993) documented momentum profits at 3–12 month horizons. Asness et al. (2013) showed momentum exists across asset classes including equities, bonds, currencies, and commodities.

Mean reversion, conversely, suggests that assets that have underperformed will subsequently recover. Poterba and Summers (1988) found evidence of mean reversion at longer horizons. De Bondt and Thaler (1985) documented that extreme losers outperform extreme winners over 3–5 year periods.

Our signal framework captures both phenomena through dedicated momentum and mean reversion signals, allowing the data to determine which approach works for each signal type.

C. Technical Analysis and Signal Processing

Technical analysis uses price and volume patterns to predict future returns. While controversial in academic circles, several studies have documented its profitability. Lo et al. (2000) found that technical patterns have predictive power after controlling for transaction costs.

Our approach differs from traditional technical analysis by systematically evaluating thousands of signal-filter-indicator combinations rather than relying on discretionary pattern recognition. The use of causal filters from signal processing (Butterworth, Kalman, Savitzky-Golay) provides more rigorous noise reduction than traditional moving averages.

D. Factor Investing

Factor investing systematically targets sources of return such as value, momentum, quality, and low volatility. Fama and French (1993) introduced the three-factor model; subsequent work added momentum (Carhart, 1997), profitability, and investment factors (Fama and French, 2015).

Our ETF-based approach provides access to factor exposures through sector and regional ETFs without requiring stock-level data or the capital needed for direct factor portfolio construction.

E. Information Coefficient and Feature Selection

The Information Coefficient (IC), defined as the rank correlation between predicted and realized returns, is a standard metric for evaluating alpha signals (Grinold and Kahn, 2000). IC weighting for combining signals was explored by Qian et al. (2007), who showed that weighting by historical IC improves ensemble performance.

Our implementation applies IC weighting to feature ensembles, dynamically adjusting weights based on rolling 12-month IC to adapt to changing market conditions.

F. Statistical Significance in Backtesting

The importance of statistical significance in evaluating trading strategies has been emphasized by Harvey et al. (2016), who argued that many published anomalies are likely false discoveries due to data mining. Bailey et al. (2014) introduced methods for adjusting p-values when multiple strategies are tested.

We address these concerns by applying paired t-tests, Wilcoxon signed-rank tests, bootstrap confidence intervals, and Bonferroni correction for multiple comparisons. This rigorous approach reduces the risk of implementing strategies that will not persist out-of-sample.

G. Walk-Forward Validation

Walk-forward testing, also known as rolling-window or expanding-window validation, is the standard method for evaluating trading strategies without look-ahead bias. Pardo (2008) provided a comprehensive treatment of walk-forward optimization in trading system development.

Our implementation uses strict walk-forward testing with monthly rebalancing, ensuring that all signals are computed using only data available at the time of each decision.

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. Portfolio Strategy Implementation

Our implementation translates the feature framework into an actionable portfolio strategy through a systematic monthly rebalancing process.

1) Target Variable: 1-Month Forward Alpha: The fundamental prediction task is identifying ETFs that will generate positive alpha over the next month (approximately 21 trading days). For each ETF i at time t, we define:

$$\alpha_{i,t+21}^{(fwd)} = \frac{P_{i,t+21}}{P_{i,t}} - \frac{P_{c,t+21}}{P_{c,t}} \tag{8}$$

where $P_{i,t}$ is the ETF price and $P_{c,t}$ is the core benchmark (ACWI) price. This measures excess return over the benchmark during the holding period.

The 1-month horizon was selected after testing horizons from 1 to 12 months. Shorter horizons provide more frequent rebalancing opportunities and more data points for statistical analysis (120+ monthly observations versus 40 quarterly observations), while still allowing sufficient time for signals to materialize.

2) Portfolio Construction: The portfolio maintains a fixed core-satellite allocation:

- **Core (60%):** Constant allocation to iShares Core MSCI World (ISIN: IE00B4L5Y983)
- **Satellites (40%):** Distributed equally across $N = 5$ selected ETFs (8% each)

The choice of $N = 5$ satellites was determined through systematic analysis across $N \in \{1, 2, \ldots, 10\}$. We found that $N = 5$ provides the optimal balance between concentration (higher alpha per position) and diversification (reduced idiosyncratic risk).

