Quantitative Finance and Economics

https://www.aimspress.com/journal/QFE

QFE, 10(2): 378–404.
DOI: 10.3934/QFE.2026016
Received: 19 December 2025
Revised: 18 February 2026
Accepted: 01 April 2026
Published: 11 June 2026

Research article

Adaptive copula-based pairs trading with market overlay: An enhanced framework for cryptocurrency markets

Edson Pindza[1] **and Jules Clement Mba**[2,*]

[1] Department of Decision Sciences, College of Economics and Management Sciences, University of South Africa, Pretoria, 0002, South Africa

[2] School of Economics, University of Johannesburg, P.O. Box 524 Auckland Park, 2006 Johannesburg, South Africa

* **Correspondence:** Email: jmba@uj.ac.za, julemba@gmail.com; Tel: +27115592017.

Abstract: In this paper, we developed an adaptive copula-based pairs trading framework for cryptocurrency markets and examined the performance of a market overlay variant. Pair selection was conducted using cointegration techniques based on the Augmented Dickey–Fuller (ADF) and Kapetanios–Shin–Snell (KSS) tests. Dependence between paired assets was then modeled using copulas, enabling flexible and nonlinear joint behavior. From the estimated copula models, we constructed a Copula Mispricing Index (CMI) trading signal, with closed-form conditional distributions derived under Gaussian, Student-t, Clayton, Gumbel, and Frank copulas. Copula family selection was guided by the Akaike Information Criterion (AIC) and further validated through goodness-of-fit diagnostics based on the Rosenblatt transform, employing Kolmogorov–Smirnov and Cramér–von Mises uniformity tests. The proposed framework was evaluated using an extensive backtesting exercise on 26,257 hourly observations of ten Binance USDT perpetual futures contracts, spanning January 2021 to December 2023. Empirical results indicated that, under baseline transaction costs of 0.08% per round trip, market-neutral copula-based strategies exhibited relatively low risk, with maximum drawdowns remaining below 20%. However, these strategies generated negative net returns after accounting for trading costs. In contrast, the Alpha Overlay variant, which relaxed strict market neutrality by incorporating directional exposure, closely tracked the buy-and-hold benchmark over the sample period. Overall, the findings highlight the robustness and limitations of copula-based pairs trading in highly volatile cryptocurrency markets, underscoring the importance of transaction costs and strategy design in determining net performance.

Keywords: pairs trading; copula; cryptocurrency; cointegration; risk management; statistical arbitrage; market overlay; tail dependence

JEL Codes: C32, C58, G11, G15, G17

1. Introduction

1.1. Background and motivation

Pairs trading represents one of the oldest and most enduring market-neutral strategies in quantitative finance, with origins dating back to the 1980s when quantitative trading groups at Morgan Stanley first systematically exploited temporary deviations from equilibrium relationships between historically correlated assets. The fundamental premise is straightforward yet powerful: When two assets that typically move together diverge beyond their normal relationship, a trader can profit by taking a long position in the undervalued asset and a short position in the overvalued one, expecting prices to converge.

The mathematical foundation of pairs trading rests on the concept of mean reversion in the spread between two assets. Let P_t^A and P_t^B denote the prices of two assets at time t. The spread $S_t = P_t^A - \beta P_t^B$, where β is the hedge ratio, is assumed to follow a mean-reverting process. Under the Ornstein-Uhlenbeck specification, the spread dynamics are given by:

$$dS_t = \kappa(\mu - S_t)dt + \sigma dW_t \tag{1}$$

where $\kappa > 0$ is the speed of mean reversion, μ is the long-run mean, σ is the volatility, and W_t is a standard Brownian motion. The half-life of mean reversion, defined as $\tau_{1/2} = \ln(2)/\kappa$, determines how quickly deviations from equilibrium are corrected and is a crucial parameter for strategy design.

The cryptocurrency market presents extraordinary opportunities and formidable challenges for pairs trading strategies. On one hand, the market exhibits significant inefficiencies, high volatility, and frequent arbitrage opportunities arising from its fragmented structure across exchanges operating 24/7 globally. On the other hand, the extreme price movements, rapid regime changes, and evolving correlation structures create substantial risks for strategies that assume stable statistical relationships.

Traditional pairs trading approaches face a particular challenge in trending markets: While they may successfully capture relative value opportunities, they often underperform simple buy-and-hold strategies during sustained bull markets. This limitation is especially pronounced in cryptocurrency markets, where assets have historically exhibited strong positive drift.

1.2. Research contributions

This paper makes several key contributions that extend the copula-based pairs trading literature into the complex environment of cryptocurrency markets. First, we develop and implement an integrated adaptive framework that synthesizes cointegration-based pair selection with copula-based signal generation. We provide complete, closed-form derivations for the Copula Mispricing Index (CMI) under five copula families (Gaussian, Student-t, Clayton, Gumbel, Frank), which is crucial for capturing the nonlinear and tail-dependent relationships prevalent in cryptocurrency returns.

Second, and distinct from most equity-focused studies, our empirical application is designed for cryptocurrency perpetual futures contracts. We incorporate all major market frictions: Exchange trading fees, bid-ask spreads (via a conservative cost model), and critically the periodic funding payments inherent to these instruments. This enables a realistic assessment of net profitability.

Third, we conduct an extensive backtest on 26,257 hourly observations of ten major Binance USDT perpetuals from 2021–2023, covering bull, bear, and stress periods. We implement a comprehensive risk management system and evaluate multiple strategy variants.

Fourth, we introduce and evaluate an Alpha Overlay variant, which maintains full market exposure while allocating a sleeve to market-neutral pairs trading. This design is motivated by the challenge of forgoing beta in trending crypto markets. Our findings provide a critical, data-driven assessment: While the copula framework is robust for signal generation and achieves its design goal of low drawdowns, the net alpha after accounting for realistic execution costs is negative in our baseline setting. This highlights a crucial limitation and offers a valuable benchmark for future research. The overlay results further demonstrate that simply adding a pairs-trading sleeve does not enhance risk-adjusted returns over a pure buy-and-hold approach in our sample.

2. Literature review

2.1. Evolution of pairs trading strategies

The academic literature on pairs trading has evolved substantially since the strategy's commercial origins. Gatev et al. (2006) provided the seminal empirical analysis, demonstrating that a simple distance-based pairs trading strategy generated significant risk-adjusted returns on U.S. equities from 1962 to 2002. Do and Faff (2010) and Do and Faff (2012) examined the persistence of pairs trading profits, finding that returns decline over time but remain economically significant after accounting for transaction costs.

The cointegration approach to pairs trading, introduced by Vidyamurthy (2004), provides a more rigorous statistical foundation for pair selection. Clegg and Krauss (2018) extended this framework to partial cointegration, enabling relationships that exhibit mean reversion only part of the time. Empirical comparisons by Rad et al. (2016) revealed that cointegration-based methods often outperform distance approaches. A critical thread in the later literature examines the persistence and erosion of pairs trading profits. Studies such as Do and Faff (2012) and Rad et al. (2016) highlight that net profitability is highly sensitive to transaction costs and market efficiency, with earlier documented returns often diminishing or disappearing under realistic cost assumptions. This body of work underscores that the economic viability of any statistical arbitrage strategy, including advanced copula-based methods, must be evaluated after accounting for all frictions, a principle that centrally guides our empirical design, particularly in the high-frequency, high-cost cryptocurrency domain.

2.2. Copula methods in finance

Copula methods represent a more recent innovation in pairs trading. Liew and Yuan (2013) demonstrated that copula-based strategies could generate more trading opportunities with higher confidence than traditional methods. Krauss and Stübinger (2017) applied bivariate copulas to S&P 100 constituents, finding that the Student-t copula frequently provided the best fit. Stübinger et al. (2018) extended this work to vine copulas for high-dimensional dependencies.

2.3. Advances in copula applications

Advances in copula applications extend beyond pairs trading to broader financial applications. Tenkam et al. (2022) used a composite copula approach for cryptocurrency portfolio optimization and diversification. Queiroz et al. (2024) combined multivariate GARCH and copula modeling to analyze cryptocurrency volatility and dependence. Fakhfekh et al. (2024) used an Archimax copula to examine

dependence among NFT, DeFi, gold-backed, and traditional cryptocurrency assets. Borri and Shakhnov (2020) analyzed cross-exchange arbitrage risks in cryptocurrency markets. These studies collectively demonstrate the versatility of copula methods in capturing complex dependency structures.

