



Article

# Optimal Market-Neutral Multivariate Pair Trading on the Cryptocurrency Platform

Hongshen Yang \* and Avinash Malik

Department of ECSE, Faculty of Engineering, The University of Auckland, Auckland CBD, Auckland 1010, New Zealand; avinash.malik@auckland.ac.nz

\* Correspondence: hyan212@aucklanduni.ac.nz

**Abstract:** This research proposes a novel arbitrage approach in multivariate pair trading, termed the Optimal Trading Technique (OTT). We present a method for selectively forming a “bucket” of fiat currencies anchored to cryptocurrency for monitoring and exploiting trading opportunities simultaneously. To address quantitative conflicts from multiple trading signals, a novel bi-objective convex optimization formulation is designed to balance investor preferences between profitability and risk tolerance. We understand that cryptocurrencies carry significant financial risks. Therefore this process includes tunable parameters such as volatility penalties and action thresholds. In experiments conducted in the cryptocurrency market from 2020 to 2022, which encompassed a vigorous bull run followed by a bear run, the OTT achieved an annualized profit of 15.49%. Additionally, supplementary experiments detailed in the appendix extend the applicability of OTT to other major cryptocurrencies in the post-COVID period, validating the model’s robustness and effectiveness in various market conditions. The arbitrage operation offers a new perspective on trading, without requiring external shorting or holding the intermediate during the arbitrage period. As a note of caution, this study acknowledges the high-risk nature of cryptocurrency investments, which can be subject to significant volatility and potential loss.

**Keywords:** pair trading; multivariate; quantitative trading; cryptocurrency market



**Citation:** Yang, Hongshen, and Avinash Malik. 2024. Optimal Market-Neutral Multivariate Pair Trading on the Cryptocurrency Platform. *International Journal of Financial Studies* 1: 0. <https://doi.org/>

Academic Editor: Babacar Sène

Received: 17 June 2024

Revised: 2 August 2024

Accepted: 5 August 2024

Published:



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The secondary market functions as a modern, efficient form of double auction where participants intending to trade assets submit their acceptable bid and ask prices. The highest bid and lowest ask are matched if the prices meet. We believe in the Efficient Market Hypothesis (EMH), which posits that in the secondary market, prices should reflect all available information (Fama 1970). However, historical economic bubbles show that investors do not always act rationally. Dale et al. (2005) illustrated how an order book filled with irrational bids and asks could drive prices in the wrong direction. Identifying inefficiencies in an efficient market helps maintain trading equilibrium. Market directional trading involves astute traders estimating more accurate prices for instruments. They then place orders with the expectation that the market will revert to rationality. Traders who bet correctly are rewarded by the market mechanism, contributing to price discovery. Quantitative trading, instead of relying on experienced accountants or financial analysts, utilizes mathematical modeling for better price estimations.

The cryptocurrency market is an emerging and yet immature secondary market. The recent cryptocurrency market trend benefits from technological advancements and global participation. Many people criticize the significant investment risk; however, its volatility, liquidity, and 24/7 trading make it an ideal platform for algorithmic trading. The cryptocurrency exchange users can trade between cryptocurrencies and fiat currencies, exploiting inefficiencies and capturing anomalies from both forex and cryptocurrency volatility.

Our research focuses on pair trading, a widespread strategy based on the Law of One Price (LOP), suggesting that homogeneous items should have the same price everywhere.

However, real world restrictions like trade frictions and legal regulations often violate this principle (Isard 1977; Miljkovic 1999). Similar assets should have similar price trends. However, their price independence can cause divergence, and convergence can often result in profit. Arbitrage-based trading approaches, grounded in LOP, are perceived as having more stable profitability. Investment banks and hedge funds have used pair trading for decades (Burgess 2000; Elliott et al. 2005; Gatev et al. 2006). While pair trading is common, the extension of pair trading to many-pair (multivariate) trading is not well studied. This paper delves into pair trading and introduces an optimization-based technique for multivariate assets, tested in the cryptocurrency market.

This research has several novel contributions: ① Unlike classical techniques (Dunis and Ho 2005; Galenko et al. 2012; Perlin 2007), our approach uses optimization to arbitrage multiple pairs simultaneously to the individual asset level without grouping assets. ② The bi-objective optimization is adjustable to investor preferences, including tunable parameters like risk tolerance and operating thresholds for maximizing return and minimizing risk. ③ Our arbitrage approach avoids holding volatile intermediate cryptocurrency, maintaining positions in fiat currencies to lower risk. ④ No borrowing is required for shorting, as all trading actions are achievable within the initial principal. ⑤ The proposed technique is robust, market-neutral, and profitable in both bull and bear markets.

The remainder of the article is structured as follows. Sections 2 review the most cited works in pair trading and outline the proposed trading idea. Sections 3 introduce the prerequisite knowledge and details of the proposed trading methodology, respectively. Section 4 presents the simulation algorithm. Finally, Sections 5 and 6 present the results and conclusions, respectively. Additional experiments and findings are detailed in Appendix A.

