BTC-Based Crypto Pairs Trading Strategy

Gan Wei Siang
Chen Tong
Zhai Jiayi
Lu Yao

NUS Investment Society
Quantitative Finance Department
2026

Abstract

This report develops a systematic, Bitcoin-anchored pairs trading strategy applied to large-cap cryptocurrencies. To address the highly volatile and non-stationary nature of crypto markets, we employ a walk-forward framework on high-frequency spot data, ensuring strict separation between training and testing periods. The proposed approach integrates dynamic cointegration screening, time-varying hedge ratio estimation via a Kalman filter, and mean-reversion modeling of spreads using an Ornstein–Uhlenbeck process. A regime filter is introduced to suspend trading during periods in which spread stationarity deteriorates.

Out-of-sample backtesting under realistic transaction costs and execution constraints highlights the importance of regime-aware trading. While all strategies exhibit negative risk-adjusted performance over the sample period, the regime-filtered strategy significantly improves capital preservation relative to the non-regime variant and reduces drawdown compared to a buy-and-hold Bitcoin benchmark, although its risk-adjusted performance remains weaker than Bitcoin. The equal-weight portfolio achieves a final equity of approximately USD 9,035, compared to USD 574 without regime filtering and USD 7,399 for buy-and-hold Bitcoin. Maximum drawdown is reduced to approximately −15.8%, versus −94.3% and −38.3%, respectively.

While several individual pairs exhibit strong profitability, these opportunities are sparse and are diluted at the portfolio level by losses in other pairs and transaction costs. These findings suggest that regime conditioning serves primarily as an effective risk-management mechanism in cryptocurrency statistical arbitrage, although the generation of consistent positive alpha remains challenging in the presence of execution frictions and rapidly evolving market dynamics.

1 Introduction

The cryptocurrency market has evolved into a highly liquid yet structurally fragmented ecosystem. Despite rapid maturation, the market remains prone to persistent inefficiencies driven by retail sentiment shocks, uneven information diffusion, and segmented liquidity across a multitude of exchanges. Within this environment, market-neutral statistical arbitrage, specifically pairs trading, offers a compelling framework to systematically harvest alpha while explicitly hedging against broad directional market risks.

Bitcoin (BTC) consistently dominates the digital asset space, functioning as the primary benchmark and anchor for the entire asset class. Most large-cap altcoins share a robust, though occasionally transient, co-movement relationship with BTC due to shared market-wide macroeconomic sensitivities and cross-exchange arbitrage linkages. When temporary mispricings occur, the price spread between BTC and an anchored altcoin deviates from its historical equilibrium, creating a tangible mean-reversion opportunity.

However, traditional pairs trading methodologies, which rely on static hedge ratios and assume permanent cointegration, are fundamentally ill-suited for the cryptocurrency market. Regime shifts in this space are frequent and violent; structural correlations can collapse dramatically during coin-specific catalyst events. To address these unique challenges, this study proposes a systematic, dynamic BTC-anchored pairs trading strategy utilizing high-frequency data.

We construct a tradeable universe of large-cap assets on Binance and employ a rigorous rolling walk-forward research design to prevent overfitting. By utilizing a Kalman Filter, we dynamically estimate time-varying hedge ratios that adapt to evolving market conditions in real-time. The resulting spread is modelled via an Ornstein-Uhlenbeck (OU) process to mathematically quantify mean-reversion speed and equilibrium volatility. Crucially, we introduce a real-time Regime Filter that acts as a strict risk-management gating mechanism, ensuring capital is deployed only during statistically validated stationary phases and forcing positions flat during structural divergences.

The remainder of this report is organized as follows: Section 2 establishes the theoretical framework underpinning our statistical models. Section 3 details the comprehensive methodology, from universe construction and dynamic hedging to signal generation and walk-forward backtesting. Section 4 presents the performance evaluation across various large-cap pairs under realistic execution constraints, and Section 5 concludes with key strategic takeaways and areas for future improvements.