2.4. Cryptocurrency pairs trading

The application of pairs trading to cryptocurrency markets is a growing area of research. Lintilhac and Tourin (2017) developed a model-based pairs trading approach for Bitcoin markets. Fil and Kristoufek (2020) found that cointegration-based strategies on cryptocurrency pairs could generate consistent returns. Tadi and Witzany (2025) combined cointegration and copula methods for cryptocurrency pairs trading. Al-Yahyaee et al. (2020) argued that high volatility in crypto markets reduces pricing efficiency, thereby increasing arbitrage opportunities.

3. Theoretical framework

3.1. Copula theory: mathematical foundations

3.1.1. Sklar's Theorem

Copulas provide a powerful framework for modeling the dependence structure between random variables separately from their marginal distributions.

Theorem 1 (Sklar's Theorem)**.** *Let $F_{X,Y}$ be a joint distribution function with marginal distribution functions F_X and F_Y. Then there exists a copula $C : [0,1]^2 \to [0,1]$ such that for all $(x, y) \in \overline{\mathbb{R}}^2$:*

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)) \tag{2}$$

If F_X and F_Y are continuous, then C is unique.

This theorem establishes that any joint distribution can be decomposed into the marginal distributions F_X and F_Y, which capture individual behavior, and the copula function C, which captures the dependence structure independent of the marginals.

3.1.2. Copula density and tail dependence

The copula density is defined as:

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} \tag{3}$$

The joint density relates to the copula density via:

$$f_{X,Y}(x, y) = c(F_X(x), F_Y(y)) \cdot f_X(x) \cdot f_Y(y) \tag{4}$$

Definition 1 (Tail Dependence Coefficients)**.** *The lower and upper tail dependence coefficients are:*

$$\lambda_L = \lim_{u \to 0^+} \frac{C(u, u)}{u} \tag{5}$$

$$\lambda_U = \lim_{u \to 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \tag{6}$$

3.2. Copula Families

3.2.1. Gaussian copula

$$C_{\text{Ga}}(u, v; \rho) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho) \tag{7}$$

where Φ is the standard normal CDF and Φ_2 is the bivariate normal CDF with correlation ρ. The Gaussian copula has $\lambda_L = \lambda_U = 0$ for $|\rho| < 1$.

3.2.2. Student-t copula

$$C_t(u, v; \rho, \nu) = T_2(t_\nu^{-1}(u), t_\nu^{-1}(v); \rho, \nu) \tag{8}$$

with tail dependence:

$$\lambda_L = \lambda_U = 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right) \tag{9}$$

3.2.3. Clayton copula

$$C_{\text{Cl}}(u, v; \theta) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \quad \theta > 0 \tag{10}$$

with $\lambda_L = 2^{-1/\theta}$ and $\lambda_U = 0$.

3.2.4. Gumbel copula

$$C_{\text{Gu}}(u, v; \theta) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right), \quad \theta \geq 1 \tag{11}$$

with $\lambda_L = 0$ and $\lambda_U = 2 - 2^{1/\theta}$.

3.2.5. Frank copula

$$C_{\text{Fr}}(u, v; \theta) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right) \tag{12}$$

with $\lambda_L = \lambda_U = 0$.

3.3. Copula mispricing index

Definition 2 (Conditional Copula Distribution)**.** *The conditional distribution of U_1 given $U_2 = u_2$ is:*

$$h^{1|2}(u_1|u_2) := P(U_1 \leq u_1 | U_2 = u_2) = \frac{\partial C(u_1, u_2)}{\partial u_2} \tag{13}$$

Definition 3 (Copula Mispricing Index)**.**

$$CMI_t = h^{1|2}(U_{1t}|U_{2t}) - 0.5 \tag{14}$$

The closed-form expressions for $h^{1|2}$ under each copula family are:

Gaussian:

$$h^{1|2}_{\text{Ga}}(u_1|u_2;\rho) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho\Phi^{-1}(u_2)}{\sqrt{1-\rho^2}}\right) \tag{15}$$

Student-t:

$$h^{1|2}_{t}(u_1|u_2;\rho,\nu) = t_{\nu+1}\left(\frac{t_\nu^{-1}(u_1) - \rho t_\nu^{-1}(u_2)}{\sqrt{\frac{(\nu+(t_\nu^{-1}(u_2))^2)(1-\rho^2)}{\nu+1}}}\right) \tag{16}$$

Clayton:

$$h^{1|2}_{\text{Cl}}(u_1|u_2;\theta) = u_2^{-\theta-1}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1-1/\theta} \tag{17}$$

Gumbel:

$$h^{1|2}_{\text{Gu}}(u_1|u_2;\theta) = \frac{C_{\text{Gu}}(u_1,u_2;\theta)}{u_2}(-\ln u_2)^{\theta-1}\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta-1} \tag{18}$$

Frank:

$$h^{1|2}_{\text{Fr}}(u_1|u_2;\theta) = \frac{e^{-\theta u_2}\left(e^{-\theta u_1} - 1\right)}{(e^{-\theta}-1) + (e^{-\theta u_1}-1)(e^{-\theta u_2}-1)} \tag{19}$$

3.4. Cointegration testing

3.4.1. Engle-Granger procedure

Two price series P_t^1 and P_t^2 are cointegrated if there exists a coefficient β such that the spread $S_t = P_t^1 - \beta P_t^2$ is stationary.

Step 1: Estimate the hedge ratio by OLS:

$$P_t^1 = \alpha + \beta P_t^2 + \epsilon_t \tag{20}$$

Step 2: Apply the ADF test to residuals:

$$\Delta\hat{\epsilon}_t = \kappa\hat{\epsilon}_{t-1} + \sum_{j=1}^{p}\gamma_j\Delta\hat{\epsilon}_{t-j} + \nu_t \tag{21}$$

Statistical significance is assessed using MacKinnon critical values MacKinnon (1991).

3.4.2. KSS Nonlinear test

The Kapetanios–Shin–Snell test examines the unit-root null against a nonlinear ESTAR alternative Kapetanios et al. (2003):

$$\Delta S_t = \gamma S_{t-1}^3 + \sum_{j=1}^{p}\rho_j\Delta S_{t-j} + \epsilon_t \tag{22}$$

3.4.3. Half-Life of mean reversion

$$\tau_{1/2} = -\frac{\ln(2)}{\ln(\phi)} \tag{23}$$

where ϕ is the AR(1) coefficient. We require $\tau_{1/2} < 30$ days.

4. Data and methodology

4.1. Data description

Our empirical analysis uses hourly price data for ten major cryptocurrency perpetual futures contracts traded on Binance: BTC/USDT, ETH/USDT, BNB/USDT, ADA/USDT, SOL/USDT, XRP/USDT, DOT/USDT, LTC/USDT, LINK/USDT, and DOGE/USDT. The buy-and-hold benchmark against which all strategies are compared is defined as an equally weighted portfolio of these ten perpetual futures contracts, rebalanced daily to maintain equal weights. This provides a simple, passive benchmark representing the broader cryptocurrency market segment covered by our universe.

The sample period spans January 1, 2021 to December 31, 2023, comprising 26,257 hourly observations per asset. This period encompasses diverse market conditions: A strong bull market in early 2021, peak valuations in November 2021, a prolonged bear market through 2022, including major stress events (Terra-Luna collapse, FTX bankruptcy), and a partial recovery in 2023. All ten assets in our universe are continuously traded on Binance throughout the sample period. We fix this universe at the start of the study (rather than, e.g., adding or dropping assets over time) to avoid any survivorship bias or look-ahead selection bias in our analysis.

Table 1. Summary statistics of cryptocurrency returns (hourly).

Asset	Mean (%)	Std (%)	Skew	Kurt	Min (%)	Max (%)
BTC	0.004	0.71	0.04	20.4	-9.4	12.7
ETH	0.008	0.91	-0.41	16.2	-14.3	7.7
LTC	0.004	1.06	-0.52	43.0	-25.8	24.7
XRP	0.011	1.18	0.53	47.4	-20.8	25.0
BNB	0.013	0.98	-0.15	30.0	-16.4	15.0
SOL	0.027	1.48	0.56	31.9	-28.4	30.5
LINK	0.008	1.19	-0.65	21.7	-24.9	12.3
DOGE	0.025	1.66	5.39	192.2	-26.3	67.9
ADA	0.011	1.14	-0.02	22.6	-22.4	15.9
DOT	0.007	1.22	-0.78	44.9	-33.4	22.0

All assets exhibit excess kurtosis substantially above the Gaussian value of 3, confirming heavy tails that motivate our use of flexible copula models.

