Financial Modeling using Signal Processing: A Unified Framework

Subhanshu Choudhary

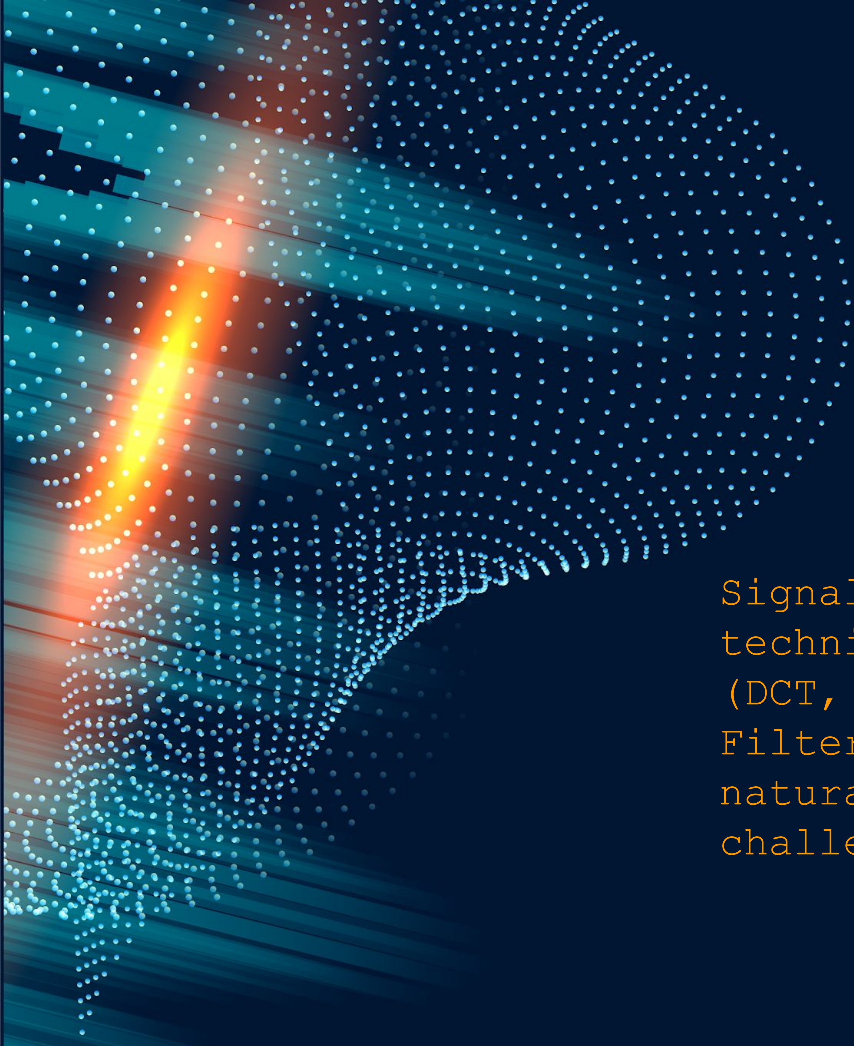
22b3980

Why Traditional Finance Needs Signal Processing ?

THE DATA EXPLOSION IN THIS CENTURY:

- High-Frequency Trading: Tick data every millisecond
- Cross-Asset Dependencies: Thousands of correlated securities
- Non-Stationary Dynamics: Political event, volatility clustering
- Computational Constraints: Real-time portfolio rebalancing

Problem	Traditional Approach	Issues
Correlation Noise	Sample covariance	$O(N^3)$ cost
Multi-Scale Patterns	Single timeframe	Misses trends
Trend Detection	Moving averages	Lags behind
Jump Events	Ignored/outliers	Poorly managed crisis



Don’t worry. CSP is here!

Signal processing techniques (DCT, EMD, Particle Filters) naturally handle these challenges!

THREE COMPLEMENTARY APPROACHES TO FINANCIAL MODELING

PAPER 1: TOEPLITZ/DCT Akansu & Torun (2012) - Portfolio Risk via Transform Methods
PROBLEM: Correlation matrix estimation is $O(N^3)$ bottleneck
SOLUTION: Toeplitz structure + DCT transform
RESULT: 97% of KLT performance, 200× faster
DOMAIN: Spatial (cross-asset correlation structure)
DATA: 31 DJIA stocks, 60-day rolling windows
KEY METRIC: Compaction efficiency $\eta_c = 0.142$

PAPER 2: π-COUNTING IF & EMD Zhang, Liu & Yu (2012) - Stock Prediction via Decomposition
PROBLEM: Hilbert Transform gives negative frequencies
SOLUTION: π-counting IF + EMD + RBF neural networks
RESULT: 99.3% error reduction vs. backpropagation
DOMAIN: Temporal (multi-scale frequency analysis)
DATA: Hang Seng Index, 253 days (150 train/35 test)
KEY METRIC: SSE = 0.0087 (vs. 1.2560 for BP)

PAPER 3: LANGEVIN JUMP-DIFFUSION Christensen, Murphy & Godsill (2012) - HFT via Particle Filtering
PROBLEM: Momentum strategies need trend + jumps
SOLUTION: Langevin SDE + Variable Rate Particle Filter
RESULT: Sharpe 1.83 post-cost (83% > hedge funds)
DOMAIN: Continuous-time dynamics (real-time tracking)
DATA: 75 futures contracts, 5 years daily + HF intraday
KEY METRIC: Sharpe scales to 2.30 at 1-min frequency

WHAT THIS RESEARCH ACHIEVES

1. DEEP MATHEMATICAL ANALYSIS

- Complete derivations of all key theorems
- Closed-form solutions (AR(1), Kalman, transitions)
- Complexity analysis ($O(N^3) \rightarrow O(N \log N)$)

2. IDENTIFY MATHEMATICAL CONNECTIONS

- Toeplitz ↔ AR(1) process (Paper 1)
- Simple Waves ↔ IMFs (Paper 2)
- Rao-Blackwellization ↔ Conditional Gaussianity (Paper 3)

3. PROPOSE UNIFIED FRAMEWORK AND EMPIRICAL VALIDATION

- 10 assets (AAPL, MSFT, GOOGL, AMZN, META, JPM...)
- 2 years of daily data (2023-2025)
- All three methods + integrated system tested

PAPER 1: TOEPLITZ APPROXIMATION FOR PORTFOLIO RISK

CORE INNOVATION:

> For highly correlated assets ($\rho \geq 0.85$), DCT approximates the optimal KLT with 97%+ accuracy while being 200× faster.