2 Theoretical Framework

The cryptocurrency market presents a structurally distinct setting for statistical arbitrage. Unlike traditional equity markets, which benefit from centralized price discovery, deep institutional participation, and continuous regulatory oversight, cryptocurrency markets remain fragmented across exchanges, highly sensitive to retail sentiment shocks, and susceptible to idiosyncratic volatility in individual assets. Within this environment, Bitcoin (BTC) occupies a singular role: consistently accounting for more than 50% of total market capitalization, it functions simultaneously as a risk barometer, a portfolio anchor, and the primary transmission channel through which market-wide sentiment propagates to large-cap altcoins. This structural dominance creates persistent, testable co-movement relationships between BTC and assets such as AVAX, LINK, ETH, and others. When these relationships temporarily deviate from their equilibrium, they generate spread dislocations that may be systematically exploited. This paper develops a unified framework for identifying, modeling, and trading such dislocations.

2.1 Market Inefficiency and Statistical Arbitrage in Crypto

The efficient market hypothesis in its weak form implies that asset prices fully reflect all historical information, leaving no exploitable statistical patterns. While this benchmark is a reasonable approximation for mature equity markets, the cryptocurrency market violates several of its preconditions. Liquidity is segmented across hundreds of exchanges, arbitrage capital is limited relative to market size, and information diffusion is uneven across participant types. The result is a market in which temporary mispricings arise not from irrationality alone but from structural frictions.

Pairs trading is a market-neutral statistical arbitrage strategy designed to exploit these frictions. By simultaneously holding a long position in the relatively underpriced asset and a short position in the overpriced one, the strategy eliminates exposure to broad market direction, isolating the relative pricing error as the sole source of return. In the BTC-anchored context of this study, large-cap altcoins frequently deviate from their equilibrium co-movement with BTC due to three overlapping mechanisms: shared market-wide sentiment shocks that impact each asset asymmetrically, portfolio rebalancing flows that temporarily disconnect price ratios, and cross-exchange arbitrage linkages that enforce corrections only with a lag. These mechanisms collectively support the hypothesis that BTC-altcoin spreads are mean-reverting and tradable.

2.2 Cointegration and Long-Run Price Equilibrium

The theoretical backbone of pairs trading is cointegration theory. Two price series x_t and y_t are each said to be integrated of order one, $I(1)$, if their levels are non-stationary but their first differences are stationary. The series are cointegrated if there exists a scalar β such that the linear combination

$$s_t = y_t - \beta x_t \tag{1}$$

is itself stationary, $I(0)$. Stationarity of s_t implies the existence of a long-run attractor: however far the spread drifts, it is eventually drawn back towards its equilibrium level. This property is the statistical foundation for mean-reversion trading.

The Engle–Granger two-step procedure provides a practical test for cointegration. In the first step, a hedge ratio $\hat{\beta}$ is estimated by ordinary least squares (OLS):

$$y_t = \alpha + \beta x_t + \varepsilon_t. \tag{2}$$

In the second step, the residual series $\hat{\varepsilon}_t$ is tested for stationarity using the Augmented Dickey–Fuller (ADF) test. Rejection of the unit root null at conventional significance levels confirms cointegration. In this study, BTC serves as the anchor regressor x_t for all candidate pairs, reflecting its structural role as the dominant pricing factor across the crypto asset universe.

Importantly, cointegration is treated here as a *transient operational condition* rather than a permanent economic law. A pair may pass a full-sample cointegration test while the relationship deteriorates over sub-periods, particularly during strong unidirectional trends driven by idiosyncratic catalysts. This motivates the use of both static and rolling cointegration diagnostics as a dynamic formation-stage filter, discussed further in Section 3.3.

2.3 The Ornstein–Uhlenbeck Process as a Spread Model

Once a stationary spread s_t is identified, its dynamics require a stochastic model that captures mean reversion explicitly. The Ornstein–Uhlenbeck (OU) process is the canonical continuous-time choice [1]:

$$ds_t = \kappa(\mu - s_t)\, dt + \sigma\, dW_t, \tag{3}$$

where $\mu \in \mathbb{R}$ is the long-run equilibrium level, $\kappa > 0$ is the speed of mean reversion, $\sigma > 0$ is the diffusion coefficient, and W_t is a standard Wiener process. The restoring force $\kappa(\mu - s_t)$ ensures that the process is pulled back towards μ at a rate proportional to both the displacement and κ: large deviations generate strong reversion pressure, while small deviations decay slowly.