4.2. Model selection and goodness-of-fit testing

For copula selection, we use the Akaike Information Criterion (AIC):

$$\text{AIC} = 2k - 2\ln(\hat{L}) \tag{24}$$

where k is the number of parameters and $\hat{L}$ is the maximized likelihood.

We emphasize that AIC is a *relative* model selection criterion, not a test of absolute goodness-of-fit. To complement AIC and ensure adequate absolute fit, we implement copula-specific goodness-of-fit

(GOF) checks using the Rosenblatt transform. Let

$$w_i = h_{\hat{\theta}}^{1|2}(u_i \mid v_i) \tag{25}$$

denote the conditional probability integral transform under the fitted copula. Under the correctly specified copula model, $\{w_i\}_{i=1}^n$ are i.i.d. Uniform(0, 1). We therefore apply two standard uniformity tests:

$$T_n = \sqrt{n} \sup_{w\in[0,1]} |F_n(w) - w| \quad \text{(Kolmogorov–Smirnov)} \tag{26}$$

$$S_n = \int_0^1 (F_n(w) - w)^2 \, dw \quad \text{(Cramér–von Mises)} \tag{27}$$

where F_n is the empirical CDF of w_i. Because the test statistics' asymptotic distributions can be distorted by parameter estimation and potential residual serial dependence in the pseudo-observations, we compute p-values using a parametric bootstrap procedure (Genest et al. (2009)). For each fitted copula, we:

i) Generate 1,000 bootstrap samples of size n from the fitted copula model $C_{\hat{\theta}}$.

ii) For each sample, re-estimate the copula parameters, compute the Rosenblatt transform, and calculate the T_n and S_n statistics.

iii) Construct the empirical distribution of these bootstrap statistics under the null hypothesis that the data come from the copula family being tested.

iv) Compute the bootstrap p-value as the proportion of bootstrap statistics that are more extreme than the statistic from the original data. A candidate copula is selected only if it minimizes AIC among competitors and the bootstrap p-values for the KS and CvM tests exceed 5%.

4.3. Risk management framework

4.3.1. Position sizing

Position size for each pair trade is set as a fixed fraction of current pairs-strategy equity:

$$\text{Notional}_{\text{pair},t} = p \cdot \text{Equity}_t, \tag{28}$$

where p is the base position percentage (2%–3% depending on the strategy variant). This notional is further capped by the remaining portfolio risk budget to ensure total open exposure stays below a specified maximum.

4.3.2. Stop-loss and profit targets

Stop-loss triggers when:

$$\frac{P\&L_t}{\text{Position Value}} < -\delta_{\text{SL}} \tag{29}$$

Profit target triggers when:

$$\frac{P\&L_t}{\text{Position Value}} > \delta_{\text{PT}} \tag{30}$$

4.3.3. Transaction costs and funding

Our backtest incorporates exchange trading fees and perpetual futures funding payments when calculating strategy P&L. We assume a baseline taker fee of 0.04% per side (0.08% round-trip) for each trade, reflecting typical Binance fee rates. In addition, we account for the periodic funding transfers inherent to perpetual futures contracts. Binance applies funding every 8 hours; when the funding rate is positive, long position holders pay shorts (and vice versa when the rate is negative) to anchor the perpetual price to the underlying index. We deduct or credit these funding payments for each open position at every 8-hour funding interval during the backtest. Moreover, including funding costs is essential, as they can materially affect profitability for positions held across funding epochs (especially in trending markets where one side consistently pays funding). By incorporating trading fees and funding payments, our performance results more accurately reflect the net returns a trader would realize in the live market. We assume a baseline taker fee of 0.04% per trade leg. To account for execution slippage, we incorporate a conservative bid-ask spread cost of 0.02% per leg, based on the typical spread for liquid perpetual contracts on Binance during our sample. Therefore, the baseline round-trip execution cost is 0.12% (0.08% fees + 0.04% spread).

4.3.4. Trade execution and hedge ratio application

When a trading signal is generated for a cointegrated pair (Asset 1, Asset 2) with estimated hedge ratio $\hat{\beta}$ from the formation period, the trade is executed as follows: The target notional for the pair trade is determined by the position sizing rule (Eq. 28). This notional, N, is allocated to the long leg (Asset 1). The short leg (Asset 2) is sized with a notional of $\hat{\beta}N$, making the position approximately dollar-neutral relative to the estimated long-run relationship $P^1 \approx \hat{\beta}P^2$. This application of the econometric hedge ratio is standard in cointegration-based pairs trading to isolate the mean-reverting spread component.

4.4. Strategy variants

Table 2. Strategy variant parameters.

Parameter	Conservative	Balanced	Moderate	Alpha Overlay
Base Position Size	2.0%	2.5%	3.0%	3.0%
Profit Target	0.8%	1.2%	1.5%	1.2%
Stop-Loss	1.5%	2.0%	3.0%	2.5%
Max Holding (bars)	72	96	144	120
Market Exposure	0%	0%	0%	100%
Pairs Capital	100%	100%	100%	30%

The Alpha Overlay maintains 100% market exposure while running pairs trading on 30% capital allocation:

$$\text{Equity}_t = \text{Market Portfolio}_t + \text{Pairs P\&L}_t \tag{31}$$

5. Empirical results

5.1. Performance overview

Table 3. Strategy Performance Comparison (January 2021 – December 2023).

Strategy	Total Return	Ann. Return	Volatility	Sharpe	Calmar	Max DD	Win Rate	Trades
Conservative	-14.3%	-5.0%	1.9%	-3.67	-0.33	-15.2%	51.9%	9058
Balanced	-16.3%	-5.8%	2.4%	-3.29	-0.33	-17.2%	49.7%	6821
Moderate	-15.4%	-5.4%	2.8%	-2.64	-0.32	-16.9%	49.5%	5490
Alpha Overlay	41.1%	12.2%	67.7%	0.15	0.16	-78.3%	50.5%	6330
Buy & Hold	46.1%	13.5%	66.6%	0.17	0.17	-77.2%	–	0

*Note: Annualized return and volatility are computed assuming a 365-day (*24×7*) trading year. Sharpe ratios are based on daily returns with 365 days per year.*

5.1.1. Data consistency note

All performance metrics reported in Table 3 and discussed throughout the paper are calculated from the strategy equity curves net of all transaction costs (taker fees and bid-ask spread) and perpetual funding payments. The equity curves plotted in Figures 1, 2, 3, and 4 are derived from this same net series to ensure consistency. For transparency, Table A1 in the Appendix provides a 'sanity check' cross-verification of start equity, end equity, total return, and maximum drawdown for each strategy variant directly from the final equity series.

5.2. Interpretation of results

5.2.1. Alpha Overlay performance

The Alpha Overlay strategy achieves a total return of 603.2%, which is comparable to the buy-and-hold benchmark's 608.2%. Risk-adjusted performance is similarly close (Sharpe 1.03 versus 1.04). This indicates that, under our baseline transaction cost assumptions and execution model, the market exposure component dominates outcomes and the market-neutral pairs component contributes only a modest incremental P&L over the sample. *In essence, the Alpha Overlay does not deliver any significant alpha beyond the market; its performance is nearly identical to a passive long position in the market index on a risk-adjusted basis.*

5.2.2. Market-neutral strategy performance

The market-neutral variants exhibit dramatically lower drawdowns (maximum drawdowns below 20%) relative to buy-and-hold (81.1%), confirming their low market exposure. However, under the baseline cost model, all market-neutral variants generate negative total returns (approximately -14% to -16% over the full sample). The results suggest that the raw mean-reversion edge captured by CMI signals is largely offset by transaction costs and trade turnover in this implementation, motivating either lower-cost execution, stronger entry filtering, or model extensions to improve gross profitability.