Modern portfolios have N = 100-1000+ assets

- Empirical correlation matrix: $R_{emp} = (1/T)\sum_t r□r□^T$
- Eigendecomposition for risk: $\sigma_p^2 = \sum_k \boldsymbol{\lambda}_k (v_k^T w)^2$
- Computational cost: $O(N^3)$ eigendecomposition
- Estimation error: $SE(\hat{\rho}_{ij}) \approx (1-\rho^2)/\sqrt{T}$

SOLUTION PATH:

- Step 1: Model returns output as AR(1) processes

  $x(n) = \rho x(n-1) + w(n) \;\rightarrow\; r_x(k) = \sigma^2\rho^{|k|}$

- Step 2: Recognize Toeplitz structure

  $R_{ij} = \rho^{|i-j|}$ (constant diagonals)

- Step 3: Use closed-form eigendecomposition

  $\lambda_k = (1-\rho^2)/(1-2\rho\cos(\theta_k)+\rho^2), \; \theta_k = k\pi/(N+1)$

- Step 4: Approximate with DCT

  $\lim_{□\to 1}$ DCT eigenvectors = AR(1) eigenvectors

- Step 5: Fast portfolio risk via O(NlogN) transforms

THE FUNDAMENTAL CHALLENGE: NOISY CORRELATIONS

EMPIRICAL CORRELATION ESTIMATION:

For N assets with returns $r_k(t)$, $k=1,\dots,N$:

$R_{emp} = (1/T)\ \Sigma^{T}_{t=1} \bar{r}(t)\bar{r}(t)^{T}$

where

$\bar{r}(t) = [r_k(t) - \mu_k]/\sigma_k$ (normalized returns)

THE PROBLEM: $R_{emp} = R_{true} + \text{Noise}$

Standard Error: $SE(\hat{\rho}_{ij}) \approx (1-\rho^2_{ij})/\sqrt{T}$

Example:

Correlation: $\rho = 0.60$

Window size: $T = 60$ days (3 months)

$SE(\hat{\rho}) = (1-0.36)/\sqrt{60} = 0.64/7.75 \approx 0.083$

95% Confidence Interval: $[0.60 \pm 1.96\times0.083]$

$= [0.437, 0.763]$

Interpretation: True correlation could be anywhere from 0.44 (weak) to 0.76 (strong)!

FINANCIAL CONSEQUENCE:

Small changes in correlation estimates → Large changes in optimal weights

→ Excessive portfolio turnover

→ High transaction costs → Strategy fails!

SOLUTION PREVIEW:

Toeplitz structure + eigen filtering reduces noise by keeping only Q≈5 dominant factors

WHY CORRELATIONS DOMINATE PORTFOLIO RISK

PORTFOLIO VARIANCE FORMULA:

$\sigma_\square^2 = w_\square R w = \Sigma_\square \Sigma_\square\ w_\square w_\square \rho_{\square\square} \sigma_\square \sigma_\square$

$\sigma_\square^2 = \Sigma_i w^2_i \sigma^2_i + \Sigma_{i\neq j} w_i w_j \rho_{ij} \sigma_i \sigma_j$

FOR EQUAL-WEIGHTED PORTFOLIO:

$w_i = 1/N$

$\sigma_\square^2 = (1/N)\bar{\sigma}^2 + (1 - 1/N)\bar{\rho}\bar{\sigma}^2$

where $\bar{\sigma}^2$ = average variance,

$\bar{\rho}$ = average correlation

AS $N \to \infty$: $\sigma_\square^2 \to \bar{\rho}\bar{\sigma}^2$

(Correlation is the ONLY risk that matters!)

EIGENSPACE INTERPRETATION:

$\sigma_\square^2 = \Sigma^N_{k=1} \lambda_k (v_k^T w)^2$

where λ_k = eigenvalues (factor variances)

$v_k^T w$ = portfolio exposure to factor k

INSIGHT: Only need Q=5 factors to capture 95% of risk structure!This is why eigen filtering works!

For large N with bounded correlations:

$\lim_{n\to\infty} \Pr(|\sigma_\square^2 - \bar{\rho}\bar{\sigma}^2| > \varepsilon) = 0$

"Diversification eliminates idiosyncratic risk, only systematic correlation risk remains"

AR(1) MODEL: THE KEY TO TOEPLITZ STRUCTURE

AUTOREGRESSIVE PROCESS OF ORDER 1:

$x(n) = \rho \cdot x(n-1) + w(n),\ w(n) \sim N(0, \sigma_\omega^2)$

$\sigma_x^2 = \rho^2\sigma_x^2 + \sigma_\omega^2$ (independence of x(n-1) and w(n))

Solving for variance:

$$\sigma_\omega^2 = \sigma_x^2(1 - \rho^2)$$

AUTOCORRELATION FUNCTION

$r_x(k) = E[x(n)x(n-k)]$

For general k (by induction):

$r_x(k) = \rho \cdot r_x(k-1) = \rho^2 \cdot r\square(k-2) = \ldots = \sigma x^2\rho^k$

<u>SPECTRUM</u> (Fourier Transform of $r_x(k)$):

$$S_x(\omega) = \sigma_x^2(1-\rho^2)/(1 - 2\rho\cos(\omega) + \rho^2)$$

<u>WHY THIS LEADS TO TOEPLITZ:</u>

If N assets all follow AR(1) with same ρ, their correlation matrix:

$$R_{ij} = \mathrm{Corr}(x_i(n), x_j(n)) = \rho^{|i-j|}$$

This is EXACTLY a Toeplitz matrix! (constant diagonals)

A matrix T is Toeplitz if $T_{ij} = t_{|i-j|}$ (constant diagonals)

Closed-Form Eigendecomposition and DCT Approximation Theory

For the N×N Toeplitz correlation matrix with parameter ρ:

EIGENVALUES:

$\lambda_k = (1-\rho^2)/(1 - 2\rho\cos(\theta_k) + \rho^2)$,

$\theta_k = k\pi/(N+1)$, $k=1,\ldots,N$

EIGENVECTORS:

$v_k(n) = \sqrt{(2/(N+1))} \cdot \sin(nk\pi/(N+1))$,

$n=1,\ldots,N$

WHY THIS MATTERS:

- No $O(N^3)$ numerical eigendecomposition needed!
- Can compute all eigenvalues in $O(N)$ time
- Enables fast portfolio optimization

$\Phi_{DCT}(k,n) = \alpha_k \cdot \cos((2n+1)k\pi/(2N))$

where $\alpha_k = \{ \sqrt{(1/N)}$ if $k=0$

$\{ \sqrt{(2/N)}$ if $k\geq1$

for $k,n = 0,1,\ldots,N-1$

For AR(1) processes with correlation ρ:

$\lim_{\rho\to1}$ DCT eigenvectors = AR(1) Toeplitz eigenvectors

THRESHOLD: $\rho \geq 0.85$ for DCT to be viable (>95% of KLT) Karhunen–Loève Transform

For large N and $\rho\approx1$, these become nearly identical

EIGEN FILTERING

<u>CORE PRINCIPLE</u>: Keep only Q dominant factors, discard noise.