## 2. Related Work

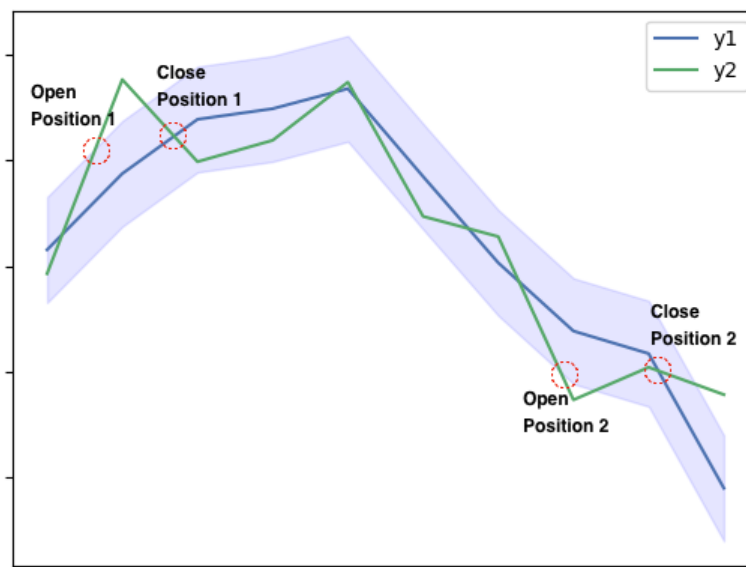
### 2.1. Pair Trading

Long-Term Capital Management L.P. (LTCM) was a highly leveraged hedge fund founded by John Meriwether in 1994 that operated on bond pair trading. It is famous for its billions of dollars bailout after the 1997 financial crisis and is known as an early real-life example of pair trading (Slivinski 2009).

Pair trading is straightforward, as shown in Figure 1, which illustrates the most cited baseline framework—the Distance Method (Gatev et al. 2006). The common trading terminology includes *long* positions as buying assets to sell later at a higher price, and *short* as borrowing and selling assets, intending to buy them back at a lower price. We can summarize pair trading in the following steps:

1. Search for two assets with the closest historical price movements,  $y_1$  and  $y_2$ , and group  $y_1$  and  $y_2$  as a *trading pair*.
2. Calculate the price difference between  $y_1$  and  $y_2$  as *distance*.
3. Open a position when the distance surpasses two times its historical standard deviation (*open threshold*, shown as the band area with  $y_1$ ) by shorting  $y_2$  and longing  $y_1$ .
4. Close the position when  $y_1$  and  $y_2$  prices converge (*close threshold*), then close the position by reversing the buy and sell actions.
5. If the price of  $y_2$  falls below the threshold, open a position by doing the opposite: long  $y_2$  and short  $y_1$  until the next crossing point.
6. Repeat this practice until the end of the trading period.

A critical property of pair trading is achieving *market neutrality* by betting equivalent amounts on the long and short positions. Since the loss on one position is compensated by the other position, the portfolio value tends to remain stable with the overall market direction. Chen et al. (2018) formed pairs from asset returns instead of prices, suggesting that a temporarily weaker return implies a stronger rebound and vice versa. Both Gatev et al. (2006) and Chen et al. (2018) based their work on correlation.



**Figure 1.** Pair trading between assets  $y_1$  and  $y_2$  with threshold band.

Besides correlation, cointegration is a key concept in time-series analysis. Cointegration represents a long-run relationship between two time-series data sets (Engle and Granger 1991). Vidyamurthy (2004) examined assets for cointegration using the Engle–Granger test and combined assets based on correlation and cointegration, providing guidance for cointegration-based pair trading.

Further studies discuss multivariate pair trading, hedging market risk by managing positions and creating artificial pairs from scenarios with more than two assets. Perlin (2007) developed a technique combining assets into an artificial index to form a pair. Instead of finding a pair, he created a synthetic index  $P_{it}^*$  to anchor a real index  $P_{it}$ . The synthetic index  $P_{it}^*$  was constructed by weighting ( $w$ )  $m$  assets based on their prices:

$$P_{it}^* = \sum_{k=1}^m w_k P_{kt}^*$$

Weights among the participating assets were tuned to match  $P_{it}$  as closely as possible. Next,  $P_{it}$  and  $P_{it}^*$  were arbitrated for pair trading. Such an artificially formed pair-trading methodology sometimes also uses cointegration techniques. Galenko et al. (2012) used a moving window method to pair trade among a group of assets, calculating the cointegration vector  $b$  to track  $Z_t$  from the asset return combined with the cointegration value  $b$ :

$$Z_t = \sum_{i=1}^N b^i r_t^i$$

Assets were then partitioned into two categories based on  $b$  results:  $L$  (long) and  $S$  (short). Quantified  $Z_t$  provided recommendations on investment amounts for each asset. Dunis and Ho (2005) adopted a similar methodology, tracking the EURO STOXX50 and split outperformers and underperformers into STOXX50+ and STOXX50-. They pair traded STOXX50+ and STOXX50- to achieve “double alpha”, a weak form of pair trading due to non-negligible market risk.