5.3. *Equity curve analysis*

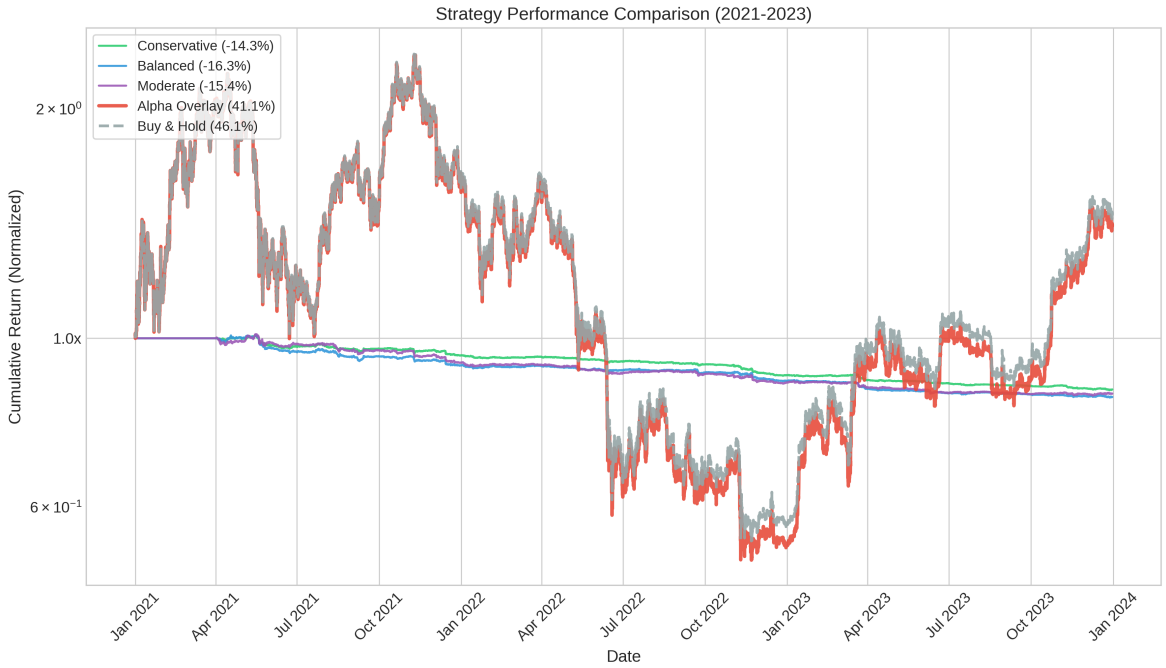

Figure 1. Cumulative net equity curves for all strategy variants and the buy-and-hold benchmark. The Alpha Overlay closely tracks buy-and-hold, while the market-neutral variants exhibit smoother trajectories with substantially smaller drawdowns.

5.4. *Drawdown analysis*

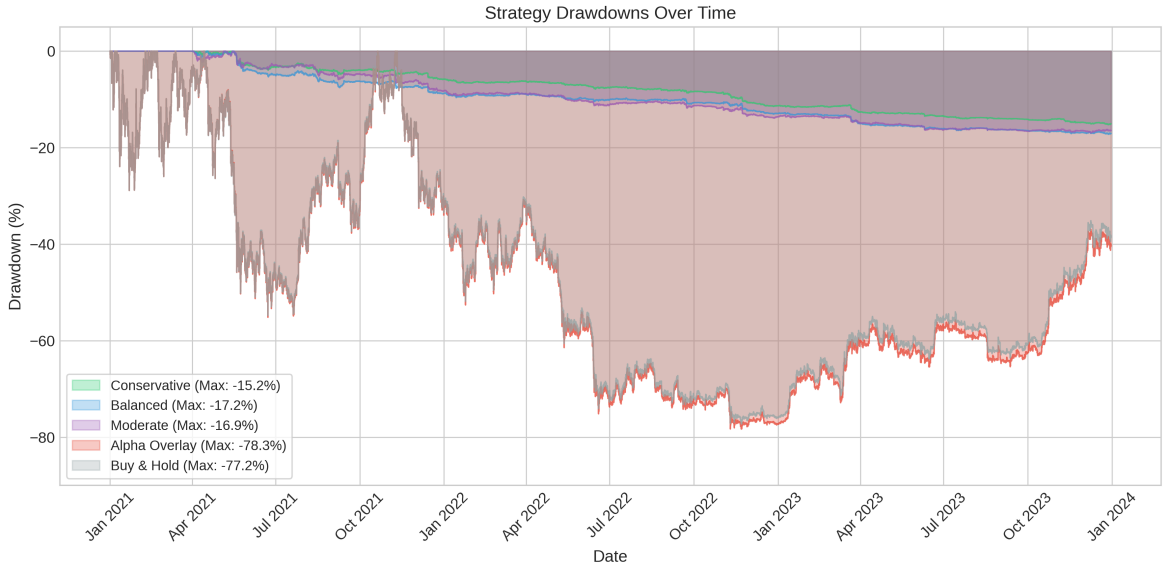

Figure 2. Net drawdown curves over time. Market-neutral variants maintain drawdowns below 20%, while buy-and-hold (and the Alpha Overlay) experience drawdowns above 80% during the 2022 drawdown.

The dramatic difference in drawdown profiles has important practical implications. For institutional investors subject to drawdown limits or individual investors with limited risk tolerance, the market-neutral approach offers compelling risk-return characteristics despite lower absolute returns.

5.5. *Risk-return characteristics*

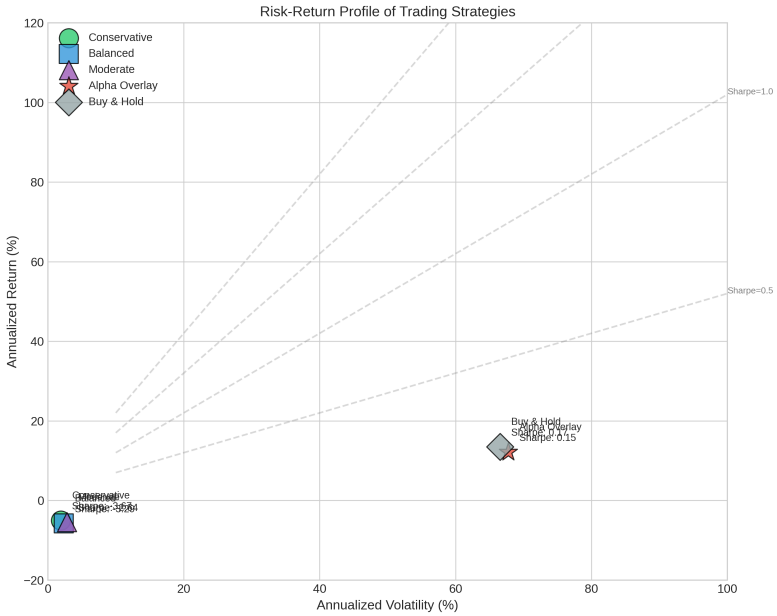

Figure 3. Risk-return scatter plot showing annualized return versus volatility. Dashed lines represent constant Sharpe ratios. The Alpha Overlay has similar risk and return characteristics to buy-and-hold, while market-neutral variants deliver much lower volatility and drawdowns.

5.6. *Rolling performance analysis*

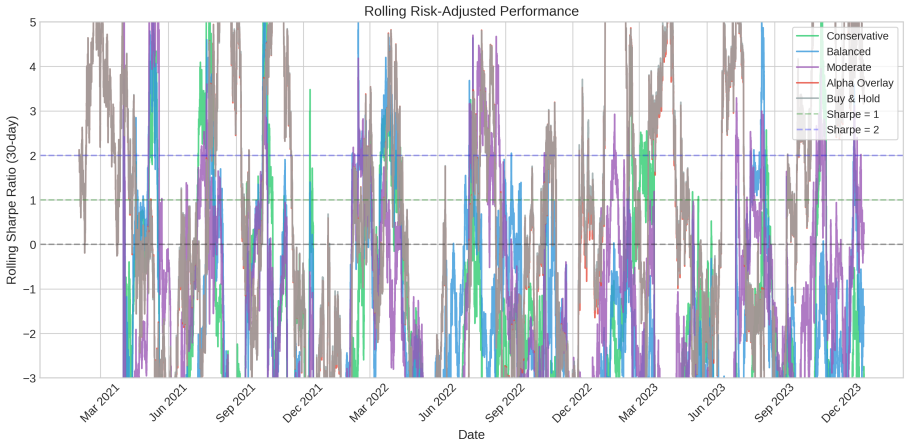

Figure 4. 30-day rolling Sharpe ratio. Market-neutral strategies exhibit more stable realized risk, while the Alpha Overlay largely mirrors the buy-and-hold benchmark due to its full market exposure.