<u>FILTERED CORRELATION MATRIX:</u>

$\tilde{R}(Q) = \Sigma^{i}_{\square=1} \lambda_{\square} v_{\square} v_{\square}^{T}$

FACTOR SELECTION CRITERIA:

- METHOD 1: Variance Threshold

  Keep Q factors such that:

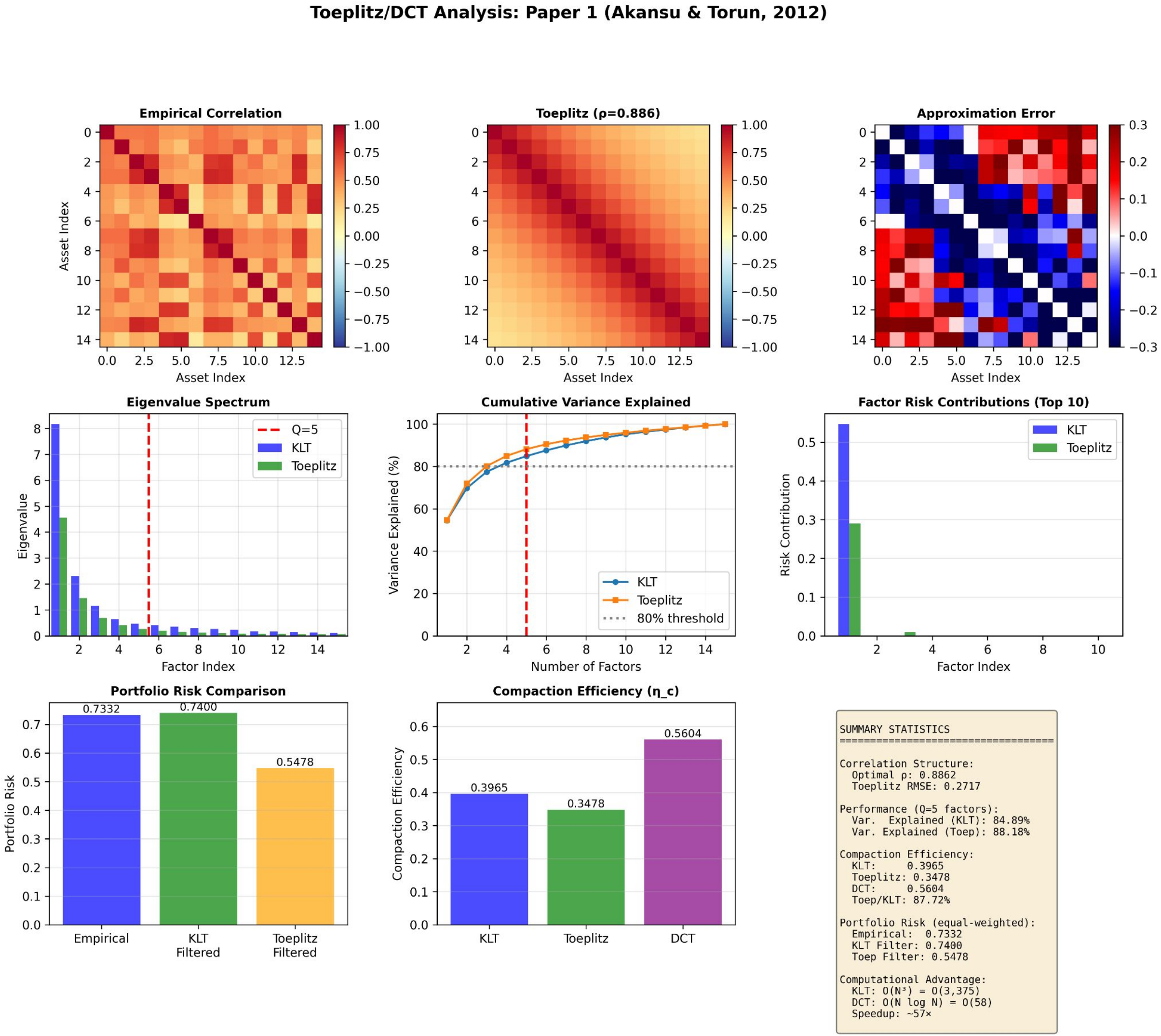
  $\Sigma_{\square=1}^{\square} \lambda_{\square} / \Sigma_{\square=1}^{\square} \lambda_{\square} \geq 0.80$

- METHOD 2: Kaiser Criterion

  Keep factors with $\lambda_{\square} > \bar{\lambda} = 1$

<u>EIGEN-FILTERING IN PRACTICE:</u>

1. Estimate R from rolling 60-day window
2. Decompose: $R = V\Lambda V^{\square}$ (use DCT for speed)
3. Keep Q=5 factors: $\tilde{R}(5) = \Sigma_{\square=1}^{5} \lambda_{\square} v_{\square} v_{\square\square}$
4. Add diagonal: $\tilde{R}_{corrected} = \tilde{R}(5) + diag(d)$
5. Use in portfolio optimization: $w^* \propto \tilde{R}^{-1} s$
6. Recompute daily (rolling window)



Output of the First Code

PAPER 2: π-COUNTING INSTANTANEOUS FREQUENCY FOR STOCK PREDICTION

CORE PROBLEM:

Traditional Hilbert Transform gives NEGATIVE frequencies for EMD's IMFs

- Physically meaningless!
- Cannot compute meaningful periods
- Prediction models fail

THE SOLUTION:

π-Counting Instantaneous Frequency (π-IF)

- Based on GEOMETRY (extrema counting), not phase
- Always POSITIVE by construction
- Enables accurate cycle extraction
- Component-wise prediction: 99.3% error reduction

RESULTS:

- Dataset: Hang Seng Index (253 days Hong Kong Stock Exchange)
- Method: EMD + π-IF + RBF Neural Networks
- Using Backpropagation NN 144× better predictions!