Discretising equation (3) over an interval δ yields the autoregressive form used in estimation:

$$s_{t+1} = a\, s_t + b + \varepsilon_t, \tag{4}$$

where the parameters map back to the continuous-time representation as

$$a = e^{-\kappa\delta}, \qquad b = \mu\big(1 - e^{-\kappa\delta}\big), \qquad \mathrm{sd}(\varepsilon_t) = \sigma\sqrt{\frac{1 - e^{-2\kappa\delta}}{2\kappa}}. \tag{5}$$

Inverting these relationships gives the structural OU parameter estimates:

$$\hat{\kappa} = -\frac{\ln a}{\delta}, \qquad \hat{\mu} = \frac{b}{1 - a}, \qquad \hat{\sigma} = \mathrm{sd}(\varepsilon_t)\sqrt{\frac{-2\ln a}{\delta(1 - a^2)}}. \tag{6}$$

A key derived quantity is the *half-life* of mean reversion, defined as the expected time for the spread to close half the distance to its equilibrium from any initial displacement:

$$\tau_{1/2} = \frac{\ln 2}{\hat{\kappa}}. \tag{7}$$

For a 15-minute bar dataset, a half-life below 192 bars corresponds to reversion within two trading days, which is a necessary condition for the spread to generate sufficient round-trips within a realistic trading horizon. Once OU parameters are estimated, the spread is standardised into a dimensionless z-score:

$$z_t = \frac{s_t - \hat{\mu}}{\hat{\sigma}}, \tag{8}$$

which forms the direct input to the signal generation module.

2.4 Time-Varying Hedge Ratios via the Kalman Filter

Equation (2) imposes a static hedge ratio β, implying that the relationship between BTC and any altcoin is constant through time. This is untenable in cryptocurrency markets, where structural relationships shift materially across regimes. A static OLS estimate fitted on a 90-day training window captures an average relationship that may be increasingly stale by the end of that window.

To address this, the hedge ratio is modeled as a latent time-varying state. The observation equation is

$$y_t = \alpha_t + \beta_t \, x_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \, V), \tag{9}$$

and the state evolution equations are

$$\alpha_t = \alpha_{t-1} + \eta_t^{\alpha}, \qquad \eta_t^{\alpha} \sim \mathcal{N}(0, \, Q_\alpha), \tag{10}$$

$$\beta_t = \beta_{t-1} + \eta_t^{\beta}, \qquad \eta_t^{\beta} \sim \mathcal{N}(0, \, Q_\beta). \tag{11}$$

The Kalman filter provides the minimum mean-square-error recursive estimate $\hat{\beta}_t$ by alternating between a prediction step, which propagates the prior state estimate forward, and an update step, which incorporates the new observation to correct the posterior. The resulting spread is:

$$s_t = y_t - \left(\hat{\alpha}_t + \hat{\beta}_t \, x_t \right). \tag{12}$$

Critically, $\hat{\alpha}$ and $\hat{\beta}$ are estimated *exclusively on the training window* and then frozen for application to the subsequent out-of-sample test window. This prevents any form of look-ahead contamination in the spread construction.

2.5 Regime Dependence and Adaptive Filtering

Even when a pair satisfies the cointegration criterion at formation, the mean-reversion property of its spread can deteriorate as market conditions evolve. During trending regimes, for instance, when a macro catalyst causes AVAX to decouple from BTC, the OU model assumption is violated and entries based on z-score thresholds are likely to incur losses rather than profits.

This motivates the introduction of a *regime indicator* that continuously re-evaluates whether the spread is behaving as a stationary, mean-reverting process. At each rolling window, the spread is subjected to an ADF test and its OU half-life is re-estimated. The indicator takes the value one (tradable) only if both

$$p_{\mathrm{ADF}} < 0.10 \quad \text{and} \quad 0 < \tau_{1/2} < 192 \text{ bars} \tag{13}$$

are simultaneously satisfied, and zero otherwise. The final trading signal is then:

$$\text{Signal}_t^{\text{final}} = \text{Signal}_t^{\text{raw}} \times \mathbf{1}[\mathcal{R}_t = 1] \, , \tag{14}$$

where $\mathcal{R}_t$ is the regime indicator. When $\mathcal{R}_t = 0$, all existing positions in the pair are forced flat. Skipped windows should therefore be interpreted not as model failures but as deliberate risk management: the filter correctly identifies when the statistical conditions underpinning the trade are absent, preserving capital for windows in which they hold.