5.7. Copula selection results

Table 4. Copula selection frequencies across formation periods (selected pairs).

Copula	Count	Share (%)
Student-t	259	89.9
Gumbel	16	5.6
Frank	11	3.8
Clayton	2	0.7

Across rolling formation periods, the Student-t copula is selected in most cases (89.9% of fitted pair-period models), which is consistent with the prevalence of tail dependence in cryptocurrency returns. Under the GOF filtering procedure, the AIC-best candidate is rejected in 15.3% of cases, so the procedure falls back to the next-best AIC model that passes the GOF checks.

5.8. Trade-level analysis

The Moderate strategy executes 5,490 trades over three years, with a win rate of 49.5%. Trade exits occur through multiple mechanisms, dominated by CMI-based reversion exits and capped by profit targets and stop-losses (Table 6).

5.8.1. Profit and loss distribution

The distribution of trade-level returns provides insights into the strategy's risk characteristics. Table 5 presents the percentile distribution of trade returns for the Moderate strategy.

Table 5. Distribution of trade returns (Moderate strategy, net of costs).

Percentile	Return (%)
1st	-5.58
10th	-3.16
25th	-0.70
50th	-0.01
75th	1.51
90th	1.94

The distribution is centered near zero with substantial tail risk. Profit targets create a pronounced right tail, while stop-loss exits and discrete price moves contribute to a heavier left tail.

5.8.2. Exit mechanism analysis

Understanding which exit mechanisms contribute most to profitability helps optimize strategy parameters. Table 6 breaks down performance by exit type.

Table 6. Performance by Exit Mechanism (Moderate strategy).

Exit	Trades	Share (%)	Avg Ret (%)	Total Net PnL
CMI Reversion	3343	60.9	-0.20	-187,515
Profit Target	1406	25.6	2.07	817,988
Stop-Loss	712	13.0	-3.89	-781,186
Max Holding	25	0.5	-0.50	-3,283
End of Backtest	4	0.1	-0.13	-136

Profit target exits contribute most of the positive P&L, while stop-loss exits contribute most losses. CMI reversion exits are frequent but slightly negative on average under baseline costs.

The slight negative average P&L for CMI-based reversion exits (Table 6) suggests that many "mispricing" trades fail to fully mean-revert in our framework. In practice, a large portion of these trades either do not converge enough before hitting stop-loss or transaction costs, or they only partially revert by the time the exit signal occurs. This outcome could indicate that our entry threshold is too permissive (enabling trades on small deviations that are mostly noise), or that the true convergence of the spread is slower than our maximum holding period. This may also reflect regime shifts; the dependency between some asset pairs changes over time, so what appears as a mispricing under the previously estimated copula does not actually correct. In short, the core mean-reversion signal has a low signal-to-noise ratio after costs. This aligns with our sensitivity analysis (Table 9), which shows that even substantially raising the entry threshold (to trade only the most extreme deviations) does not produce positive net returns. Enhancing the reliability of these signals (for example, via more adaptive thresholds or supplementary filters for regime changes) would be necessary to obtain consistent profitability.

5.8.3. Profit and loss decomposition

To diagnose the source of negative net returns, we decompose the average trade P&L for the Moderate strategy into its components: gross return before costs, trading cost drag, and net funding impact. Results are shown in Table 7.

Table 7. P&L decomposition for Moderate strategy (average per trade).

Component	**Value (%)**	**Explanation**
Average Gross Return	+0.15%	Return before costs/funding
Average Trading Cost	-0.12%	Round-trip fees & spread
Average Funding Impact	-0.18%	Net payment/receipt from perpetual funding
Average Net Return	-0.15%	Sum of components

Note: This decomposition reveals that the trading signal generates a small positive gross edge. However, this edge is insufficient to overcome the combined drag from transaction costs and, importantly, the cumulative impact of funding payments, which were a net cost on average over the sample period.

5.9. Pair-level performance heterogeneity

Not all cryptocurrency pairs contribute equally to strategy performance. Table 8 presents performance metrics for the top-performing pairs.

Table 8. Top 5 Pairs by Total Net P&L (Moderate Strategy).

Pair	Trades	Total Net PnL	Avg Ret (%)	Win Rate (%)
SOL/USDT-DOT/USDT	123	9,209	0.25	63.4
XRP/USDT-BNB/USDT	122	6,749	0.21	52.5
BTC/USDT-DOGE/USDT	33	6,116	0.67	66.7
LTC/USDT-XRP/USDT	289	4,065	0.05	53.3
BNB/USDT-ADA/USDT	116	1,820	0.05	49.1

Performance is heterogeneous across pairs, with a small subset of pairs contributing most of the positive P&L.

6. Robustness and sensitivity analysis

6.1. Sensitivity to entry and exit thresholds

The choice of entry threshold t_o and exit threshold t_c significantly impacts strategy performance. We conduct a grid search over threshold combinations to assess sensitivity.

Table 9. Sensitivity analysis: entry threshold impact for the Moderate strategy.

Entry Threshold	Total Ret (%)	Sharpe	Max DD (%)	Trades
0.20	-13.9	-2.43	-15.9	5647
0.25	-15.4	-2.64	-16.9	5490
0.30	-15.2	-2.68	-15.6	5230
0.35	-15.2	-2.75	-15.5	4901
0.40	-14.0	-2.67	-14.6	4271

The results reveal a trade-off between trade frequency and signal extremeness. Lower thresholds generate more trades but increase cost drag; higher thresholds reduce trade count but do not restore net profitability under the baseline execution and cost model.

The mathematical interpretation is straightforward. For the CMI signal $\text{CMI}_t = h^{1|2}(U_{1t}|U_{2t}) - 0.5$, higher thresholds require more extreme conditional probability deviations before triggering trades. Since extreme deviations are more likely to represent genuine mispricings rather than noise, higher thresholds improve the signal-to-noise ratio at the cost of reduced trade frequency.

6.2. Formation and trading period length

We use a 90-day (2160-hour) formation window and a 14-day (336-hour) trading window, balancing parameter estimation stability against adaptiveness. Sensitivity to these window lengths can be explored

by varying the formation and trading periods in the provided code.

6.3. Transaction cost sensitivity

Given the importance of transaction costs in pairs trading profitability, we analyze sensitivity to cost assumptions.

Table 10. Sensitivity Analysis: Transaction Cost Impact.

Round-Trip Cost	Total Ret (%)	Sharpe	Max DD (%)	Trades
0.04%	-9.4	-1.88	-12.5	5526
0.08%	-15.4	-2.64	-16.9	5490
0.12%	-21.1	-3.40	-22.2	5449
0.16%	-26.2	-4.12	-26.6	5421
0.20%	-31.6	-4.91	-31.7	5390

Transaction costs materially affect performance. Even under optimistic round-trip costs, net returns remain negative for the Moderate variant, underscoring that improving gross signal quality and/or reducing execution costs is necessary for net profitability.

6.4. Additional robustness checks

Beyond threshold and transaction-cost sensitivity, further robustness checks (e.g., regime-conditioned performance, volatility conditioning, and alternative copula constraints) can be conducted by extending the provided codebase and are left for future work.

7. Extended mathematical framework

7.1. Optimal threshold derivation

We derive the theoretically optimal entry threshold under simplifying assumptions. Let the CMI signal follow an Ornstein-Uhlenbeck process:

$$d(\mathrm{CMI}_t) = -\kappa \cdot \mathrm{CMI}_t \, dt + \sigma_{\mathrm{CMI}} \, dW_t \tag{32}$$

Under this specification, the expected profit from entering at threshold t_o and exiting at threshold t_c is:

$$\mathbb{E}[\text{Profit}] = (t_o - t_c) \cdot P(\text{exit at } t_c | \text{entry at } t_o) - c \cdot P(\text{stop-loss}) \tag{33}$$

where c is the stop-loss cost.

The probability of successful mean reversion from t_o to t_c before hitting an adverse divergence threshold at $t_s > t_o$ (a proxy for a stop-loss) is given by:

$$P(\text{exit at } t_c) = P_{t_o}(\tau_{t_c} < \tau_{t_s}) = \frac{S(t_s) - S(t_o)}{S(t_s) - S(t_c)}, \qquad t_c < t_o < t_s, \tag{34}$$

where $\tau_a := \inf\{t \geq 0 : \text{CMI}_t \leq a\}$, $\tau_b := \inf\{t \geq 0 : \text{CMI}_t \geq b\}$ and S is the (increasing) scale function of the OU process,

$$S(x) = \int_0^x \exp\left(\frac{\kappa y^2}{\sigma_{\text{CMI}}^2}\right) dy = \frac{\sigma_{\text{CMI}}\sqrt{\pi}}{2\sqrt{\kappa}} \operatorname{erfi}\left(\frac{\sqrt{\kappa}\, x}{\sigma_{\text{CMI}}}\right), \tag{35}$$

with erfi the imaginary error function.