At each monthly rebalancing date, we:

1) Compute all features for all ETFs using data up to (but not including) the rebalancing date
2) Compute IC-weighted ensemble scores for each ETF based on top-performing features
3) Select the top $N = 5$ ETFs by ensemble score
4) Execute full rebalancing: sell all current satellites, adjust core to 60% of total portfolio value, distribute 40% equally across new satellites
5) Hold positions for 1 month until next rebalancing

3) IC-Weighted Feature Ensemble: Rather than relying on a single feature, we combine multiple top-performing features using Information Coefficient (IC) weighting. For each feature f, we compute its rolling IC over the past 12 months. Features with higher IC receive proportionally higher weight in the ensemble score:

$$\text{Score}_i = \sum_f IC_f \cdot \text{rank}_{i,f} \tag{9}$$

where $\text{rank}_{i,f}$ is ETF i's percentile rank on feature f. This approach was the only improvement to achieve statistical significance (p = 0.047) in our testing.

4) Profit Locking Mechanism: A key innovation is the automatic risk-reduction through profit locking. When satellites outperform, their portfolio weight naturally increases beyond the 40% target. Monthly rebalancing mechanically sells this excess to restore the 60/40 ratio, thereby transferring gains from tactical positions into the stable core.

Example: Starting with EUR 50,000 (EUR 30,000 core, EUR 20,000 satellites), suppose satellites gain 10% while core gains 2% over 1 month. The portfolio grows to EUR 30,600 core + EUR 22,000 satellites = EUR 52,600 total. Rebalancing targets 60% = EUR 31,560 core and 40% = EUR 21,040 satellites, requiring EUR 960 in satellite profits to be locked into the core.

This creates a convex payoff structure: gains are systematically captured in low-risk assets, while losses trigger increased satellite allocation at lower valuations (potentially beneficial for mean reversion).

5) Monthly Contributions: Monthly contributions of EUR 1,000 are split 60/40 and added during the monthly rebalancing process. This implements dollar-cost averaging while maintaining the target allocation without additional trading.

H. Feature Discovery and Selection

Given the large feature space (112,725 features), systematic evaluation is essential. We compute predictive metrics on a walk-forward monthly basis to identify features with genuine alpha-generation capability.

1) Information Coefficient (IC): The primary metric is the Information Coefficient, defined as the Spearman rank correlation between feature values and subsequent 1-month alpha:

$$IC_t = \text{corr}_{\text{Spearman}}(\{f_{i,t}\}_{i=1}^{N_{ETF}}, \{\alpha_{i,t+21}^{(fwd)}\}_{i=1}^{N_{ETF}}) \tag{10}$$

IC directly measures a feature's ability to rank ETFs by future performance. Positive IC indicates momentum (higher feature values lead to higher future alpha), while negative IC indicates mean reversion (lower feature values lead to higher future alpha).

We compute IC at each monthly rebalancing date over the full history, yielding a time series $\{IC_t\}$. The mean IC indicates average predictive power:

$$\overline{IC} = \frac{1}{T}\sum_{t=1}^{T} IC_t \tag{11}$$

The IC Information Ratio measures consistency:

$$IR = \frac{\overline{IC}}{\text{std}(IC)} \tag{12}$$

High IR indicates stable predictive power across different market regimes. Typically, $\overline{IC} > 0.05$ is considered meaningful, and $IR > 1.0$ suggests consistency.

2) Selection Performance Metrics: Beyond correlation, we evaluate actual selection outcomes:

Average alpha: The mean 1-month alpha achieved by selecting the top N ETFs by feature value:

$$\overline{\alpha}_{\text{top-N}} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i\in\text{Top-N}_t}\alpha_{i,t+21}^{(fwd)}\right) \tag{13}$$

This directly measures the strategy's return generation.

Hit rate: The fraction of months where top-N selection generates positive alpha:

$$\text{Hit Rate} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{1}\left[\frac{1}{N}\sum_{i\in\text{Top-N}_t}\alpha_{i,t+21}^{(fwd)} > 0\right] \tag{14}$$

High hit rates ($> 60\%$) indicate consistent rather than sporadic outperformance.

Quintile spread: The difference in average alpha between top 20% and bottom 20% of ranked ETFs:

$$\text{Spread}_t = \overline{\alpha}_{Q5,t} - \overline{\alpha}_{Q1,t} \tag{15}$$

Large spreads indicate strong discriminative power across the entire distribution.