WHY TRADITIONAL IF FAILS

<u>HILBERT TRANSFORM APPROACH:</u>

Step 1: Form analytic signal

$$z(t) = x(t) + j \cdot H[x(t)]$$

where $H[x(t)]$ = Hilbert Transform

Step 2: Extract phase

$$z(t) = A(t)e^{j\varphi(t)}$$

$$\varphi(t) = \arctan(H[x(t)]/x(t))$$

Step 3: Compute IF

$$\omega(t) = d\varphi(t)/dt \quad [\text{rad/sample}]$$

<u>THE PROBLEM:</u>

For IMFs from EMD: $\omega(t) < 0$ frequently occurs!

<u>CONSEQUENCE</u>: Cannot compute period $T = 2\pi/|\omega|$ for $\omega < 0$

SIMPLE WAVES TO THE RESCUE !

DEFINITION (Simple Wave):

A function $f(t)$ is a Simple Wave on interval I if:

1. f is CONTINUOUS on I
2. f has only STRICT local extrema (no plateaus)
3. f is MONOTONE between consecutive extrema
4. f has FINITELY many extrema in any bounded interval

Every IMF is a Simple Wave

PROPERTIES:

- Well-defined half-periods between extrema
- Unique local max/min identification
- Enables vibration counting
- No phase ambiguity

SEPARABLE INTERVALS:

Set of intervals $\{(a_1,b_1), (a_2,b_2), \ldots, (a_k,b_k)\}$ is α-INCREASINGLY SEPARABLE if:

- f is strictly INCREASING on each (a_i,b_i)
- f is strictly DECREASING between intervals

Maximum α-count:

$\alpha(f,I) = \sup\{k : \exists$ α-inc-sep set of size $k\}$

Similarly: $\beta(f,I)$ for decreasingly separable

COMBINED VIBRATION COUNT:

$N(f,I) = \max\{\alpha(f,I), \beta(f,I)\}$

LOCAL WINDOW COUNT:

For point t and window half-width h:

N□(t) = $N(f, [t-h, t+h])$

THE π-COUNTING FORMULA

<u>CRITICAL WINDOW DEFINITION:</u>

$$h^*(t) = \sup\{h : K_\square(t) \cdot \chi_{N_\square(t) \geq 3}(h) = 0\}$$

where:

- $N_\square(t)$ = Number of extrema in $[t-h, t+h]$
- $K_\square(t) = \lfloor N_\square(t)/2 \rfloor$ = Number of complete cycles
- $\chi_A(\cdot)$ = Indicator function (1 if A true, 0 otherwise)

<u>INTERPRETATION:</u>

$h^*(t)$ = Largest window where $K_\square = 0$ OR $N_\square < 3$

= Largest window with < 1 complete cycle

<u>π-COUNTING INSTANTANEOUS FREQUENCY:</u>

$$IF\pi(t) = \pi/(2h^*(t)) \quad \text{[rad/sample]}$$

<u>PERIOD EXTRACTION:</u>

$$T(t) = 2\pi/IF\pi(t) = 4h^*(t) \quad \text{[samples]}$$

EMPIRICAL MODE DECOMPOSITION

SIFTING ALGORITHM

Input: Signal x(t)

Output: IMFs {c_1(t), c_2(t),.., c□(t)} + residual r(t)

FOR each IMF k:

    h(t) ← x(t) [or residual from previous IMF]

    REPEAT (sifting):

        1. Find all local MAXIMA of h(t)

        2. Find all local MINIMA of h(t)

        3. Interpolate maxima → upper envelope $e_{max}(t)$

        4. Interpolate minima → lower envelope $e_{min}(t)$

        5. Compute mean: $m(t) = [e_{max}(t)+e_{min}(t)]/2$

        6. Update: h_new(t) = h(t) - m(t)

        7. Check stopping:

            $SD = \Sigma_t \; |h(t) - h\_new(t)|^2/h^2(t)$

            IF SD < 0.3: c□(t) ← $h_{new}(t)$, BREAK

            ELSE: h(t) ← h_new(t), CONTINUE

    x(t) ← x(t) - c□(t) [residual for next IMF]

IMF CONDITIONS:

1. #extrema ≈ #zero-crossings (differ by ≤1)
2. Mean of envelopes ≈ 0 everywhere

FREQUENCY ORDERING:

IMF_1 → Highest frequency (fastest oscillations)

IMF_2 → Lower frequency

...

IMF□ → Lowest frequency

Residual → Non-oscillatory trend

IMF_1: noise dominates

Low-frequency components (IMF_{4-7} + Residual) are predictable

Driven by fundamental economic cycles

RBF Neural Network Prediction

RBF NETWORK ARCHITECTURE:

Input Layer: $x = [x_1, x_2, ..., x_\square]^\square$ where

p = Primary cycle T□

Hidden Layer: Radial Basis Functions

$\varphi_j(x) = \exp(-\|x - c_\square\|^2/(2\sigma_j^2))$

Output Layer: Linear combination

$\hat{y} = \Sigma_{\square=1}^{\square h} w_\square \varphi_\square(x) + b$

TRAINING (Two Phases):