There are many practices for pair trading and even multivariate pair trading techniques, but most existing pair trading techniques focus on compounding pairs rather than focusing on the individual granularity. Our research aims to provide a technique that neither creates an artificial index nor binds a group of assets. Using optimization tech-

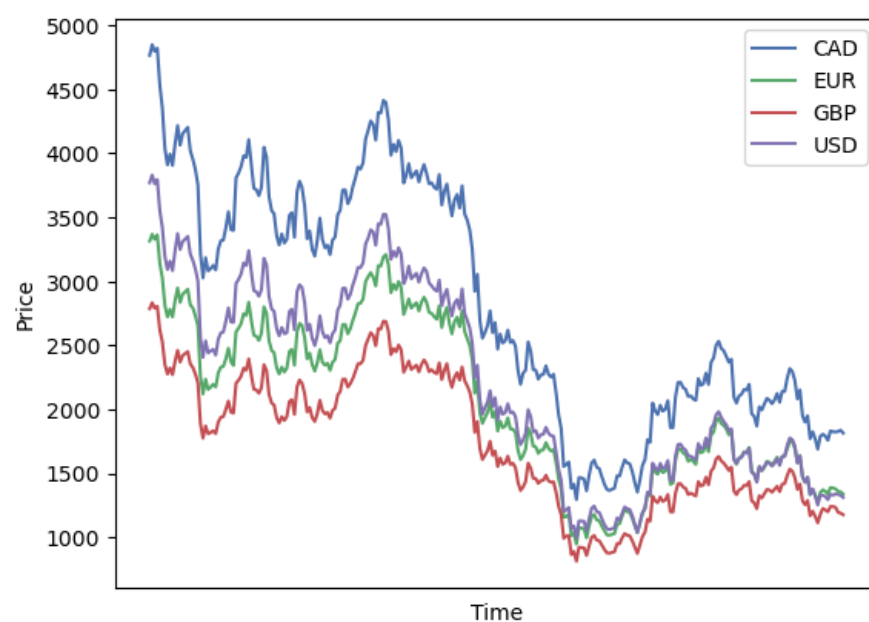
niques, we can operate directly on all participating assets to provide an optimal solution for multivariate pair trading.

## 2.2. Cryptocurrency

The work in [Fang et al. \(2022\)](#) provides a comprehensive survey of the history and recent research on cryptocurrencies. The cryptocurrency began with [Nakamoto \(2008\)](#)'s groundbreaking work in 2008, introducing a decentralized monetary system based on a peer-to-peer ledger recording system. The token that operates within the monetary system is called Bitcoin (BTC). By incentivizing Bitcoin users to record a full transaction ledger on their own computers, the ledgers are distributed across countless individual nodes instead of a centralized server. This property enables frictionless international participation in Bitcoin. Furthermore, international participation provides excellent liquidity for Bitcoin trades. Shortly after, Bitcoin's success led to the creation of imitators. Ethereum (ETH), the second most famous cryptocurrency, was created by [Buterin \(2014\)](#) for its programmable extensibility. ETH gained its reputation from its fast development and widespread adoption in recent years ([Bai et al. 2020](#)).

Popular pair trading techniques from Section 2.1 have also been introduced in the cryptocurrency market. [Fil and Kristoufek \(2020\)](#) applied the distance method based on the work of [Gatev et al. \(2006\)](#), achieving a 3% monthly profit in the cryptocurrency market. [van den Broek \(2018\)](#) formed pairs based on 60-day trading periods to profit from arbitrage opportunities, demonstrating existing inefficiencies of the cryptocurrency market.

Similar to gold, ETH has a global market against multiple fiat currencies. ETH can be traded independently with most major fiat currencies. And the unsynchronization of ETH against different fiat currencies is the inefficiencies in the market that provide trading opportunities. If a large volume of ETH is bought in USD, the value of ETH against USD will appreciate relative to ETH against GBP and CAD. This value discrepancy is the source of profit. There are some advantages of picking ETH as an anchor over gold because it is completely fungible and has low international transaction frictions. As shown in Figure 2, the CAD, EUR, GBP, and USD follow a similar trend but do not have 100% correlation. Our technique aims to capture minor volatility movements and arbitrage market inefficiencies.



**Figure 2.** The price of ETH in multiple currencies.

Though the cryptocurrency system operates in a decentralized manner, users can deposit their cryptocurrency into a centralized exchange for trading. The exchange func-

tions similarly to stock or forex exchanges, allowing users to freely trade cryptocurrency for another cryptocurrency or fiat currency, either on spot or margin trading. Due to the decentralized nature of participants, cryptocurrency exchanges operate 24/7 with liquidity coming from global participation. These properties have led to experiments on cryptocurrency exchanges since as early as 2014 (Brandvold et al. 2015; Cheah and Fry 2015; Shah and Zhang 2014).

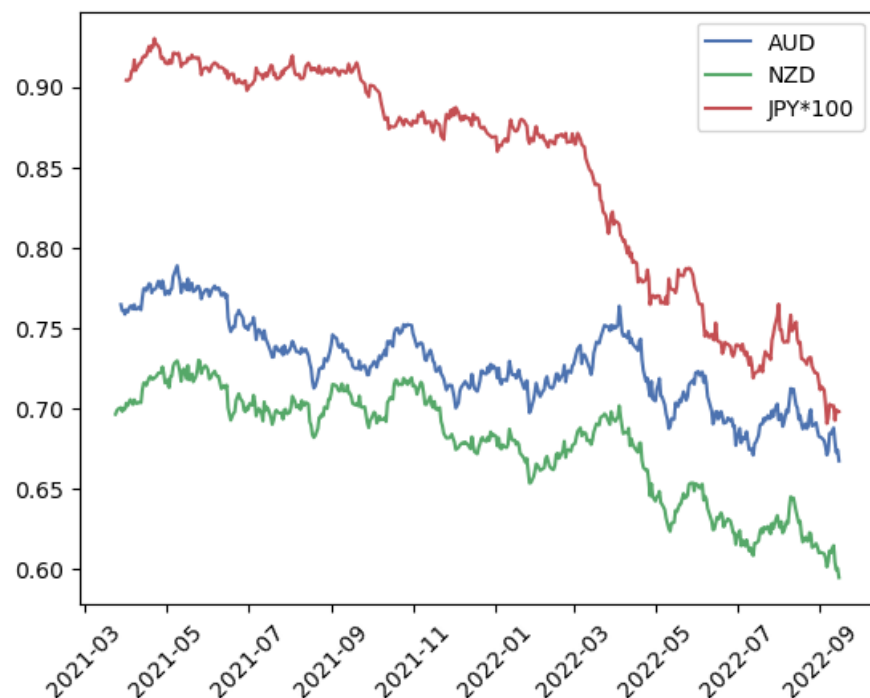
However, it is important to caution about the risks associated with cryptocurrency investment. Kerr et al. (2023) state that cryptocurrency carries a significant risk of financial fraud. Both Kerr et al. (2023) and Pakhnenko et al. (2023) compare cryptocurrencies to financial bubbles, highlighting the potential for significant market volatility and instability. This paper does not encourage investment in cryptocurrencies; rather, we only use them as intermediates for arbitrage in pair trading.