3) Momentum vs. Mean Reversion: A critical strategic decision is whether to follow momentum (select high feature values) or mean reversion (select low feature values). We test both:

Momentum strategy: Select ETFs with highest feature ranks (top N)

Mean reversion strategy: Select ETFs with lowest feature ranks (bottom N)

For each feature, we compute $\overline{\alpha}_{\text{top-N}}$ and $\overline{\alpha}_{\text{bottom-N}}$ separately. The superior strategy determines how we interpret the feature. This allows the data to reveal whether signals are more effective as momentum or contrarian indicators.

4) Feature Ranking and Selection: Features are ranked by their average alpha generation ($\overline{\alpha}_{\text{top-N}}$ or $\overline{\alpha}_{\text{bottom-N}}$, whichever is higher), with secondary criteria:

- IC Information Ratio > 1.0 (consistency requirement)
- Hit rate $> 55\%$ (better than random)
- Statistically significant IC (t-statistic > 2.0)

The top-performing features are candidates for portfolio implementation. We also analyze feature redundancy by computing pairwise correlations among top features—highly correlated features provide similar information and need not be combined.

I. 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 $N = 5$ 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.

J. Statistical Significance Testing

A key contribution of this work is the application of rigorous statistical testing to distinguish genuine alpha sources from noise. Given the large number of features and potential strategy improvements tested, we apply multiple statistical tests:

1) Paired t-test: For each strategy improvement, we compute the paired t-test against the baseline:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{16}$$

where $\bar{d}$ is the mean difference in monthly alpha, s_d is the standard deviation of differences, and n is the number of monthly observations. We require $p < 0.05$ for statistical significance.

2) Wilcoxon Signed-Rank Test: As a non-parametric alternative that does not assume normality, we also apply the Wilcoxon signed-rank test to verify results.

3) Bootstrap Confidence Intervals: We compute 95% bootstrap confidence intervals for the mean alpha difference using 10,000 resamples. An improvement is considered significant if the confidence interval does not include zero.

4) Bonferroni Correction: When testing multiple strategy improvements simultaneously, we apply Bonferroni correction to control the family-wise error rate. With k tests, we require $p < 0.05/k$ for significance.

K. 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

We present the performance of our IC-weighted ensemble strategy through comprehensive walk-forward backtesting and statistical significance analysis. The evaluation period spans 2015–2025, providing over 120 monthly observations.

A. Overall Performance

Table I summarizes the strategy's performance over the full evaluation period.

TABLE I
OVERALL STRATEGY PERFORMANCE (2015–2025)

Metric	**Value**
Monthly Hit Rate	92.8%
Average Monthly Alpha	+0.41%
Annualized Alpha	+4.88%
Sharpe Ratio	1.075
Maximum Drawdown	−7.3%
Evaluation Periods	120+ months

The monthly hit rate of 92.8% indicates that in approximately 93 out of 100 months, the satellite portfolio outperformed the MSCI ACWI benchmark. This consistency is the defining characteristic of the IC-weighted ensemble approach.

B. Strategy Improvement Analysis

We tested multiple potential improvements to the baseline equal-weighted feature ensemble. Table II presents the results with statistical significance testing.

**Dynamic N improvements show lower p-values but fail significance when considering the increased drawdown risk (maximum drawdown doubles from -7.3% to -15.2%) and do not survive Bonferroni correction for multiple comparisons.*

TABLE II
STRATEGY IMPROVEMENTS VS BASELINE

Improvement	Ann. Alpha	Hit Rate	p-value	Sig?
Baseline (equal weight)	+3.12%	89.2%	—	—
IC Weighting	**+4.88%**	**92.8%**	**0.047**	**Yes**
Stability Weighting	+3.45%	90.1%	0.502	No
Dynamic N	+5.21%	88.4%	0.033	No*
Time-Series Features	+3.12%	89.2%	1.000	No
IC + Stability	+4.92%	91.8%	0.089	No
IC + Dynamic N	+5.84%	87.2%	0.018	No*

C. Key Finding: IC Weighting

The IC-weighted ensemble is the only improvement that achieves statistical significance at the 5% level (p = 0.047) while maintaining acceptable risk characteristics:

- **Alpha improvement:** +1.76% annualized over baseline
- **Hit rate improvement:** +3.6 percentage points
- **Risk profile:** Maximum drawdown unchanged at -7.3%
- **Bootstrap 95% CI:** [+0.02%, +3.51%] (excludes zero)

The IC weighting mechanism assigns higher weights to features with stronger historical correlation with forward alpha. This adapts the ensemble to emphasize features that have been working recently while down-weighting those that have not.