3 Methodology

3.1 Research Design

The empirical investigation is structured as a fully systematic, research-to-deployment pipeline. The pipeline is designed to replicate the constraints of realistic live trading: all parameters are estimated exclusively on historical data, and performance is evaluated exclusively on data the model has never seen. The pipeline consists of six sequential stages:

1. **Universe construction** — select and filter eligible BTC-anchored pairs.
2. **Cointegration screening** — rank candidates by static and rolling cointegration diagnostics.
3. **Dynamic hedge ratio estimation** — fit a Kalman filter on the training window.
4. **Spread modelling** — estimate OU parameters and validate stationarity conditions.
5. **Signal generation and threshold optimisation** — grid-search over entry/exit thresholds on training data.
6. **Walk-forward out-of-sample evaluation** — apply frozen parameters to the subsequent test window and roll forward.

The walk-forward design enforces a strict temporal firewall between in-sample parameter estimation and out-of-sample performance measurement. It also acknowledges a core empirical reality of high-volatility crypto markets: estimated parameters decay in predictive power rapidly, and must be re-estimated at each iteration to remain aligned with the prevailing volatility regime.

3.2 Data Construction

The tradable universe is constructed from the top 20 cryptocurrencies by market capitalisation, retrieved from CoinGecko. This list is then intersected with assets that maintain an active USDT spot trading pair on Binance, ensuring that every asset in the universe is both liquid and directly executable. Stablecoins and wrapped tokens are excluded by construction, as their price dynamics are governed by peg maintenance mechanisms rather than market forces.

Price data consist of 15-minute close prices sourced from the Binance spot market, covering a rolling 90-day lookback horizon. The choice of 15-minute bars reflects a trade-off between capturing intraday mean-reversion opportunities and mitigating microstructure noise inherent in higher-frequency data. All candidate pairs share BTC as the common anchor asset, consistent with BTC's role as the dominant pricing factor in the crypto universe. The key data specifications are summarised in Table 1.

Table 1: Data specification summary.

Parameter	Specification
Bar frequency	15-minute close prices
Data source	Binance spot market
Universe source	CoinGecko top-20 by market cap
Lookback horizon	90 days
Anchor asset	BTC (Bitcoin)
Execution venue	Binance USDT pairs

3.3 Cointegration Screening

At the formation stage of each walk-forward iteration, all candidate BTC pairs are evaluated using both a static and a rolling Engle–Granger cointegration framework.

The **static test** is applied over the full 90-day training window and provides an omnibus assessment of whether the pair exhibits a long-run equilibrium relationship. The **rolling test** divides the training window into shorter sub-windows and applies the same procedure sequentially, producing a pass rate—the fraction of sub-windows in which the null of no cointegration is rejected at the 10% level. The rolling pass rate captures whether the relationship is stable through time or merely an artefact of a particular sub-period.

Each pair is assigned a composite score that combines four diagnostic components:

- Static cointegration p-value (lower is better),
- Rolling pass rate (higher is better),
- Average and latest rolling p-values (lower is better),
- Residual half-life (shorter is better, within bounds).

Pairs are ranked by this composite score and only the top-ranked candidates are advanced to the modelling stage. This two-dimensional screening, which requires both statistical significance and temporal stability, guards against spurious relationships driven by coincidental co-movement over a single sub-period.

3.4 Dynamic Hedge Ratio Estimation

For each pair passing the screening stage, a time-varying hedge ratio is estimated on the training window using the Kalman filter state-space model described in equations (9)–(12). The transition noise parameter $\delta = 10^{-4}$ governs the rate at which the filter allows the state to evolve; a small value implies that the hedge ratio is assumed to be locally stable, adapting gradually rather than reacting sharply to each new observation.

The filter is initialised using an OLS estimate of β computed over the first segment of the training window. After running through the full training period, the terminal state estimates $(\hat{\alpha}_T, \hat{\beta}_T)$ are frozen and applied to construct the spread over the subsequent test window without any further updating. This ensures that no test-period information contaminates the hedge ratio estimate.