### 3. Methodology

There are two stages in our trading technique. The first stage involves screening the assets. We select assets that are correlated and cointegrated with each other for a trading bucket. Once the bucket is formed, we monitor the differences between each asset and conduct trades based on opening and closing trading signals.

#### 3.1. Screening

In the screening stage, we analyze historical data to ensure the assets are suitable for arbitrage. Since movements are impacted by the volatility of both cryptocurrency and fiat currency, severe inflation or deflation in a fiat currency could mislead our trading signals. Therefore, one major criterion is that the volatility of fiat currencies should be small compared to that of cryptocurrency. An example of strong fiat depreciation is illustrated in Figure 3: the difference between AUD and NZD is relatively stable compared to the rapid drop of JPY, which would send a short signal based on its value change without mean reversion. To select appropriate fiat currencies for the trading bucket, we reviewed the historical trends of fiat currencies to ensure they have the desired statistical relationships.



**Figure 3.** The volatility of forex rate against USD.

The first desired statistical property is *correlation*. We examine correlation using Pearson's correlation coefficient (Cohen et al. 2009), which ranges between  $[-1, 1]$ . The higher the absolute value of the coefficient, the stronger the association between the two variables. A coefficient of  $-1$  represents an absolute negative association,  $1$  represents an absolute positive association, and  $0$  indicates no association between datasets.

The second desired statistical property is *cointegration*. We refer to the price difference between two assets as *spread*. In the long term, a widening spread exposes arbitrageurs to bankruptcy risk. Therefore, the spread between two assets must have a constant mean and variance. A time-series dataset with these characteristics is called *stationary*. Cointegration describes the relationship between two time-series datasets whose difference is stationary. Cointegration is vital in pair trading because forming a pair with cointegration in the cryptocurrency market means that longing one asset and shorting another effectively holds the price difference. If the spread is stationary with a constant mean, convergence should always follow divergence. This property is called *mean reversion*. In this research, the Engle–Granger test (Engle and Granger 1987) is used to examine the cointegration relationship.

### 3.2. Long and Short Trading Signals

After selecting the assets, we need to establish rules for trading the multivariate assets. Key questions include: How do we ensure market neutrality for multiple assets? How should we define and react to trading signals? How do we allocate our principal capital?

#### 3.2.1. Multivariate Pairs

We define a consistent notation for *pairs*:

- ETH/USD or ETH/CAD denote the price of ETH in USD or CAD, respectively.
- A pair between ETH/USD and ETH/CAD, is referred to as ETH/USD:ETH/CAD, abbreviated as USD:CAD.

Theoretically, we can form a maximum of  $k(k-1)/2$  pairs from  $k$  assets.

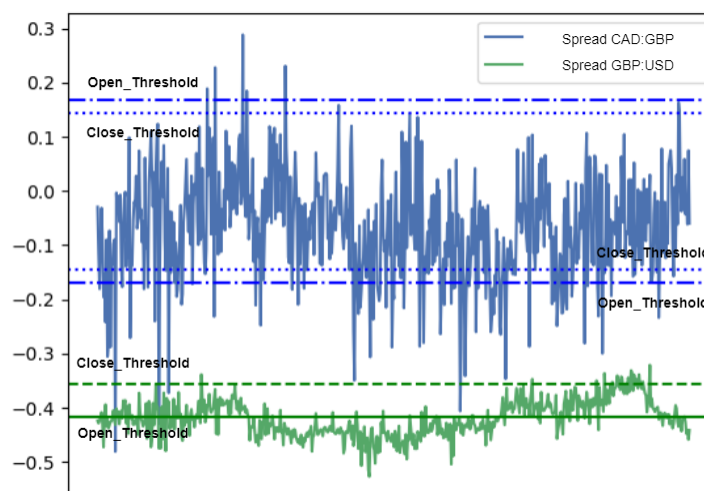
#### 3.2.2. Buy/Sell Signaling

Arbitrage relies on anomalies. We extract anomalies by comparing the normalized spread with the standard deviation. We use two parameters, *open\_threshold* and *close\_threshold*, to tune the trading signals. For a two-asset trading scenario, the trading process between assets  $i$  and  $j$  is as follows:

- Calculate the difference of the logarithmized *price* to obtain the *spread* array.
- Normalize the spread array using the *z-score*.
- Open a new position if there is no existing position and the  $z\text{-score} > \text{open\_threshold} \times \text{standard\_deviation}$ .
- Close the position if there is an open position and the  $z\text{-score} < \text{close\_threshold} \times \text{standard\_deviation}$ .
- Repeat until the end of the arbitrage period.