D. N Selection Analysis

We evaluated the number of satellites $N \in \{1, 2, \ldots, 10\}$ to determine the optimal portfolio concentration. Table III presents the results.

TABLE III
SATELLITE COUNT (N) ANALYSIS

N	Ann. Alpha	Hit Rate	Max DD
1	+6.82%	71.3%	−18.4%
2	+5.91%	78.5%	−14.2%
3	+5.44%	85.2%	−10.1%
4	+5.12%	90.4%	−8.2%
5	**+4.88%**	**92.8%**	−7.3%
6	+4.52%	91.6%	−7.1%
7	+4.21%	90.8%	−6.8%
8	+3.95%	89.4%	−6.5%
9	+3.72%	88.1%	−6.2%
10	+3.51%	87.3%	−5.9%

The analysis reveals a clear trade-off between alpha magnitude and consistency:

- Lower N (1–3): Higher alpha but lower hit rates and larger drawdowns
- Higher N (7–10): Lower alpha but more stable performance
- **N=5**: Optimal balance with 92.8% hit rate and acceptable +4.88% alpha

We select N=5 as it maximizes the hit rate while maintaining meaningful alpha generation.

TABLE IV
DYNAMIC N VARIANTS

Variant	Ann. Alpha	Hit Rate	Max DD	p vs IC Only
IC Only (N=5)	+4.88%	92.8%	−7.3%	—
DynN Original	+5.84%	87.2%	−15.2%	0.089
DynN Floor-2	+5.21%	89.4%	−9.8%	0.142
DynN Floor-3	+5.02%	90.8%	−8.1%	0.231
DynN VeryConserv	+4.95%	91.2%	−7.8%	0.412

E. Dynamic N Analysis

We investigated whether dynamically adjusting N based on signal confidence could improve performance. Several variants were tested:

Key finding: No Dynamic N variant achieves statistical significance versus IC Only after Bonferroni correction. The apparent alpha improvements come with substantially increased drawdown risk. The Friedman test indicates variants are statistically different from each other (p = 0.0004), but pairwise comparisons cannot identify a clear winner.

F. Robustness Analysis

To verify that results are not driven by early lucky periods compounding over time, we conducted sub-period analysis.

TABLE V
YEAR-BY-YEAR PERFORMANCE

Year	Ann. Alpha	Hit Rate	Beats Baseline
2015	+4.21%	91.7%	Yes
2016	+5.82%	100.0%	Yes
2017	+3.94%	91.7%	Yes
2018	+4.56%	83.3%	Yes
2019	+5.12%	100.0%	Yes
2020	+6.38%	91.7%	Yes
2021	+3.21%	83.3%	Yes
2022	+4.87%	91.7%	Yes
2023	+4.02%	100.0%	Yes
2024	+5.24%	91.7%	Yes
All Years	**+4.88%**	**92.8%**	**10/10**

1) Year-by-Year Performance: The strategy outperforms the baseline in **every single year**, demonstrating consistent rather than episodic alpha generation.

2) Sub-Period Analysis: We divided the backtest into three equal periods to test for temporal stability:

TABLE VI
SUB-PERIOD PERFORMANCE

Period	Ann. Alpha	Hit Rate	Beats Baseline
Early (2015–2018)	+4.63%	91.7%	87% of months
Middle (2018–2021)	+4.82%	91.7%	85% of months
Late (2021–2025)	+5.18%	95.0%	89% of months

Performance is consistent across all three periods, with the late period actually showing slight improvement. This rules out the hypothesis that results are driven by early luck compounding.

3) Rolling Alpha Analysis: We computed 12-month rolling alpha difference versus baseline:

- Percentage of time with positive rolling alpha: 87.4%
- Average rolling alpha: +4.92% annualized
- Worst 12-month period: −1.2% (single occurrence)
- Best 12-month period: +11.4%

The strategy maintains positive rolling alpha in the vast majority of 12-month windows, confirming robustness.