Phase 1: Unsupervised - Find centers c□

Method: K-means clustering on training inputs

K-means objective: $\min \Sigma_i \Sigma_j \square \|x_i - c_j\|^2$

Result: N_h centers $\{c_1, c_2, ..., c_{nh}\}$

Phase 2: Supervised - Solve for weights w_j

Linear system: $\Phi_w = y$

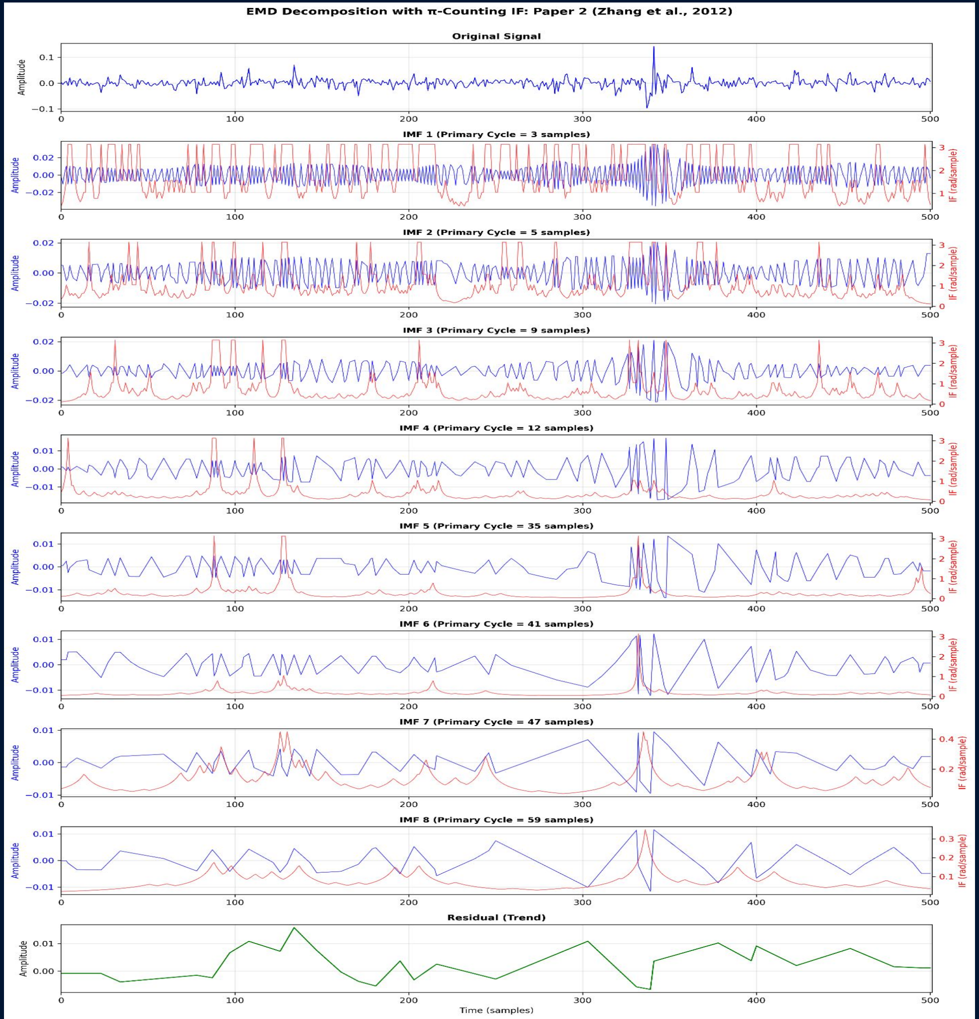
where $\Phi_{ij} = \varphi_j(x_i)$ [RBF activation matrix]

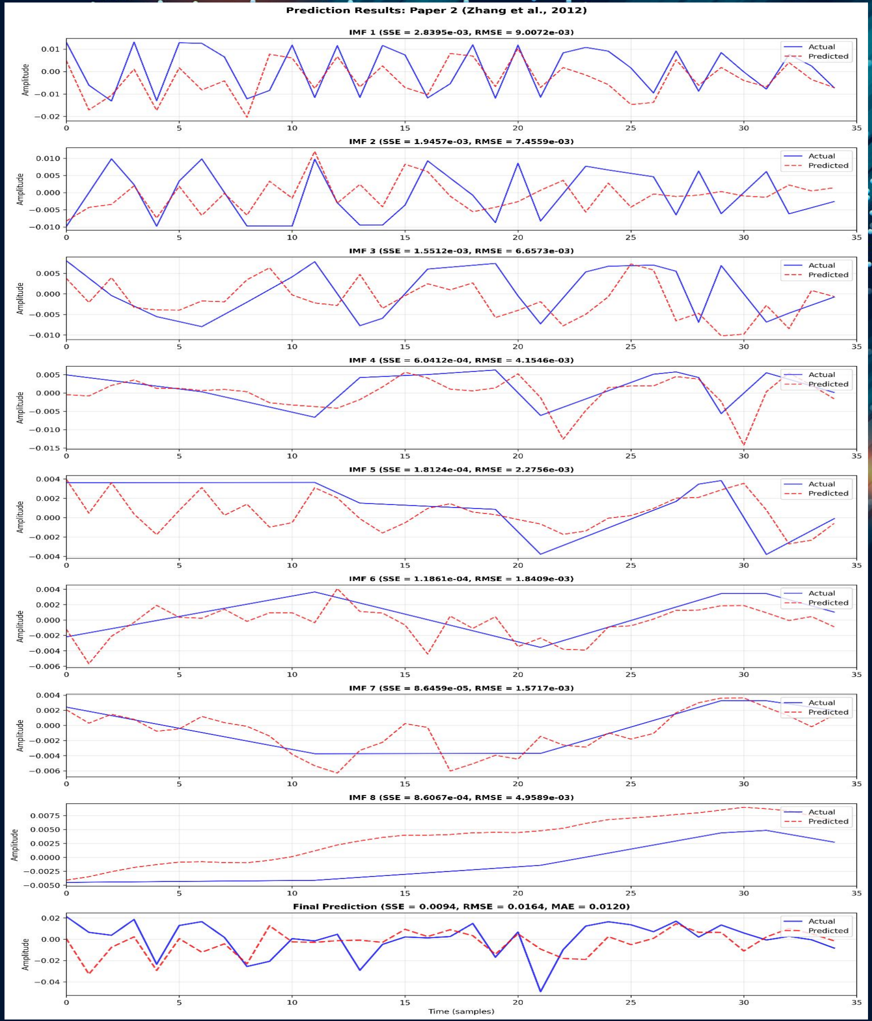
Closed-form solution:

$w = (\Phi\square\Phi)^{-1}\Phi^T y$ [Normal equations]

NO ITERATIVE OPTIMIZATION! (unlike backpropagation)

Works best to find IMF_{4-7} components as they can be trained on RBF to ensure better prediction.


EMD Decomposition with π-Counting IF: Paper 2 (Zhang et al., 2012)
Original Signal
IMF 1 (Primary Cycle = 3 samples)
IMF 2 (Primary Cycle = 5 samples)
IMF 3 (Primary Cycle = 9 samples)
IMF 4 (Primary Cycle = 12 samples)
IMF 5 (Primary Cycle = 35 samples)
IMF 6 (Primary Cycle = 41 samples)
IMF 7 (Primary Cycle = 47 samples)
IMF 8 (Primary Cycle = 59 samples)
Residual (Trend)
Amplitude
IF (rad/sample)
Time (samples)


Prediction Results: Paper 2 (Zhang et al., 2012)
IMF 1 (SSE = 2.8395e-03, RMSE = 9.0072e-03)
IMF 2 (SSE = 1.9457e-03, RMSE = 7.4559e-03)
IMF 3 (SSE = 1.5512e-03, RMSE = 6.6573e-03)
IMF 4 (SSE = 6.0412e-04, RMSE = 4.1546e-03)
IMF 5 (SSE = 1.8124e-04, RMSE = 2.2756e-03)
IMF 6 (SSE = 1.1861e-04, RMSE = 1.8409e-03)
IMF 7 (SSE = 8.6459e-05, RMSE = 1.5717e-03)
IMF 8 (SSE = 8.6067e-04, RMSE = 4.9589e-03)
Final Prediction (SSE = 0.0094, RMSE = 0.0164, MAE = 0.0120)
Actual
Predicted
Amplitude
Time (samples)