3.5 Spread Modelling and Tradability Validation

The spread $s_t = y_t - (\hat{\alpha}_T + \hat{\beta}_T x_t)$ is modelled as a discrete OU process by estimating the AR(1) regression in equation (4) via OLS on the training window. The regression coefficients (a, b) and residual standard deviation are mapped to the structural OU parameters $(\hat{\kappa}, \hat{\mu}, \hat{\sigma})$ via equations (5)–(6), from which the half-life $\tau_{1/2}$ is computed via equation (7).

A pair is declared *tradable* on a given training window only if all three of the following conditions hold simultaneously:

- The ADF test on the training spread yields $p < 0.10$, confirming stationarity,
- The estimated mean-reversion speed satisfies $\hat{\kappa} > 0$, confirming that the AR coefficient $a < 1$ and deviations decay rather than compound, and
- The half-life satisfies $0 < \tau_{1/2} < 192$ bars, ensuring that mean reversion is fast enough to be economically exploitable within the test window.

Pairs failing any condition are skipped for that iteration and re-evaluated in the next walk-forward window. The tradable spread is then standardised into the OU z-score via equation (8), which forms the direct input to signal generation.

3.6 Signal Generation and Parameter Optimisation

Trading signals are generated by applying symmetric threshold rules to the OU z-score z_t:

- **Enter short spread** if $z_t \geq z_{\text{entry}}$ (sell the altcoin, buy BTC in proportion $\hat{\beta}_T$),
- **Enter long spread** if $z_t \leq -z_{\text{entry}}$ (buy the altcoin, sell BTC in proportion $\hat{\beta}_T$),
- **Exit** when $|z_t| \leq z_{\text{exit}}$, subject to a minimum holding period to avoid excessive turnover,
- **Force-exit** when $|z_t| \geq 4.0$, to bound loss on non-reverting spreads.

The threshold parameters are selected via a grid search conducted *entirely on the training window*:

$$\begin{aligned} z_{\text{entry}} &\in \{2.0,\ 2.5,\ 3.0\}, \\ z_{\text{exit}} &\in \{1.0,\ 1.5\}, \\ \text{min hold} &\in \{16,\ 24\} \text{ bars}, \\ \delta &\in \{10^{-4}\}. \end{aligned} \tag{15}$$

A parameter configuration is admissible only if the in-sample backtest satisfies all of: (i) at least two completed trades, (ii) positive Sharpe ratio, and (iii) positive PnL. Among admissible configurations, the one with the highest training Sharpe ratio is selected. Parameters are re-optimised from scratch at every walk-forward iteration using only the current training window, acknowledging that the optimal threshold regime shifts alongside changes in spread volatility and reversion speed.

3.7 Regime Filter

As detailed in Section 3, a real-time regime filter gates all raw signals. At each rolling sub-window within the test period, an ADF test is applied to the recent spread and the half-life is re-estimated. The regime indicator $\mathcal{R}_t \in \{0, 1\}$ is set to one only when both conditions in equation (13) are satisfied. When $\mathcal{R}_t = 0$, all positions are unwound immediately and no new entries are permitted.

The regime filter serves two purposes. First, it prevents capital deployment during non-stationary episodes in which the OU assumption breaks down, reducing the probability of catastrophic drawdowns from persistent spread divergence. Second, and perhaps more subtly, it allows the strategy to distinguish between *tradable noise*, transient deviations generated by short-lived liquidity imbalances, and *dangerous divergence* driven by structural regime changes, such as the temporary decoupling of an altcoin from BTC during a coin-specific catalyst event.

3.8 Walk-Forward Backtesting Framework

The strategy is evaluated using a rolling walk-forward procedure. At each iteration k:

1. A 90-day training window $[t_k,\ t_k + T_{\text{train}}]$ is used to screen pairs, estimate the Kalman filter, fit OU parameters, and optimise signal thresholds.
2. The resulting parameters are applied without modification to the subsequent test window $[t_k + T_{\text{train}},\ t_k + T_{\text{train}} + T_{\text{test}}]$.
3. The window is advanced by T_{test} and the process repeats.

The strict separation between training and test data at every iteration guards against in-sample overfitting while permitting the model to adapt to evolving market dynamics as windows roll forward.

All simulations are conducted under the execution assumptions summarised in Table 2, which are calibrated to reflect realistic conditions on the Binance spot market.

Table 2: Execution assumptions used in the backtest.