We provide a graphical example to explain the proposed approach. Table 1 gives the statistics and graphs for a three-asset scenario. Since each asset moves independently, each pair has a different standard deviation after normalization. The spreads are volatile, with thresholds represented by auxiliary lines. Every touch of the *open\_threshold* indicates a position-opening signal. Conversely, if an open position exists and the spread touches the *close\_threshold*, it signals a position closing.



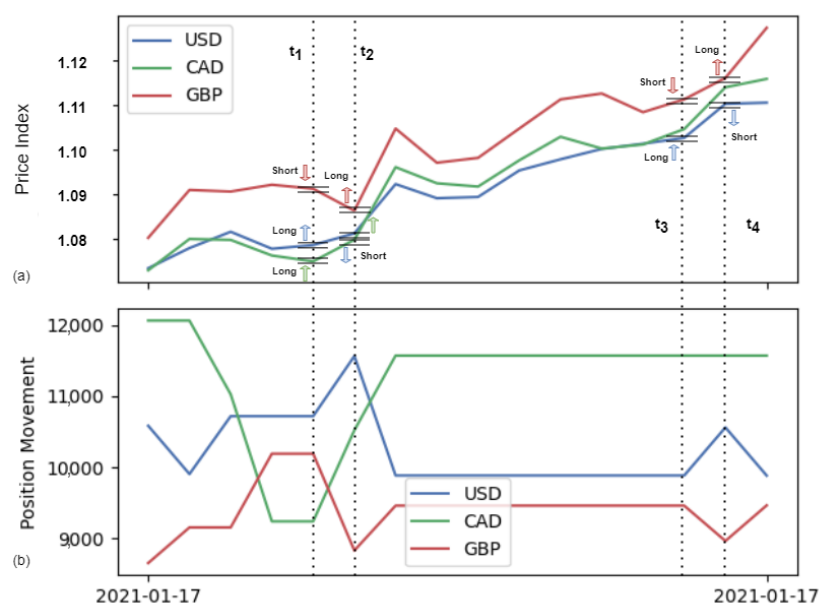
**Table 1.** Spread trading signaling.

Spread	Standard Deviation	Open_Threshold	Close_Threshold
USD:CAD	0.024017	0.168122	0.144102
USD:GBP	0.059545	0.416817	0.357270
CAD:GBP	0.049811	0.348677	0.298866

In a cut-off period on 17 January 2021, a visualized spread-trading signaling based on the *open\_threshold* as 7 and *close\_threshold* as 6.

### 3.2.3. Position Movements

Suppose we return from the spreads to the actual asset granularity (Figure 4). The initial principal is distributed among the fiat currencies we are trading. Traditionally, we long the loser and short the winner for every positive signal. However, multiple spreads might produce conflicting signals. The example in Figure 4a shows the normalized ETH price growth in USD, CAD, and GBP, while Figure 4b tracks the principal movements of our fiat currency holdings.



**Figure 4.** Actions and position movements for a three-currency scenario with plot (a) as price indices with trading positions and plot (b) as position movements over time.

At time  $t_1$ , GBP starts a rapid decline. Both relevant spreads GBP:CAD and GBP:USD shrink, triggering trading signals. Our corresponding actions are longing ETH with GBP and shorting ETH to USD and CAD. At a later time  $t_2$ , there is a rebound in the ETH price against all currencies. Although all fiat currencies are rising, the USD reversal is weaker than the others, making the USD:CAD and USD:GBP spreads greater than the open threshold. This triggers the use of USD to buy ETH and sell ETH for CAD and GBP. By design, it is possible for a fiat currency to be longed and shorted simultaneously in different spreads.

The classical scenario of two-asset trading still exists because each spread is monitored individually. For example, the movement at time  $t_3$  does not trigger trading in CAD. The latency of rebound in USD makes only the  $z$ -score of USD:GBP greater than the open threshold, so a position is opened for the spread USD:GBP. Later after time  $t_4$ , when the spread widens, the position opened at time  $t_3$  is closed.

There are more benefits to including multivariate assets. If a fiat currency receives a long signal from one pair and an equivalent short signal from another, the actions can cancel out on paper, saving fees. Furthermore, when one fiat currency reveals a strong anomalous movement, all relevant pairs will display signals for the principal flowing in or out of that fiat currency. If there is an entry signal in USD:GBP, all the USD can be used to conduct the operation, ensuring no USD sits idle. Though ideally, this mitigation saves on transaction fees, signals that cancel each other are rare. More often, we have multiple short and long signals happening simultaneously. Furthermore, one short signal may come from a position closure and another from a new position opening. Therefore, we need to devise a mechanism to determine how funds should be distributed precisely.

### 3.3. An Optimal Approach to Principal Allocation

Resource distribution addresses how much to invest in each asset whenever a trading opportunity arises. At every trading interval, as long as a trading opportunity is detected, the optimization process is executed to make a quantitative decision on the amount to invest. Optimization is unnecessary if we only trade a two-asset pair since all the capital (or a pre-determined amount) can be allocated for the long and short positions. However, with more than two assets, the multilateral relationships necessitate distributing resources for each signal result to achieve the optimal outcome. The optimization process aims to maximize return with minimum risk under given constraints. In this research, a bi-objective optimization formulation inspired by the Markowitz portfolio allocation (Markowitz 1952) is employed to find the optimal investment solution.