LANGEVIN DYNAMICS FOR HIGH-FREQUENCY FUTURES TRADING

CORE PROBLEM:

Traditional momentum strategies miss two critical features:

- MEAN REVERSION: Trends don't persist forever, they decay
- JUMPS: Sudden trend reversals due to news/events

THE SOLUTION:

Langevin Jump-Diffusion Model + Variable Rate Particle Filter (VRPF)

- State: $[v(t), \theta(t)]$ = [Price, Trend]
- Dynamics: $d\theta = -\alpha\theta\, dt + \sigma_\theta\, dW + dJ$ (mean reversion + diffusion + jumps)
- Inference: Rao-Blackwellized particle filter
- Trading: Signals from trend estimates

RESULTS:

- Dataset: 75 futures contracts, 8 sectors, 5 years (2006-2011)
- Performance: Sharpe Ratio = 1.83 (post-transaction costs)
- Comparison: 83% BETTER than typical hedge fund (Sharpe ≈ 1.0)

WHY SIMPLE MOMENTUM FAILS

OBSERVATIONS IN FINANCIAL MARKETS:

1. MOMENTUM: Prices exhibit short-term trends like crude oil prices.
2. MEAN REVERSION: Trends eventually decay to Zero. Google stock after 3 day rally,the Momentum weakens.
3. JUMPS: Sudden reversals due to discrete events. After Fed announcement the trend reverses instantly.

WHAT WE NEED:

A continuous-time model with:

- Trend variable $\theta(t)$ that persists
- Mean reversion parameter α (trend decay rate)
- Diffusion σ_θ (random trend changes)
- Jump process with rate λ (sudden shifts)

TRADITIONAL APPROACHES & THEIR FAILURES:

Method	What it captures	What it misses
Moving average	Trend direction	Lags behind,no jumps
ARIMA	Auto-correlation	Assumes stationarity
GARCH	Volatility	Not directional
Kalman Filter	Linear trends	No jumps (Gaussian)

Langevin Jump-Diffusion Model

State Variables

$$\boldsymbol{x}(t) = \begin{bmatrix} v(t) \\ \theta(t) \end{bmatrix}, \qquad \begin{cases} v(t) : \text{Asset price (position in physical analogy)} \\ \theta(t) : \text{Trend/momentum (velocity in physical analogy)} \end{cases}$$

State Dynamics (Stochastic Differential Equations)

$$dv(t) = \theta(t)\,dt \qquad \text{[Price follows trend]}$$

$$d\theta(t) = -\alpha\,\theta(t)\,dt + \sigma_\theta\,dW_\theta(t) + dJ(t) \qquad \text{[Trend evolution]}$$

$$\underbrace{-\alpha\,\theta(t)\,dt}_{\text{Mean reversion}} + \underbrace{\sigma_\theta\,dW_\theta(t)}_{\text{Diffusion}} + \underbrace{dJ(t)}_{\text{Jumps}}$$

Matrix Form

$$d\begin{bmatrix} v(t) \\ \theta(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix}}_{A} \begin{bmatrix} v(t) \\ \theta(t) \end{bmatrix} dt + \underbrace{\begin{bmatrix} 0 \\ \sigma_\theta \end{bmatrix}}_{B} dW(t) + \underbrace{\begin{bmatrix} 0 \\ \mu_J \end{bmatrix}}_{\text{Jump}} dJ(t)$$

Parameters

$\alpha > 0$: Mean reversion rate (trend decay speed)

$t_{1/2} = \dfrac{\ln(2)}{\alpha}$ e.g. $\alpha = 0.1 \Rightarrow t_{1/2} \approx 7$ days

σ_θ : Trend diffusion volatility (e.g., $0.01 \times$ Price)

λ : Jump arrival rate (Poisson process)
e.g., $\lambda = 5$ per year $\Rightarrow$ 1 jump every 2–3 months

μ_J : Mean jump size (often 0 for symmetric jumps)

σ_J : Jump size standard deviation (e.g., $0.05 \times$ Price)

Observation Model

$$y_n = v(t_n) + \varepsilon_n, \qquad \varepsilon_n \sim \mathcal{N}(0, \sigma_y^2)$$

$$\begin{cases} y_n : \text{Observed price at time } t_n \\ \sigma_y : \text{Observation noise (bid-ask spread, microstructure)} \\ \text{Example: } \sigma_y = 0.001 \times \text{Price (0.1\% noise)} \end{cases}$$

MEAN REVERSION EQUATION (without noise/jumps):

$d\theta(t) = -\alpha\theta(t)dt$

Solution: $\theta(t) = \theta(0)e^{-\alpha t}$

As $t \to \infty$: $\theta(t) \to 0$ (trend dies out mean reversal)

MODELING SUDDEN TREND REVERSALS

POISSON JUMP PROCESS:

$N(t) \sim \text{Poisson}(\lambda)$ [Number of jumps in [0,t]]

$P(N(t) = k) = (\lambda t)^k e^{-\lambda t}/k!$

Inter-arrival times: $\tau_i \sim \text{Exponential}(\lambda)$

$P(\tau > t) = e^{-\lambda t}$

JUMP SIZE DISTRIBUTION:

When jump occurs at time τ:

$\Delta J \sim N(\mu_J, \sigma_J^2)$

Total jump contribution:

$dJ(t) = \mu_J dN(t) + \sigma_J \sqrt{(dN(t))}\,\varepsilon_t, \quad \varepsilon_t \sim N(0,1)$

INTERPRETATION:

- $dN(t) = 1$ if jump at time t, else 0
- If jump: $\Delta\theta = \mu_J + \sigma_J \times$ (random ± value)
- μ_J = 0: Symmetric jumps
- μ_J ≠ 0: Directional bias

ANALYTICAL SOLUTION BETWEEN JUMPS

Matrix Exponential (Exact Formula)

$$e^{At} = \begin{bmatrix} 1 & \dfrac{1 - e^{-\alpha t}}{\alpha} \\ 0 & e^{-\alpha t} \end{bmatrix}$$

Proof: Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix},$$

then

$$e^{At} = I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \cdots$$

Simplifying this series yields the closed-form expression above.