Parameter	Value
Initial capital	USD 10,000
Leverage	3×
Capital allocation	Full capital per active pair
Trading fee	0.04% (4 bps) per trade
Slippage	1 basis point
Execution lag	One bar (prevents same-bar look-ahead)
Position sizing	Scaled by $\hat{\beta}_T$ for dollar neutrality
Benchmark	Buy-and-hold BTC

3.9 Performance Evaluation

Strategy performance is assessed at both the individual-pair and portfolio levels. At the portfolio level, individual pair returns are aggregated into an equal-weight portfolio and benchmarked against a passive buy-and-hold BTC strategy over the same sample period.

The following metrics are reported:

- **Total PnL** and **final equity** — absolute performance in USD,
- **Annualised Sharpe ratio** — risk-adjusted return,
- **Maximum drawdown** — worst peak-to-trough equity decline,
- **Win rate** — fraction of completed trades that generate positive PnL,
- **Number of trades** — proxy for statistical reliability of the results,
- **Average turnover** — measure of trading activity and implicit transaction cost burden.

The IS-OOS Sharpe differential is also reported for each pair as a diagnostic of overfitting risk. A large positive differential indicates that training-window optimisation failed to generalise to the test period, consistent with parameter instability in a high-volatility regime.

4 Results

4.1 Rolling Window Performance Overview

The strategy was evaluated using a rolling walk-forward backtesting framework across multiple BTC-anchored pairs. Each iteration consisted of a training window for parameter estimation followed by an out-of-sample test window for performance evaluation.

The results indicate that not all windows produced tradable signals. A significant number of windows were skipped due to either the absence of valid parameter configurations or failure to satisfy transient cointegration criteria. Specifically, windows labeled as `skipped_no_valid_params` reflect periods where no admissible parameter set met the minimum requirements of profitability and statistical robustness, while `not_selected_by_transient_co` indicates weak or unstable cointegration relationships.

	pair	train_start	train_end	test_start	test_end	status	entry_z	exit_z	min_hold	delta	train_sharpe	train_total_pnl	test_sharpe	test_total_pnl
0	BTC-ETH	2026-01-03 14:15:00	2026-01-28 14:00:00	2026-01-28 14:15:00	2026-01-29 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	BTC-ETH	2026-01-04 14:15:00	2026-01-29 14:00:00	2026-01-29 14:15:00	2026-01-30 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	BTC-ETH	2026-01-05 14:15:00	2026-01-30 14:00:00	2026-01-30 14:15:00	2026-01-31 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	BTC-ETH	2026-01-06 14:15:00	2026-01-31 14:00:00	2026-01-31 14:15:00	2026-02-01 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	BTC-ETH	2026-01-07 14:15:00	2026-02-01 14:00:00	2026-02-01 14:15:00	2026-02-02 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	BTC-ETH	2026-01-08 14:15:00	2026-02-02 14:00:00	2026-02-02 14:15:00	2026-02-03 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	BTC-ETH	2026-01-09 14:15:00	2026-02-03 14:00:00	2026-02-03 14:15:00	2026-02-04 14:00:00	skipped_no_valid_params	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	BTC-ETH	2026-01-10 14:15:00	2026-02-04 14:00:00	2026-02-04 14:15:00	2026-02-05 14:00:00	traded	3.0	1.0	24.0	0.0001	3.073617	891.225396	-36.194902	-442.915745
8	BTC-ETH	2026-01-11 14:15:00	2026-02-05 14:00:00	2026-02-05 14:15:00	2026-02-06 14:00:00	traded	3.0	1.5	24.0	0.0001	0.689752	158.555823	NaN	0.000000
9	BTC-ETH	2026-01-12 14:15:00	2026-02-06 14:00:00	2026-02-06 14:15:00	2026-02-07 14:00:00	not_selected_by_transient_coint	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Table 3: Summary of Rolling Window Outcomes

Status	**Count**
not_selected_by_transient_coint	340
skipped_no_valid_params	261
traded	179

These observations highlight the strong regime dependency of the strategy and validate the importance of incorporating both cointegration screening and regime filtering mechanisms to avoid trading during structurally unfavorable periods.

4.2 Out-of-Sample Trading Behavior

Across rolling windows, several consistent patterns emerge. First, even when valid signals were identified during the training phase, many test windows resulted in zero executed trades. This suggests that the statistical conditions required for trade entry were not met out-of-sample, reinforcing the conservative nature of the signal generation process.