#### 3.3.1. Parameters

1. Let  $\mathbf{c}$  be a vector of participating currencies, where  $|\mathbf{c}| \in \mathbb{Z}^+$  represents the number of currencies.
2. Let  $\mathbf{n}$  be a vector of pairs formed from currencies, where  $|\mathbf{n}| \in \mathbb{Z}^+$  represents the number of pairs.
3. Let  $\mathbf{t} \in [1, 2, 3, \dots, T]$  represent the timestamp of execution, and  $T$  denotes the total length of the arbitrage period.
4. Let  $\mathbf{P} \in \mathbb{R}^{|\mathbf{c}| \times T}$  be a matrix representing ETH prices in  $|\mathbf{c}|$  fiat currencies for the arbitrage period  $T$ .

$$\begin{cases} \text{ETH/USD} : [P_{USD,t_1}, P_{USD,t_2}, \dots, P_{USD,t_T}] \\ \text{ETH/CAD} : [P_{CAD,t_1}, P_{CAD,t_2}, \dots, P_{CAD,t_T}] \\ \text{ETH/GBP} : [P_{GBP,t_1}, P_{GBP,t_2}, \dots, P_{GBP,t_T}] \end{cases}$$

5. Let  $\mathbf{r} \in \mathbb{R}^{|\mathbf{c}|}$  records the average returns of the ETH price in  $|\mathbf{c}|$  fiat currencies throughout the arbitrage period, and let expected profit  $\mathbf{EP} \in \mathbb{R}^{|\mathbf{n}| \times 2}$  be a matrix for  $|\mathbf{n}|$  pairs recording the average expected profit once an arbitrage position is opened. Since each spread consists of two different assets, the speeds of mean reversion should differ



from pair to pair. Therefore, we measure a mean-reversion vector  $\mathbf{mr} \in \mathbb{R}^{|\mathbf{n}|}$  to record the average reversion time to the mean from historical data.  $\mathbf{EP}$  is calculated based on the combination of returns and mean-reversion speed as follows:

$$\begin{bmatrix} mr_{USD,CAD} \times r_{USD} & mr_{USD,CAD} \times r_{CAD} \\ mr_{USD,GBP} \times r_{USD} & mr_{USD,GBP} \times r_{GBP} \\ mr_{CAD,GBP} \times r_{CAD} & mr_{CAD,GBP} \times r_{GBP} \end{bmatrix}$$

6. Let covariance  $\mathbf{COV} \in \mathbb{R}^{|\mathbf{n}| \times 2 \times 2}$  be a three-dimensional matrix. It contains  $2 \times 2$  matrices for  $|\mathbf{n}|$  pairs. The covariance matrix is leveraged to compute the risk of opening a position. An individual  $2 \times 2$  matrix from the three-dimensional slice is a variance–covariance matrix for the two fiat currencies in a pair. An example of the  $2 \times 2$  matrix is shown below:

$$\begin{bmatrix} \text{var}(USD) & \text{cov}(USD, CAD) \\ \text{cov}(USD, CAD) & \text{var}(CAD) \end{bmatrix}$$

7. Let transaction cost  $tc \in \mathbb{R}^{\geq 0}$  be a constant, representing the percentage deducted from each investment amount as commission.
8. Let trading weights  $\mathbf{tw} \in \mathbb{R}^{|\mathbf{c}|}$  be a vector representing the weight occupation of each currency. Since our arbitrage is a continuous process, every trade decision incorporates historical information. A strong buy signal guides the optimization process to invest all available USD into ETH. Without using margin trading, the maximum investment is 100%. For example, if 40% of USD is already used in an open position (long USD), the maximum available for the next signal is 60%. This 40% represents a weight that is actively being traded, hence the term trading weights, which can be positive or negative.

### 3.3.2. Variable

Let Weights  $\mathbf{W} \in \mathbb{R}^{N \times 2}$  be the variable of the optimization in the form of a matrix for weights from every pair. The first column is the long currencies with positive weights, and the second column is the short currencies with negative weights. The negative weights capture the fact that the capital is reduced when going long, and the capital increases when going short. For instance:

$$\begin{bmatrix} W_{USD} & W_{CAD} \\ W_{USD} & W_{GBP} \\ W_{CAD} & W_{GBP} \end{bmatrix}$$

### 3.3.3. Objective

The objective of the optimization is to maximize the return while minimizing the risk. We receive trading signals from spreads, so the decisions for investment weights are in pairs. When we need to long USD and short CAD in the pair USD:CAD, we are effectively applying the weight to bet on the rise of  $r_{USD}$  and the drop of  $r_{CAD}$ . The expected profit is calculated as the weights multiplied by the expected profit, presented as:

$$\sum_{i=1}^{|\mathbf{n}|} W_n \cdot (\mathbf{EP}_n \odot [1, -1])', \forall n \in \mathbf{n} \quad (1)$$

where  $\odot$  is the Hadamard product, with long weights positive and short weights negative. We also aim to minimize risk, meaning the portfolio should have as little volatility as possible. According to Modern Portfolio Theory (Markowitz 1952), the portfolio volatility is  $\mathbf{W} \times \mathbf{COV}_{2 \times 2} \times \mathbf{W}'$  for any pair spread in  $\mathbf{n}$ . Since long weights are positive and short weights are negative, we need to take the Hadamard product with  $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$  to correctly

capture the portfolio covariance in the covariance–variance matrix. Overall, for  $\mathbf{n}$  pairs, we obtain Equation (2).