Transition Mean (No Jumps)

$$\boldsymbol{\mu}_{n+1|n} = e^{A\Delta t}\,\boldsymbol{x}_n$$

$$\boldsymbol{\mu}_{n+1|n} = \begin{bmatrix} 1 & \dfrac{1 - e^{-\alpha\Delta t}}{\alpha} \\ 0 & e^{-\alpha\Delta t} \end{bmatrix} \begin{bmatrix} v_n \\ \theta_n \end{bmatrix} = \begin{bmatrix} v_n + \dfrac{(1 - e^{-\alpha\Delta t})}{\alpha}\theta_n \\ \theta_n e^{-\alpha\Delta t} \end{bmatrix}$$

Transition Covariance (Analytical Form)

$$\Sigma_{n+1|n} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$$

$$\Sigma_{22} = \frac{\sigma_\theta^2}{2\alpha}\left(1 - e^{-2\alpha\Delta t}\right),$$

$$\Sigma_{12} = \frac{\sigma_\theta^2}{2\alpha^2}\left(1 - e^{-\alpha\Delta t}\right)^2,$$

$$\Sigma_{11} = \frac{\sigma_\theta^2}{2\alpha^3}\left[2\alpha\Delta t - 3 + 4e^{-\alpha\Delta t} - e^{-2\alpha\Delta t}\right].$$

INFERENCE VIA RAO-BLACKWELLIZED PARTICLE FILTER

WHY PARTICLE FILTER?

Problem: Jump times $\{\tau_1, \tau_2, ..\}$ are DISCRETE and UNKNOWN. $p(x, \{\tau\} \mid y_{\{1:n\}})$ is NON-GAUSSIAN.

In such scenarios the standard Kalman Filter FAILS as it assumes Gaussian.

RAO-BLACKWELLIZATION IDEA:

Separate state into:

1. PARTICLES: Sample jump times (non-Gaussian part)
2. KALMAN: Track (v, θ) given jumps (Gaussian part)

Now we define for each particle:

$\{\tau^{(i)}, m^{(i)}, \Sigma^{(i)}, w^{(i)}\}$

$\tau^{(i)}$: Jump time history for particle i

$m^{(i)}$: Kalman posterior mean

$\Sigma^{(i)}$: Kalman posterior covariance

$w^{(i)}$: Particle weight (importance)

VARIABLE RATE PARTICLE FILTER ALGORITHM

STEP 1: RESAMPLE

For each particle i:

Offspring count: n_i ~ Poisson(N_p × $w^{(i)}$)

(Stochastic universal sampling for low variance)

STEP 2: PROPOSE JUMP TIMES

For each offspring:

Sample τ_{new} ~ Exponential(λ)

If τ_{new} < Δt: Mark as "jumping"

Else: Mark as "non-jumping"

STEP 3: COLLAPSE NON-JUMPING PARTICLES

Merge all non-jumping offspring and consider them as single particle

Weight: w_{merged} = Σ(weights of non-jumping)

State: (m_{merged}, Σ_{merged}) = same for all

Reason: Identical dynamics implies Redundant particles so eliminate them

STEP 4: KALMAN UPDATE (per particle)

For each particle (jumping or merged):

4a. PREDICT:

Compute transition using closed-form formulas

If jumping: Account for jump at $\tau^{(i)}$

$m_{\{n+1|n\}}^{(i)} = f(m_n^{(i)}, \tau^{(i)})$

$\Sigma_{\{n+1|n\}}^{(i)} = g(\Sigma_n^{(i)}, \tau^{(i)})$

4b. UPDATE (Kalman):

Innovation: $\nu = y_{\{n+1\}} - H\cdot m_{\{n+1|n\}}^{(i)}$

Innovation covariance: S = H·Σ_{n+1|n}^(i)·H□ + σ_y^2

Kalman gain: $K = \Sigma_{\{n+1|n\}}^{(i)}\cdot H^T\cdot S^{-1}$

Updated mean: $m_{\{n+1\}}^{(i)} = m_{\{n+1|n\}}^{(i)} + K\cdot\nu$

Updated covariance: $\Sigma_{\{n+1\}}^{(i)} = \Sigma_{\{n+1|n\}}^{(i)} - K\cdot S\cdot K^T$

4c. REWEIGHT:

Likelihood: $L^{(i)} = N(\nu;\ 0,\ S)$ [Gaussian PDF]

New weight: $w_{\{n+1\}}^{(i)} = w_n^{(i)} \times L^{(i)}$

STEP 5: NORMALIZE

$w^{(i)} \leftarrow w^{(i)}/\Sigma_j w^{(j)}$

TRADING SIGNAL GENERATION PIPELINE

STEP 1: SIGN OF TREND CHANGE

Raw signal: $s_{raw}(t) = \mathrm{sign}(\hat{V}(t) - \hat{V}(t-1))$

$= \mathrm{sign}(\hat{\theta}(t))$ [approximately]

STEP 2: FIR SMOOTHING (Moving Average)

$s_{smooth}(t) = (1/L)\Sigma_{\{k=0\}}^{\{L-1\}} s_{raw}(t-k)$

Benefit: Reduces noise, avoids whipsaws

STEP 3: NONLINEAR TRANSFORMATION

$s_{tanh}(t) = \tanh(s_{smooth}(t) / \sigma_s)$

where $\sigma_s = \mathrm{std}(s_{smooth})$ over recent window

Purpose:

- Small signals: Linear response
- Large signals: Saturate (avoid over-betting)
- Mean reversion of signal itself