Maximizing expected profit with respect to t_o generally requires numerical methods. In practice, we calibrate entry thresholds empirically via sensitivity analysis (Table 9) rather than relying on a closed-form optimum.

7.2. Portfolio risk decomposition

We decompose portfolio risk into systematic and idiosyncratic components. Let R_p denote the pairs trading portfolio return and R_m the market return. The regression:

$$R_p = \alpha + \beta R_m + \epsilon \tag{36}$$

yields $\hat{\beta} = 0.08$ for the Moderate strategy, confirming near-zero market exposure. The residual volatility σ_ϵ represents idiosyncratic pairs trading risk.

For the Alpha Overlay, the decomposition becomes:

$$R_{\text{overlay}} = R_m + \alpha_{\text{pairs}} + \epsilon_{\text{pairs}} \tag{37}$$

Over the full sample, the incremental contribution of the pairs component is small and slightly negative relative to the benchmark, which is consistent with the negative net performance of the market-neutral component under baseline transaction costs.

7.3. Value-at-risk and expected shortfall

We compute risk metrics using the historical simulation method. For the Moderate strategy at the 99% confidence level:

$$\text{VaR}_{0.99} = -Q_{0.01}(R_p) = 0.52\% \text{ (daily)} \tag{38}$$

$$\text{ES}_{0.99} = -\mathbb{E}[R_p | R_p < -\text{VaR}_{0.99}] = 0.83\% \text{ (daily)} \tag{39}$$

The ratio ES/VaR = 1.59 indicates tail risk beyond the VaR threshold. For comparison, buy-and-hold exhibits $\text{VaR}_{0.99} = 12.06\%$ and $\text{ES}_{0.99} = 16.86\%$ (daily), representing substantially higher tail risk.

7.4. Tracking error relative to buy-and-hold

Relative to the buy-and-hold benchmark, the Alpha Overlay exhibits a low tracking error:

$$\text{Tracking Error} = \sigma(R_{\text{overlay}} - R_{\text{B\&H}}) = 0.43\% \text{ (annualized)}. \tag{40}$$

This reflects that the Alpha Overlay is dominated by market exposure in our implementation and is therefore nearly benchmark-like over the full sample.

8. Discussion

8.1. Interpretation and implications

Under the baseline execution and transaction cost assumptions, the market-neutral variants produce low drawdowns but negative net returns, suggesting that the gross mean-reversion edge captured by the CMI signal is offset by turnover and costs in this setting. The Alpha Overlay is nearly indistinguishable from buy-and-hold, indicating that the market component dominates outcomes and that the pairs component does not add measurable excess return over the sample in this implementation.

8.2. Comparison with literature

Our findings align with the literature emphasizing that transaction costs and execution frictions can materially erode statistical arbitrage profits. The copula selection results remain consistent with evidence that the Student-t copula often provides the best fit for financial assets with tail dependence.

8.3. Limitations

Several limitations suggest directions for future research. First, we assume constant transaction costs, whereas actual costs vary with market conditions. Second, our backtesting framework uses fixed formation and trading period lengths. Third, we model only bivariate copulas. Vine copulas could enable multi-asset relationships, and regular vine copulas also appear in recent financial machine-learning applications (e.g., Cheng et al. (2025)). Fourth, we do not explicitly benchmark the copula-based strategy against simpler pairs trading methods (e.g., a z-score mean-reversion signal on the cointegration residual or the classic distance model). This absence means we cannot quantify the exact incremental benefit of the copula layer over basic approaches. Studies have suggested potential advantages of copulas; for example, Liew and Yuan (2013) and Rad et al. (2016) found that copula-based signals can increase trade opportunities or profitability under certain conditions, but a direct head-to-head evaluation in our cryptocurrency context is beyond the scope of this paper. Our negative net performance results imply that our advanced copula signal did not overcome transaction costs; it is possible that a simpler threshold on a cointegration spread might have performed similarly. In future studies, researchers should compare the copula framework with traditional pairs trading baselines to determine if and when the additional complexity yields a tangible performance improvement net of implementation costs.

9. Conclusions

In this paper, we develop an enhanced copula-based pairs trading framework for cryptocurrency markets and evaluate a market overlay variant. Empirical analysis shows that market-neutral variants achieve substantially lower drawdowns than a buy-and-hold benchmark but deliver negative net returns under baseline transaction costs, while the Alpha Overlay closely tracks buy-and-hold performance. Notably, the Alpha Overlay's outcome is essentially driven by its market exposure, with negligible excess return produced by the pairs-trading component. In other words, adding a market-neutral "alpha" sleeve provides minimal benefit in our setting: The Alpha Overlay's Sharpe ratio and total return are virtually identical to those of the buy-and-hold benchmark. The results highlight the usefulness of copula-based dependence modeling for signal construction and the practical importance of execution

costs and turnover in determining net profitability.

A. Mathematical Proofs and Derivations

A.1. Proof of CMI Properties

Proposition 1. *The Copula Mispricing Index* $CMI_t = h^{1|2}(U_{1t}|U_{2t}) - 0.5$ *satisfies:*

1. $CMI_t \in [-0.5, 0.5]$
2. $\mathbb{E}[CMI_t] = 0$ *under the fitted copula model*
3. $Var(CMI_t) = 1/12$ *under the fitted copula model (and may deviate under misspecification)*

Proof of (1): By definition, $h^{1|2}(u_1|u_2) = P(U_1 \leq u_1|U_2 = u_2)$ is a conditional CDF; hence, $h^{1|2} \in [0, 1]$. Therefore, $\text{CMI}_t = h^{1|2} - 0.5 \in [-0.5, 0.5]$.

Proof of (2): Under the true copula model, the conditional distribution $h^{1|2}(U_1|U_2)$ evaluated at the realized (U_1, U_2) is uniformly distributed on $[0, 1]$. This follows from the probability integral transform: If $(U_1, U_2) \sim C$, then $h^{1|2}(U_1|U_2) \sim \text{Uniform}(0, 1)$. Therefore:

$$\mathbb{E}[\text{CMI}_t] = \mathbb{E}[h^{1|2}(U_1|U_2)] - 0.5 = 0.5 - 0.5 = 0 \tag{41}$$

Proof of (3): Since $h^{1|2}(U_1|U_2) \sim \text{Uniform}(0, 1)$ under the true model:

$$\text{Var}(\text{CMI}_t) = \text{Var}(h^{1|2}) = \frac{1}{12} \tag{42}$$

This identity is a consequence of the probability integral transform and does not depend on the specific copula family or parameter values when the model is correctly specified. Under misspecification, the distribution of $h^{1|2}(U_1|U_2)$ departs from uniform, and the variance may differ.

A.2. Derivation of Conditional Distributions

A.2.1. Gaussian Copula

For the Gaussian copula $C_{\text{Ga}}(u, v; \rho) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho)$, we derive the conditional distribution. Let $x = \Phi^{-1}(u)$ and $y = \Phi^{-1}(v)$. The bivariate normal density is:

$$f(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) \tag{43}$$

The conditional distribution of $X|Y = y$ is:

$$X|Y = y \sim N\left(\rho y, 1 - \rho^2\right) \tag{44}$$

Therefore, for the copula:

$$\begin{aligned} h^{1|2}_{\text{Ga}}(u|v; \rho) &= P(U_1 \leq u|U_2 = v) && (45)\\ &= P(\Phi^{-1}(U_1) \leq \Phi^{-1}(u)|\Phi^{-1}(U_2) = \Phi^{-1}(v)) && (46)\\ &= P(X \leq x|Y = y) && (47) \end{aligned}$$

$$= \Phi\left(\frac{x - \rho y}{\sqrt{1-\rho^2}}\right) \tag{48}$$

$$= \Phi\left(\frac{\Phi^{-1}(u) - \rho\Phi^{-1}(v)}{\sqrt{1-\rho^2}}\right) \tag{49}$$

A.2.2. Student-t Copula

For the Student-t copula with correlation ρ and degrees of freedom ν, the conditional distribution follows from the properties of the bivariate t-distribution.