$$\sum_{i=1}^{|\mathbf{n}|} W_n \cdot \mathbf{COV}_n \odot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \cdot W'_n, \forall n \in \mathbf{n} \quad (2)$$

### 3.3.4. Constraints

The first constraint is that we can never invest more than what we have. Therefore, in every trading pair, the long weights cannot surpass 100%, and the short weights shall not be less than  $-100\%$ :

$$0 \leq W_{n,long} \leq 1, -1 \leq W_{n,short} \leq 0, \forall n \in \mathbf{n} \quad (3)$$

Another constraint comes from the weight at the currency level. Every currency should not be spent more than 100% at any time of the trade. Every time the optimization process runs, it calculates the summation of the current trading weight, long weight, and short weight to ensure that it does not exceed 100% for each individual currency.

$$tw_c + \sum_{n \in \mathbf{n}} W_{c,long} - \sum_{n \in \mathbf{n}} W_{c,short} \leq 1, \forall c \in \mathbf{c} \quad (4)$$

Market neutrality is the ability to generate stable income regardless of the market condition, and it is one of the major incentives for investors to select pair trading from other quantitative trading methods. Market neutrality requires an equivalent distribution of investment in the long and short positions. In that, we need to translate the weights back to the amount of ETH. The price of currency  $c$  at time  $t$  is attainable as  $P_{c,t}$ . In our case, whenever an ETH is bought in one currency, that ETH should be sold instantaneously. Hence, we expect the last constraint is that every pair position reflecting back on assets is equalized in the unit of ETH. Transaction cost tweaks the balance slightly. The amount of ETH received from longing is reduced because of transaction costs. However, the amount of ETH required to obtain a certain amount of fiat currency increases. The market neutrality constraint is presented in Equation (5):

$$-W_{n,short} \cdot (1 + tc)P_{c_1,t} = W_{n,long} \cdot (1 - tc)P_{c_2,t} \\ \forall t \in \mathbf{t}, \forall (c_1, c_2) \in \mathbf{c}, \forall n \in \mathbf{n} \quad (5)$$

### 3.3.5. Optimization Equation

Given the above, for the purpose of maximizing profitability and minimizing risk, one can optimize the allocation of the principal capital by solving the bi-objective optimization problem in Equation (6). The ultimate goal of the optimization process is to calculate a set of quantitative weights when opening any position on the pair spreads.

We understand that market participants have different tendencies toward risk aversion. Therefore, we use a coefficient  $\lambda$  (default to 1) to measure the degree of risk aversion in the objective. The greater the  $\lambda$ , the greater the risk aversion. In extreme cases,  $\lambda = 0$  stands for risk neutrality. The  $\lambda$  and portfolio variance form a penalty factor in the objective to cater to investors' risk preferences.

$$\begin{cases} \text{obj : } \sum_{i=1}^{|\mathbf{n}|} W_n \cdot (\mathbf{EP}_n \odot [1, -1])' \\ \quad - \lambda \sum_{i=1}^{|\mathbf{n}|} W_n \cdot \mathbf{COV}_n \odot \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \cdot W'_n, \forall n \in \mathbf{n} \\ \text{s.t. } 0 \leq W_{n, \text{long}} \leq 1, -1 \leq W_{n, \text{short}} \leq 0, \forall n \in \mathbf{n} \\ \text{s.t. } tw_c + \sum_{n \in \mathbf{n}} W_{c, \text{long}} - \sum_{n \in \mathbf{n}} W_{c, \text{short}} \leq 1, \forall c \in \mathbf{c} \\ \text{s.t. } -W_{n, \text{short}} \cdot (1 + tc)P_{c_1, t} = W_{n, \text{long}} \cdot (1 - tc)P_{c_2, t} \\ \quad \forall t \in \mathbf{t}, \forall (c_1, c_2) \in \mathbf{c}, \forall n \in \mathbf{n} \end{cases} \quad (6)$$

where:

- $\mathbf{n}$ : set of all pairs;
- $W_n$ : weight vector for the  $n$ -th pair;
- $\mathbf{EP}_n$ : expected profit vector for the  $n$ -th pair;
- $\odot$ : Hadamard product (element-wise multiplication);
- $\lambda$ : risk aversion coefficient;
- $\mathbf{COV}_n$ : covariance matrix for the  $n$ -th pair;
- $W_{n, \text{long}}$ : long weight for the  $n$ -th pair;
- $W_{n, \text{short}}$ : short weight for the  $n$ -th pair;
- $tw_c$ : trading weight for currency  $c$ ;
- $\mathbf{c}$ : set of all currencies;
- $tc$ : transaction cost;
- $P_{c, t}$ : price of currency  $c$  at time  $t$ ;
- $\mathbf{t}$ : set of all time periods.

#### 4. Simulation

This section describes the process and structure of our simulation (Algorithm 1) on historical data. The process consists of multiple stages: *screening*, *hist\_analysis*, *trading*, and *optimization*.