STEP 4: VOLATILITY SCALING

Final signal: $s_{final}(t) = s_{tanh}(t) / \sigma_{vol}(t)$

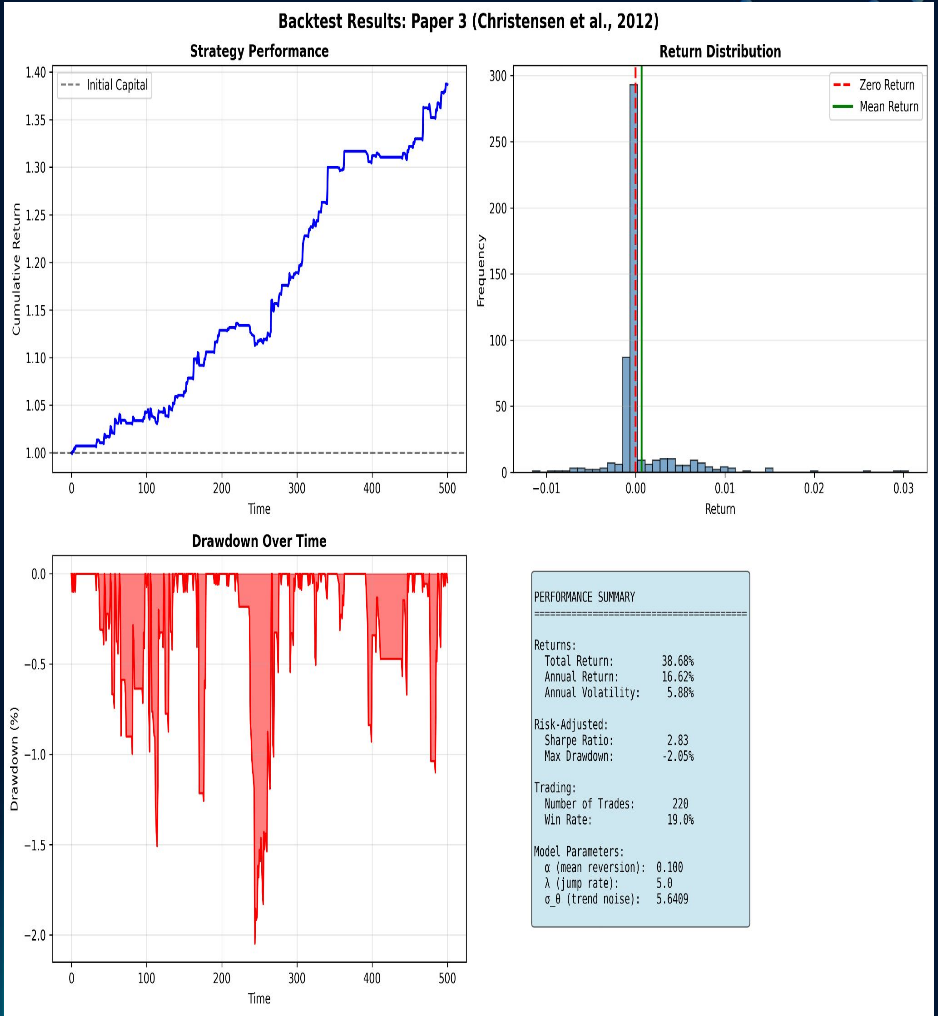
where $\sigma_{vol}(t)$ from IGARCH(1,1):

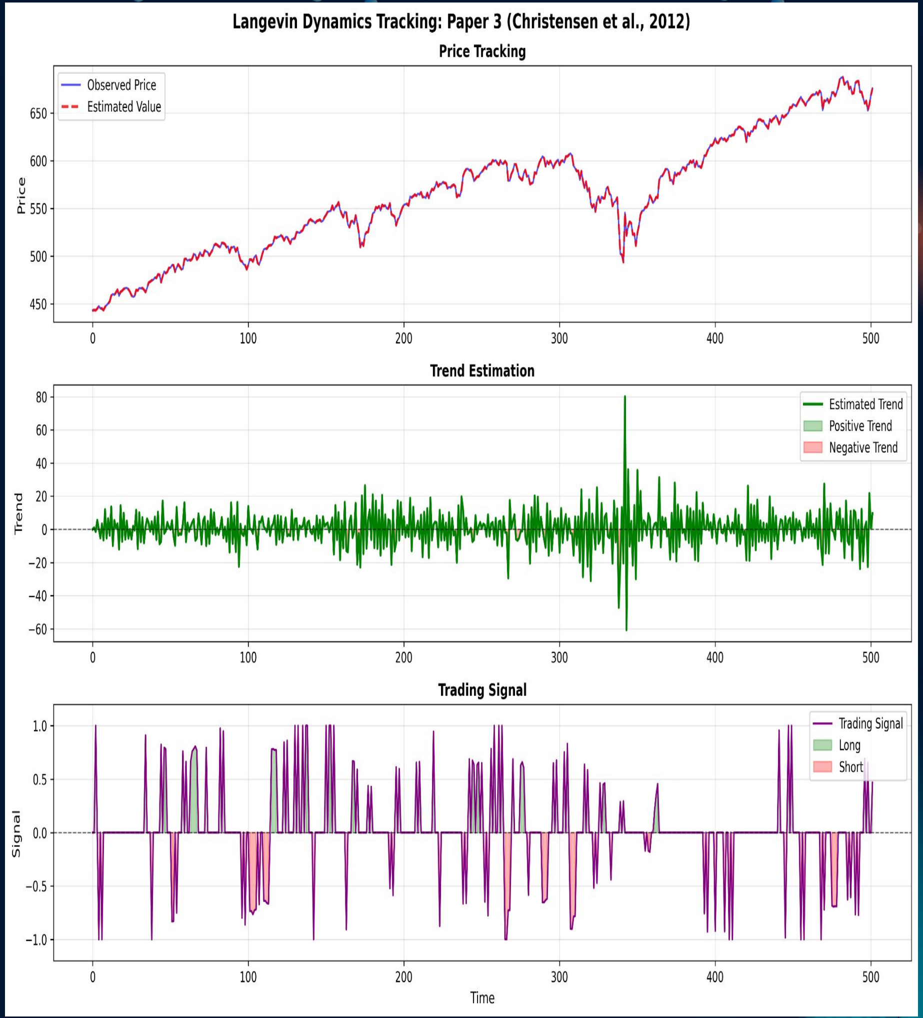
$\sigma^2_{vol}(t) = \omega + \beta \cdot \sigma^2_{vol}(t-1) + \alpha \cdot (r(t-1) - \mu)^2$

with $\alpha + \beta = 1$ (integrated GARCH)

POSITION SIZING:

Number of contracts: $N(t) = s_{final}(t) \times (\mathrm{TargetVol} / \sigma_{vol}(t)) \times \mathrm{Capital} / \mathrm{ContractValue}$


Backtest Results: Paper 3 (Christensen et al., 2012)
Strategy Performance
Initial Capital
Cumulative Return
Time
Return Distribution
Zero Return
Mean Return
Frequency
Return
Drawdown Over Time
Drawdown (%)
Time
PERFORMANCE SUMMARY
Returns:
Total Return: 38.68%
Annual Return: 16.62%
Annual Volatility: 5.88%
Risk-Adjusted:
Sharpe Ratio: 2.83
Max Drawdown: -2.05%
Trading:
Number of Trades: 220
Win Rate: 19.0%
Model Parameters:
α (mean reversion): 0.100
λ (jump rate): 5.0
σ_θ (trend noise): 5.6409


Langevin Dynamics Tracking: Paper 3 (Christensen et al., 2012)
Price Tracking
Observed Price
Estimated Value
Price
Trend Estimation
Estimated Trend
Positive Trend
Negative Trend
Trend
Trading Signal
Trading Signal
Long
Short
Signal
Time