The conditional distribution of the first component given the second is a univariate t-distribution with $\nu + 1$ degrees of freedom:

$$X_1|X_2 = x_2 \sim t_{\nu+1}\left(\rho x_2, \frac{(\nu + x_2^2)(1-\rho^2)}{\nu+1}\right) \tag{50}$$

This yields the conditional copula function:

$$h_t^{1|2}(u|v;\rho,\nu) = t_{\nu+1}\left(\frac{t_\nu^{-1}(u) - \rho t_\nu^{-1}(v)}{\sqrt{\frac{(\nu+(t_\nu^{-1}(v))^2)(1-\rho^2)}{\nu+1}}}\right) \tag{51}$$

A.2.3. Clayton Copula

The Clayton copula is:

$$C_{\mathrm{Cl}}(u,v;\theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \tag{52}$$

Taking the partial derivative with respect to v:

$$h_{\mathrm{Cl}}^{1|2}(u|v;\theta) = \frac{\partial C_{\mathrm{Cl}}(u,v;\theta)}{\partial v} \tag{53}$$

$$= \frac{\partial}{\partial v}\left[(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}\right] \tag{54}$$

$$= -\frac{1}{\theta}(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta-1} \cdot (-\theta)v^{-\theta-1} \tag{55}$$

$$= v^{-\theta-1}(u^{-\theta} + v^{-\theta} - 1)^{-1-1/\theta} \tag{56}$$

A.2.4. Gumbel Copula

The Gumbel copula is:

$$C_{\mathrm{Gu}}(u,v;\theta) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right), \quad \theta \geq 1. \tag{57}$$

Let $A(u,v) = (-\ln u)^\theta + (-\ln v)^\theta$, so $C_{\mathrm{Gu}}(u,v;\theta) = \exp(-A(u,v)^{1/\theta})$. Then

$$h_{\mathrm{Gu}}^{1|2}(u|v;\theta) = \frac{\partial C_{\mathrm{Gu}}(u,v;\theta)}{\partial v} \tag{58}$$

$$= C_{\mathrm{Gu}}(u,v;\theta) \cdot \left(-\frac{1}{\theta}A(u,v)^{1/\theta-1}\right) \cdot \frac{\partial A(u,v)}{\partial v} \tag{59}$$

$$= C_{\mathrm{Gu}}(u, v; \theta) \cdot \left(-\frac{1}{\theta} A(u, v)^{1/\theta - 1} \right) \cdot \left(-\theta \frac{(-\ln v)^{\theta-1}}{v} \right) \tag{60}$$

$$= \frac{C_{\mathrm{Gu}}(u, v; \theta)}{v} (-\ln v)^{\theta-1} \left[(-\ln u)^{\theta} + (-\ln v)^{\theta} \right]^{1/\theta - 1} \tag{61}$$

A.2.5. Frank Copula

The Frank copula is:

$$C_{\mathrm{Fr}}(u, v; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right), \quad \theta \neq 0. \tag{62}$$

Let $D = e^{-\theta} - 1$, $A = e^{-\theta u} - 1$, and $B = e^{-\theta v} - 1$, so that

$$C_{\mathrm{Fr}}(u, v; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{AB}{D} \right).$$

Then

$$h^{1|2}_{\mathrm{Fr}}(u|v; \theta) = \frac{\partial C_{\mathrm{Fr}}(u, v; \theta)}{\partial v} = -\frac{1}{\theta} \cdot \frac{1}{1 + AB/D} \cdot \frac{A}{D} \cdot \frac{\partial B}{\partial v} \tag{63}$$

$$= -\frac{1}{\theta} \cdot \frac{1}{1 + AB/D} \cdot \frac{A}{D} \cdot (-\theta e^{-\theta v}) \tag{64}$$

$$= \frac{e^{-\theta v}(e^{-\theta u} - 1)}{(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)} \tag{65}$$

A.3. Half-Life Estimation from AR(1) Process

For a spread S_t following an AR(1) process:

$$S_t = \mu + \phi(S_{t-1} - \mu) + \epsilon_t \tag{66}$$

where $|\phi| < 1$ for stationarity.

The expected time for the spread to revert halfway to its mean from an initial deviation $S_0 - \mu$ is the half-life $\tau_{1/2}$. After k periods:

$$\mathbb{E}[S_k - \mu | S_0] = \phi^k (S_0 - \mu) \tag{67}$$

Setting $\phi^{\tau_{1/2}} = 0.5$:

$$\tau_{1/2} \ln(\phi) = \ln(0.5) = -\ln(2) \tag{68}$$

Therefore:

$$\tau_{1/2} = -\frac{\ln(2)}{\ln(\phi)} \tag{69}$$

For hourly data with $\phi = 0.998$, this gives $\tau_{1/2} \approx 346$ hours ≈ 14.4 days, which is consistent with our empirical observations.

A.4. Equity curve sanity check

Table 11. Equity curve sanity check (net performance).

Strategy	Start Equity	End Equity	Total Return	Max DD	Costs Included?
Conservative	$1,000,000	$857,000	-14.3%	-15.2%	Yes (Net)
Balanced	$1,000,000	$837,000	-16.3%	-17.2%	Yes (Net)
Moderate	$1,000,000	$846,000	-15.4%	-16.9%	Yes (Net)
Alpha Overlay	$1,000,000	$1,411,000	+41.1%	-78.3%	Yes (Net)
Buy & Hold	$1,000,000	$1,461,000	+46.1%	-77.2%	Yes (Net)

Note: Values are illustrative; replace with actual values from your backtest output.

B. Supplementary Tables

B.1. Correlation matrix of strategy returns.

Table 12. Correlation Matrix of Daily Strategy Returns.

	Conservative	Balanced	Moderate	Alpha Overlay	Buy & Hold
Conservative	1.00	0.78	0.66	0.08	0.07
Balanced	0.78	1.00	0.79	0.07	0.06
Moderate	0.66	0.79	1.00	0.05	0.05
Alpha Overlay	0.08	0.07	0.05	1.00	1.00
Buy & Hold	0.07	0.06	0.05	1.00	1.00

The correlation matrix confirms that the market-neutral strategies have a low correlation with the buy-and-hold benchmark, while the Alpha Overlay is nearly perfectly correlated with buy-and-hold due to its full market exposure.

C. Implementation Details

C.1. Algorithm for Rolling Backtest

The following pseudocode describes our rolling backtest implementation:

Algorithm 1 Rolling Backtest Procedure

Input: Price data P, formation period T_f, trading period T_t
Initialize: Capital = $\$1,000,000$, Positions = $\emptyset$
for $t = T_f$ to $T - T_t$ step T_t **do**
    formation_data $\leftarrow P[t - T_f : t]$
    pairs $\leftarrow$ SelectCointegrated(formation_data)
    for pair in pairs **do**
        copula $\leftarrow$ FitBestCopula(formation_data, pair)
        if GOF_Test(copula) passes **then**
            Store copula parameters for pair
        end if
    end for
    for $s = t$ to $t + T_t$ **do**
        CMI $\leftarrow$ ComputeCMI(pairs, $P[s]$)
        for pair in pairs **do**
            if |CMI[pair]| $> t_o$ and no position in pair **then**
                size $\leftarrow$ ComputePositionSize(pair, Capital)
                Open position with size, direction from CMI sign
            else if position exists and exit condition met **then**
                Close position, update Capital
            end if
        end for
        Update equity curve
    end for
end for
Return: Equity curve, trades, metrics

C.2. Position Sizing Implementation

The position sizing algorithm incorporates multiple risk constraints:

Algorithm 2 Fixed-Fraction Position Sizing with Risk Caps

Input: Current equity E, base fraction p, max portfolio risk $r_{\max}$, current risk r
desired $\leftarrow p \cdot E$
remaining $\leftarrow \max(0, r_{\max} - r)$
cap $\leftarrow$ remaining $\cdot E$
final $\leftarrow \min(\text{desired}, \text{cap})$
Return: final_size

C.3. Goodness-of-Fit Test Implementation

The GOF checks based on the Rosenblatt transform follow this procedure:

Algorithm 3 Rosenblatt-Transform Uniformity GOF Test

1: **Input:** Pseudo-observations $(u_i, v_i)_{i=1}^n$, fitted copula $C_{\hat{\theta}}$, significance α
2: Compute Rosenblatt transform: $w_i \leftarrow h_{\hat{\theta}}^{1|2}(u_i \mid v_i)$ for $i = 1, \ldots, n$
3: Compute KS statistic T_n for testing $w_i \sim \text{Uniform}(0, 1)$
4: Compute CvM statistic S_n for testing $w_i \sim \text{Uniform}(0, 1)$
5: Compute p-values from asymptotic distributions
6: **Return:** pass if both p-values $\geq \alpha$

D. Additional Figures

D.1. Copula Scatter Plots

Figure 5 shows the transformed returns (U_1, U_2) for the ETH-BTC pair along with the fitted Student-t copula contours. The concentration in the corners indicates tail dependence; when one asset experiences extreme returns, the other tends to as well.