G. Summary of Results

The walk-forward backtest and statistical analysis yield the following conclusions:

1) **IC weighting works:** The only statistically significant improvement (p = 0.047), adding +1.76% annualized alpha over baseline.
2) **N=5 is optimal:** Provides the best balance of alpha (+4.88%) and consistency (92.8% hit rate).
3) **Dynamic N not recommended:** Fails statistical significance and increases drawdown risk.
4) **Other improvements fail:** Stability weighting, time-series features, and combinations do not achieve significance.
5) **Results are robust:** Consistent performance across all years, sub-periods, and rolling windows.

V. Discussion

This section interprets the key findings from our walk-forward backtest and addresses important considerations for practical implementation.

A. Why IC Weighting Works

The Information Coefficient (IC) weighting mechanism achieved the only statistically significant improvement over the baseline. Several factors explain its effectiveness:

Adaptive feature selection: IC weighting naturally emphasizes features that have been predictive in recent history. Markets are non-stationary, and the features that predict alpha may change over time. By weighting features by their rolling IC, the ensemble adapts to current market conditions.

Noise reduction: Equal weighting treats all features identically, including those with weak or negative predictive power. IC weighting down-weights noisy features, effectively filtering the signal from the noise.

Robustness to overfitting: Rather than selecting a single "best" feature (which risks overfitting), IC weighting combines many features while emphasizing those with proven track records. This provides diversification across signal sources.

B. Why Dynamic N Failed Significance Tests

Despite showing higher raw alpha in backtests, Dynamic N selection did not achieve statistical significance. The key reasons are:

Increased variance: Dynamic N introduces additional variability in portfolio composition. When signal confidence is low, selecting fewer satellites increases concentration risk. When confidence is high, selecting more satellites dilutes alpha. This variability inflates the standard error of alpha estimates.

Risk-return trade-off: The apparent alpha improvement (+0.96% annualized) comes with doubled maximum drawdown (−15.2% vs −7.3%). On a risk-adjusted basis, the improvement is less compelling.

Multiple comparisons: When testing multiple Dynamic N variants (Original, Floor-2, Floor-3, VeryConservative), Bonferroni correction requires $p < 0.01$ for significance. None of the variants meet this threshold.

Practical implication: Given the lack of statistical significance and increased risk, we recommend the simpler IC-weighted approach with fixed N=5.

C. The Importance of Statistical Testing

A key contribution of this work is the application of rigorous statistical testing to strategy development. Without such testing, one might conclude that Dynamic N or other improvements are beneficial based solely on backtest performance.

The statistical tests reveal that:

- Only 1 of 6 tested improvements (IC weighting) achieves significance at the 5% level
- Backtest alpha can be misleading without significance testing
- Multiple comparison correction is essential when testing many strategies
- Bootstrap confidence intervals provide intuitive interpretation of uncertainty

This approach guards against overfitting and increases confidence that the selected strategy will perform well out-of-sample.

D. Robustness Considerations

The sub-period analysis provides strong evidence that results are not driven by early luck:

Temporal consistency: Alpha is positive in all 10 years tested, with no year showing underperformance versus baseline. This rules out the hypothesis that cumulative alpha is driven by a few exceptional years.

No degradation over time: The late period (2021–2025) actually shows slightly higher alpha than earlier periods. This suggests the strategy is not suffering from alpha decay as more market participants adopt similar approaches.

Rolling window analysis: Positive rolling alpha 87.4% of the time indicates consistent outperformance rather than occasional large gains offsetting frequent small losses.

E. Limitations

Several limitations should be acknowledged:

Sample period: While 120+ months provides reasonable statistical power, it represents only about 10 years of market history. The strategy has not been tested through all possible market regimes (e.g., prolonged bear markets, inflationary periods).

Transaction costs: We assume EUR 1–3 per trade, which is realistic for European retail investors using low-cost brokers. Higher transaction costs would reduce net alpha.

Market impact: With monthly rebalancing of 5 satellites, each representing 8% of portfolio value, market impact should be negligible for portfolios under EUR 1 million. Larger portfolios may experience some slippage.

Survivorship bias: Our ETF universe consists of currently available ETFs. ETFs that were delisted may not be fully represented, potentially introducing survivorship bias.

Data snooping: Despite walk-forward testing, the choice of signal bases, filters, and indicators was informed by domain knowledge of what has worked historically. True out-of-sample testing would require implementing the strategy prospectively.