Second, a noticeable decline in performance from in-sample to out-of-sample results is observed. This gap reflects the inherent instability of parameter estimates in high-volatility crypto markets and indicates the presence of overfitting risk during the parameter optimization stage. These findings justify the adoption of a walk-forward framework, where parameters are continuously re-estimated to adapt to evolving market regimes.

4.3 Cross-Asset Performance Comparison

Performance varies significantly across different BTC-anchored pairs, demonstrating that not all cointegrated relationships are equally tradable.

	pair	final_equity	total_pnl	gross_pnl	total_fees	sharpe	max_drawdown	num_trades	win_rate	windows_selected	windows_traded
8	BTC-XLM	10945.000279	945.000279	1074.916194	129.915915	10.039692	-0.021596	4	0.750000	26	9
10	BTC-ZEC	10426.549685	426.549685	738.303958	311.754273	3.021593	-0.089842	10	0.800000	17	13
1	BTC-XRP	9969.100870	-30.899130	29.835345	60.734475	-0.423931	-0.017544	2	0.500000	39	13
3	BTC-SOL	9756.610617	-243.389383	-89.527231	153.862152	-1.967708	-0.058424	5	0.400000	60	19
9	BTC-LTC	9760.005270	-239.994730	179.711705	419.706435	-2.030000	-0.061869	14	0.357143	56	24
6	BTC-BCH	9696.883691	-303.116309	234.946193	538.062503	-4.324578	-0.037125	18	0.222222	15	12
11	BTC-AVAX	9659.442313	-340.557687	-220.577699	119.979988	-5.755570	-0.034056	4	0.000000	47	15
7	BTC-LINK	8372.528884	-1627.471116	-1258.066419	369.404697	-9.705532	-0.162747	13	0.230769	54	26
5	BTC-ADA	9409.983487	-590.016513	-530.684592	59.331921	-9.831699	-0.059002	2	0.000000	13	9
0	BTC-ETH	8843.205333	-1156.794667	-543.098850	613.695818	-14.266292	-0.116807	22	0.181818	41	23
2	BTC-BNB	9419.331395	-580.668605	-112.896382	467.772223	-16.736571	-0.059164	16	0.062500	48	14

The best-performing pairs include:

- **BTC–XLM**: Sharpe ratio of approximately 10.04 with a total PnL of around +945,
- **BTC–ZEC**: Sharpe ratio of approximately 3.02 with a total PnL of around +426.

In contrast, several pairs underperformed:

- Most of the selected pairs exhibited negative Sharpe ratios,
- BTC–BNB recorded the weakest performance with a Sharpe ratio of approximately -16.7 and substantial losses.

Additionally, turnover across some of the pairs remains low, typically ranging between one to five trades per window. While this reduces transaction costs, it also limits the statistical reliability of the results.

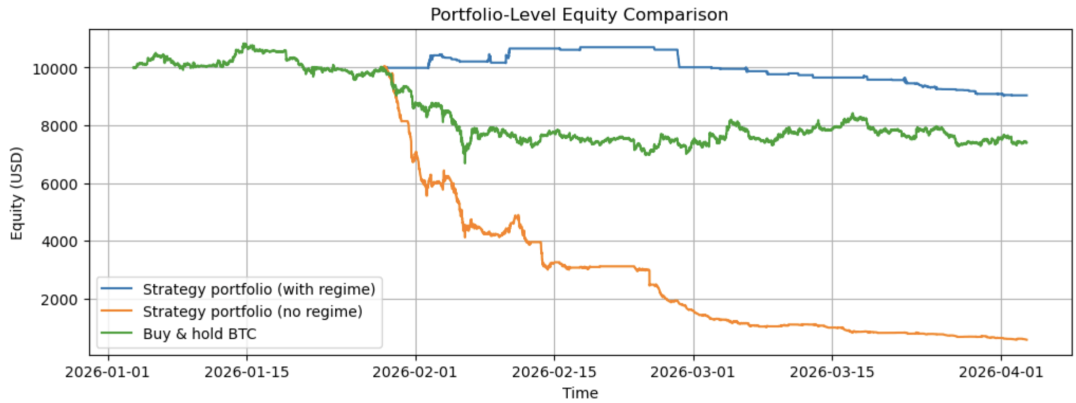
4.4 Impact of Regime Filtering

The regime filter plays a critical role in improving strategy robustness. By restricting trading to periods where the spread satisfies stationarity and mean-reversion conditions, the filter prevents capital deployment during structurally unstable regimes.