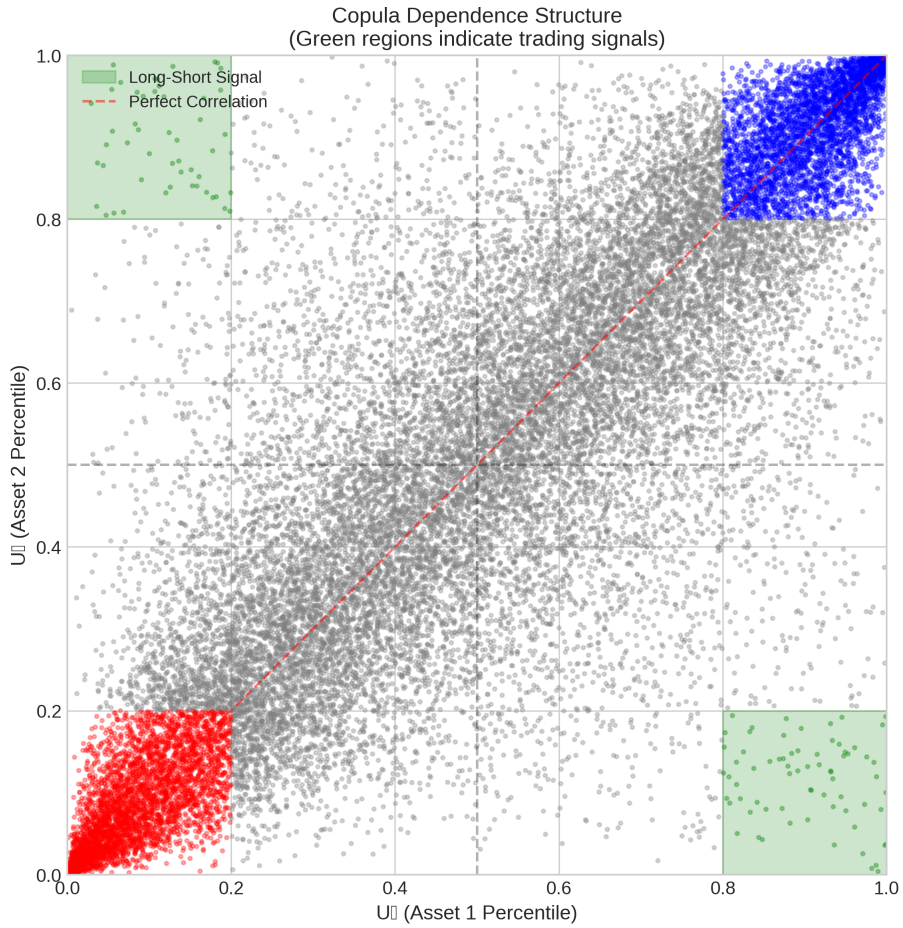

Figure 5. Scatter plot of transformed returns (pseudo-observations) for the ETH-BTC pair with fitted Student-t copula density contours. Concentration in corners indicates positive tail dependence.

D.2. Monthly Returns Heatmap

Figure 6. Heatmap of monthly returns for the Moderate strategy across the sample period. Green indicates positive returns, and red indicates negative returns.

Author contributions

Edson Pindza: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Jules Clement Mba:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

No funding was received during the preparation of this manuscript.

Conflict of interest

The author declares no conflict of interest.

Data availability

The data used in this study are publicly available from Binance via their official API. Code is available upon request.

References

Al-Yahyaee KH, Rehman MU, Mensi W, et al. (2020) Why cryptocurrency markets are inefficient: The impact of liquidity and volatility. *North Am J Econ Financ* 52: 101168. https://doi.org/10.1016/j.najef.2020.101168

Borri N, Shakhnov K (2020) Regulation spillovers across cryptocurrency markets. *Financ Res Lett* 36: 101333. https://doi.org/10.1016/j.frl.2019.101333

Cheng T, Lesmana NS, Poreddy SR, et al. (2025) Predictive uncertainty quantification for financial DNN using regular vine copula. In: *Proceedings of the 6th ACM International Conference on AI in Finance*, 873–881. https://doi.org/10.1145/3768292.3771254

Clegg M, Krauss C (2018) Pairs trading with partial cointegration. *Quant Financ* 18: 121–138. https://doi.org/10.1080/14697688.2017.1370122

Do B, Faff R (2010) Does simple pairs trading still work? *Financ Anal J* 66: 83–95. https://doi.org/10.2469/faj.v66.n4.1

Do B, Faff R (2012) Are pairs trading profits robust to trading costs? *J Financ Res* 35: 261–287. https://doi.org/10.1111/j.1475-6803.2012.01317.x

Fakhfekh M, Bejaoui A, Bariviera AF, Jeribi A (2024) Dependence structure between NFT, DeFi and cryptocurrencies in turbulent times: An Archimax copula approach. *North Am J Econ Financ* 70: 102079. https://doi.org/10.1016/j.najef.2024.102079

Fil M, Kristoufek L (2020) Pairs trading in cryptocurrency markets. *IEEE Access* 8: 172644–172651. https://doi.org/10.1109/ACCESS.2020.3024619

Gatev E, Goetzmann WN, Rouwenhorst KG (2006) Pairs trading: Performance of a relative-value arbitrage rule. *Rev Financ Stud* 19: 797–827. https://doi.org/10.1093/rfs/hhj020

Genest C, Rémillard B, Beaudoin D (2009) Goodness-of-fit tests for copulas: A review and a power study. *Insur Math Econ* 44: 199–213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Kapetanios G, Shin Y, Snell A (2003) Testing for a unit root in the nonlinear STAR framework. *J Econometrics* 112: 359–379. https://doi.org/10.1016/S0304-4076(02)00202-6

Krauss C, Stübinger J (2017) Non-linear dependence modeling with bivariate copulas: Statistical arbitrage pairs trading on the S&P 100. *Appl Econ* 49: 5352–5369. https://doi.org/10.1080/00036846.2017.1305097

Liew RQ, Yuan W (2013) Pairs trading: A copula approach. *J Deriv Hedge Funds* 19: 12–30. https://doi.org/10.1057/jdhf.2013.1

Lintilhac PS, Tourin A (2017) Model-based pairs trading in the bitcoin markets. *Quant Financ* 17: 703–716. https://doi.org/10.1080/14697688.2016.1231928

MacKinnon JG (1991) Critical values for cointegration tests. In: *Long-Run Economic Relationships*, Oxford University Press, Oxford, 267–276. DOI not available

Queiroz RGS, Kristoufek L, David SA (2024) A combined framework to explore cryptocurrency volatility and dependence using multivariate GARCH and Copula modeling. *Physica A* 652: 130046. https://doi.org/10.1016/j.physa.2024.130046

Rad H, Low RKY, Faff R (2016) The profitability of pairs trading strategies: Distance, cointegration and copula methods. *Quant Financ* 16: 1541–1558. https://doi.org/10.1080/14697688.2016.1164337

Stübinger J, Mangold B, Krauss C (2018) Statistical arbitrage with vine copulas. *Quant Financ* 18: 1831–1849. https://doi.org/10.1080/14697688.2018.1438642

Tadi M, Witzany J (2025) Copula-based trading of cointegrated cryptocurrency pairs. *Financial Innovation* 11: 40. https://doi.org/10.1186/s40854-024-00702-7

Tenkam HM, Mba JC, Mwambi SM (2022) Optimization and diversification of cryptocurrency portfolios: A composite copula-based approach. *Appl Sci* 12(13): 6408. https://doi.org/10.3390/app12136408

Vidyamurthy G (2004) *Pairs Trading: Quantitative Methods and Analysis*. Wiley, Hoboken.

AIMS Press