F. Comparison with Alternative Approaches

Our signal-based approach differs from common alternatives:

Factor investing: Academic factor portfolios (value, momentum, quality) typically use stock-level data and require significant capital. Our ETF-based approach is more accessible to retail investors.

Technical analysis: Traditional technical analysis relies on discretionary pattern recognition. Our approach systematically evaluates thousands of signal combinations and applies statistical testing.

Machine learning: More complex ML models (neural networks, gradient boosting) could potentially capture non-linear relationships. However, they require more data to avoid overfitting and are less interpretable. The linear IC-weighting approach provides a good balance of performance and transparency.

G. Implementation Recommendations

Based on our findings, we recommend the following implementation:

1) **Use IC-weighted ensemble:** Weight features by their rolling 12-month IC when computing ensemble scores.
2) **Fix N=5:** Select exactly 5 satellites per month, avoiding dynamic adjustment.
3) **Maintain 60/40 allocation:** Keep 60% in ACWI core, 40% in satellites (8% each).
4) **Rebalance monthly:** Execute on the last trading day of each month.
5) **Monitor performance:** Track rolling alpha and hit rate to detect potential strategy degradation.
6) **Avoid over-optimization:** Resist the temptation to add complexity without statistical evidence of improvement.

VI. Conclusion

This document has presented a systematic approach to core-satellite portfolio management using signal-based ETF selection. Through comprehensive feature engineering, walk-forward backtesting, and rigorous statistical testing, we have identified an effective and robust strategy for generating consistent alpha.

A. Key Findings

1. IC-weighted feature ensembles generate statistically significant alpha.

The Information Coefficient (IC) weighting mechanism is the only tested improvement that achieves statistical significance at the 5% level (p = 0.047). By weighting features according to their rolling correlation with forward alpha, the ensemble adapts to changing market conditions while maintaining robustness.

2. Simplicity outperforms complexity.

Despite testing numerous enhancements—including dynamic satellite count, stability weighting, time-series features, and regime-adaptive signals—none achieved statistical significance beyond the basic IC-weighted approach. This supports the principle that simpler strategies are often more robust out-of-sample.

3. N=5 satellites provides optimal balance.

Analysis across $N \in \{1, \ldots, 10\}$ reveals that N=5 maximizes the hit rate (92.8%) while maintaining meaningful alpha (+4.88% annualized). Lower N increases alpha but reduces consistency; higher N provides stability at the cost of alpha dilution.

4. Results are temporally robust.

The strategy outperforms the baseline in all 10 years tested, with consistent alpha across early, middle, and late sub-periods. Rolling alpha is positive 87.4% of the time. This rules out early-luck compounding as the source of returns.

B. Final Strategy Specification

Based on our analysis, the recommended strategy is:

TABLE VII
FINAL STRATEGY PARAMETERS

Parameter	Value
Core allocation	60% (iShares MSCI ACWI)
Satellite allocation	40% (5 ETFs at 8% each)
Selection method	IC-weighted feature ensemble
Number of satellites	Fixed N=5
Rebalancing frequency	Monthly
Holding period	1 month

Expected performance:

- Annualized alpha: approximately +4.9% over ACWI
- Monthly hit rate: approximately 93%
- Maximum drawdown: approximately -7% relative to benchmark
- Sharpe ratio: approximately 1.07

C. Practical Implications

For long-term investors with a 30+ year horizon, this strategy offers several advantages:

Systematic execution: Clear rules eliminate emotional decision-making and ensure consistent implementation.

Tax efficiency: Irish-domiciled accumulating ETFs minimize dividend taxes. Infrequent trading (monthly) reduces capital gains events.

Low costs: With EUR 1–3 per trade and low-expense ETFs (TER $< 0.30\%$), transaction costs do not significantly erode alpha.

Scalability: The strategy works for portfolios from EUR 10,000 to EUR 1,000,000+ without modification.

Profit locking: Monthly rebalancing automatically transfers satellite gains into the stable core, reducing portfolio risk over time.

D. Limitations and Future Work

While the strategy demonstrates robust historical performance, several areas warrant further investigation:

Extended out-of-sample testing: Prospective implementation would provide the strongest evidence of strategy validity.

Alternative asset classes: The signal framework could be extended to bonds, commodities, or factor ETFs to further diversify alpha sources.

Machine learning integration: Non-linear models might capture additional predictive relationships, though care must be taken to avoid overfitting.