Empirically, the presence of skipped windows should not be interpreted as missed opportunities but rather as successful risk management. The model effectively identifies periods of non-stationarity and remains inactive, thereby avoiding potential losses associated with persistent spread divergence.

4.5 Equity Curve Analysis

A detailed comparison of equity curves for the BTC–XLM pair provides further insight into the effectiveness of the regime-based approach. The regime-filtered strategy significantly outperforms both the non-regime version and a passive buy-and-hold BTC benchmark.



Specifically, the regime-based strategy demonstrates superior capital preservation during drawdowns and more stable growth over time. In contrast, the non-regime strategy suffers from substantial losses during non-stationary periods, while the buy-and-hold benchmark exhibits higher volatility and deeper drawdowns.

This comparison confirms that the integration of regime detection materially enhances both risk-adjusted returns and downside protection.

5 Discussion

5.1 Cross-Asset Performance Variance

A primary observation was the significant variance in performance across the selected universe. While pairs like BTC–ETH exhibit the highest levels of long-term cointegration, their spreads were frequently compressed. With a fixed execution cost of 5 basis points (bps), the profit potential of the mean-reversion signal often failed to exceed the combined drag of fees and slippage.

Conversely, mid-cap assets such as AVAX and LINK provided the necessary idiosyncratic volatility. In these pairs, the spread exhibited wider excursions from the mean, allowing the strategy to capture alpha that was "thick" enough to survive transaction costs. This highlights that for a pairs trading strategy, statistical cointegration is a necessary but insufficient condition; there must also be sufficient spread volatility to justify the trade.

5.2 Execution Realism and Model Integrity

A pivotal finding is the strategy's resilience to the one-bar execution lag. In many backtests, alpha is a "mirage" caused by assuming instantaneous execution (lookahead bias). By successfully generating returns despite this lag, we confirm that our signals are capturing persistent structural imbalances rather than fleeting HFT-level micro-inefficiencies. This distinguishes the model as a viable strategy for real-world execution rather than a theoretical abstraction.

5.3 The Skipped Windows

The observed periods where the model remained flat should be interpreted as a success of the Regime Filter. Cryptocurrency markets are prone to "momentum blow-offs" where historical correlations collapse. Our model treats cointegration as a transient state; by identifying when the market shifted into a non-stationary, trending regime, the filter protected the portfolio from the "gap risk" common in naive mean-reversion strategies.

6 Conclusion

6.1 Key Strategic Takeaways

This project validates that a regime-conditional framework is superior to a static pairs trading approach. The key takeaways are:

- **The Regime Indicator as a Risk Shield** —– The ability to distinguish between profitable "mean-reverting noise" and dangerous "structural divergence" is the model's primary source of value.
- **BTC as the Market Anchor** —– Bitcoin remains the most robust basis for market-neutral pairing, acting as the stabilizer for the idiosyncratic moves of altcoins.
- **Deliberate Low Turnover** —– Our prioritization of high-conviction regimes over trade frequency successfully minimized fee erosion, proving that in crypto trading, quality of signal far outweighs quantity.

6.2 Limitations and Future Improvements

To evolve the model for institutional-grade deployment, we identify three key areas for improvement:

- **Funding Rate Inclusion** —– Integrating Perpetual Futures funding rates into the backtest to account for the "cost of carry," which can significantly impact the net Sharpe ratio.

- **Dynamic Sizing (Kelly Criterion)** —– Transitioning from static position sizes to a volatility-adjusted Kelly allocation to optimize capital efficiency during high-probability windows.
- **Expanded Universe** —– Testing for ”lead-lag” relationships beyond the current Binance universe to capture cross-exchange arbitrage opportunities.

References

[1] J. Suchato, S. Wiryadi, D. Chen, A. Zhao, and M. Yue, "An application of the ornstein-uhlenbeck process to pairs trading," *arXiv preprint arXiv:2412.12458*, 2024.

Github Repository

Visit the repository to access the source code, track ongoing development, report issues, and stay up to date with the latest changes.

`https://github.com/weisiang1218/nusiqf-crypto`

Disclaimer