Transaction cost optimization: More sophisticated rebalancing rules could reduce turnover while maintaining signal freshness.

E. Closing Remarks

The core-satellite framework, combined with systematic signal-based selection, offers a compelling approach to long-term wealth accumulation. By prioritizing statistical rigor over backtest optimization, we have identified a strategy that balances alpha generation with robustness and simplicity.

The key insight from this work is that **statistical significance matters more than backtest performance**. Many apparent improvements fail to achieve significance when properly tested. By accepting only IC weighting—the sole statistically significant enhancement—we reduce the risk of implementing strategies that will not persist out-of-sample.

For investors seeking consistent outperformance over a multi-decade horizon, the IC-weighted core-satellite strategy provides a principled, evidence-based approach to portfolio construction.

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

We compute 167 signal bases from raw price data, organized into 27 categories. Each signal captures a different aspect of ETF behavior relative to the core benchmark (ACWI). Signals use day-based window parameters (21d, 63d, 126d, 252d) corresponding approximately to 1, 3, 6, and 12 months respectively.

Categories include:

- Returns and alpha (daily returns, excess returns, price ratios)
- Relative strength (multi-horizon momentum vs benchmark)
- Risk metrics (beta, volatility, drawdown)
- Risk-adjusted performance (Sharpe, Sortino, Information ratios)
- Trend indicators (price vs moving averages, golden cross)
- Mean reversion signals (z-scores, Bollinger bands, RSI, oversold indicators)
- Higher moments (skewness, kurtosis)
- Correlation and dispersion
- Momentum dynamics (acceleration, deceleration)
- Signal disagreement (short vs long-term momentum divergence)
- Regime-adaptive signals (volatility-boosted, trend-boosted)

Smoothing Filters

We apply 27 causal smoothing filter configurations to each signal base. All filters are strictly causal—they only use past data at each time step, ensuring no look-ahead bias.

Filter types:

- Raw (no filter) — 1 configuration
- Exponential Moving Average (EMA) — 21d, 63d
- Double EMA (DEMA) — 21d, 63d
- Triple EMA (TEMA) — 21d, 63d
- Zero-Lag EMA (ZLEMA) — 21d, 63d
- Hull Moving Average — 21d, 63d
- Triangular MA (TRIMA) — 21d, 63d
- Gaussian MA — 21d, 63d
- Butterworth Low-Pass — 21d, 63d
- Kalman Filter — fast, slow configurations
- Savitzky-Golay — 21d, 63d
- Median Filter — 21d, 63d
- Kaufman Adaptive MA (KAMA) — 21d
- Regime-Switching Filter — adaptive

Indicator Transformations

Each filtered signal undergoes 25 indicator transformations that capture different aspects of signal behavior:

- Level (raw filtered value)
- Momentum (5d, 21d, 63d, 126d changes)
- Z-scores (21d, 63d, 126d windows)
- Cross-sectional rank and z-score
- Trend ratios (MA ratios at multiple horizons)
- Mean reversion (distance from rolling means)
- Velocity and acceleration
- Drawdown (distance from rolling maximum)
- Range position (within high-low range)
- Volatility ratios (short vs long-term)
- Percentile (within rolling distribution)
- Relative volatility (vs benchmark)

Feature Space Summary

Combining all components yields the complete feature space:

- **Signal Bases**: 167 fundamental signals
- **Smoothing Filters**: 27 causal filter configurations
- **Indicators**: 25 transformations
- **Total Features**: $167 \times 27 \times 25 = 112{,}725$

Each feature is named using the convention: `signal__filter__indicator`, for example `rel_strength_126d__hull_63d__xs_rank`.

Computational Pipeline

The signal computation pipeline processes all 112,725 features efficiently using:

- Parallel processing with Python multiprocessing
- Vectorized numpy/pandas operations
- Parquet file caching for intermediate results
- SQLite database for signal storage
- Walk-forward architecture ensuring strict causality

Feature Selection for Ensemble

From the 112,725 features, we select the top-performing features for the IC-weighted ensemble based on:

- Historical average alpha generation
- IC Information Ratio > 0.5 (consistency)
- Hit rate $> 55\%$ (better than random)
- Low correlation with other selected features (diversification)

The final ensemble typically includes 50–100 features, weighted by their rolling 12-month IC values.