The *screening* (cf. Section 3.1) determines which assets should be included within our trading bucket. ETH pricing data are extracted from crypto exchanges against fiat currencies as the input (Algorithm 1 line 1). We choose the trading period by splitting the pricing data into formation data and trading data. In the formation period, ETH prices against fiat currencies are screened to fill the bucket with fiat currencies that have strong correlations and high cointegration test passing rates. Once we drop the fiat currencies with weak correlation/cointegration relationships, the screened currencies are included in *participating\_assets*. Then, we call the function *hist\_analysis*.

In the *hist\_analysis* process (Algorithm 1 Line 1), standard deviation, expected return, and expected risk are computed using the historical formation data. Parameters *open\_threshold* and *close\_threshold* are computed using the formation data (cf. Section 3.2.2). Knowing when to trade is not sufficient; we also need to calculate how much to bid/ask on each order. The *trading* function (Algorithm 1 line 1) simulates by amount. The first step is to set up *orig\_amount*, which is the initial amount of each participating fiat currency at the beginning of the trade. Secondly, every two currencies are combined into spreads. Next, we loop through all the trading times and constantly check if any open position meets the closing criteria:

- If yes, we close the open position to harvest profit. The trading weight vector  $\mathbf{tw}$  is also updated as described in Section 3.3.
- If no, we leave it as is and check again in the next time sample.

**Algorithm 1:** Arbitrage Technique Simulation

---

```

Data: Pricing_Data
/* Historical pricing data for cryptocurrency against fiat currencies */

def optimisation(trading_data, expected_return, expected_risk):
    invest_weights ← trading_data, expected_return, expected_risk
    /* Optimize investment weights from historical data using Equation (6) */
    return invest_weights

def trading(open_date, close_date, trading_data, open_threshold, close_threshold,
expected_return, expected_risk):
    orig_amount()
    /* Set the starting amount of each asset */
    spread_data ← trading_data
    /* Calculate spread data from asset trading period pricing data */
    for time in [open_date, close_date] do
        if |spread_data[time]| < close_threshold & position_open = True then
            close_position()
            /* Close position for a pair when there is an open position and the
            spread is smaller than the closing threshold */
        end
        if |spread_data[time]| > open_threshold & position_open = False then
            invest_weights ← optimisation(trading_data, expected_return,
            expected_risk)
            open_position(invest_weights, orig_amount)
            /* Open position when the spread has no current position and is greater
            than the opening threshold */
        end
    end
end

def hist_analysis(formation_data, trading_data):
    std_dev, expected_return, expected_risk ← formation_data
    /* Calculate standard deviation, expected risk/return among spreads */
    open_threshold, close_threshold ← std_dev, formation_data
    /* Set trading thresholds between spreads */
    trading(open_date, close_date, trading_data, open_threshold, close_threshold,
    expected_return, expected_risk)
    /* Call trading function to trade from open_date to close_date */

def screening(pricing_data):
    formation_data, trading_data ← pricing_data
    /* Split the data into formation period and trading period */
    assets_correlation, assets_cointegration ← formation_data
    /* Calculate the correlation and cointegration between assets */
    participating_assets ← assets_correlation, assets_cointegration
    /* Select assets above threshold */
    formation_data, trading_data ← formation_data, trading_data,
    participating_assets
    /* Filter the data to only participating assets */
    hist_analysis(formation_data, trading_data)
screening(Pricing_Data)

```

---

Similarly, all the spreads that do not have an ongoing open position need to be checked at every time sample to see if an open signal is triggered. Different from closing a position, open-position actions must consider how much *orig\_amount* to spend on each open signal.

To calculate the investment amount, we call the *optimization* function (Algorithm 1 line 1). This function uses trading data and properties such as expected return and expected risk to determine the weights based on Equation (6). Once the optimal weights are calculated, they are returned to the *trading* function to continue position opening until the next call.

## 5. Experiment

In this section, we apply our technique on historical ETH datasets against USD, EUR, CAD, and GBP. We select three market conditions to test the applicability of the Optimal Trading Technique (cf. Algorithm 1), including a bull-run market (1 January 2021 to 1 January 2022, 365 days), a bear-run market (1 January 2022 to 1 January 2023, 365 days), and a full-cycle market (15 January 2021 to 1 October 2022, 625 days) for trading intervals 1 min, 5 min and 60 min.

### 5.1. Experimental Setup

An overarching view of our method is to pair trade fiat currencies with a cryptocurrency as an intermediate. We adopt the trading bucket from the Kraken<sup>1</sup> exchange because it provides freely accessible asset trading against multiple fiat currencies for individual researchers. Kraken is also one of the top exchanges offering the most fiat options for platform users without geographical restrictions.

We trade a bucket of fiat currency pairs referenced to ETH for several reasons: ① Cryptocurrency markets have substantial liquidity and depth; ② the volatility in the cryptocurrency market overwhelms the inherent volatility in fiat currencies, providing opportunities for arbitrage. ③ for fiat currency pairs (equivalently spreads) referenced to cryptocurrencies, ETH has a strong cointegrated and stationary relationship (cf. Table 2);

Currencies with the most consistent trading volumes available on the Kraken exchange are USD, CAD, GBP, and EUR. The period from 15 January 2022 to 1 October 2022 represents a perfect full-cycle market where ETH prices against USD climb roughly quadruple in the first half and fall back to the beginning price in the second half. For more verifications, we also examined a whole-year bull market (1 January 2021 to 1 January 2022) and a whole-year bear market (1 January 2022 to 1 January 2023).

**Table 2.** ETH correlation and cointegration among currency pairs.

Interval	1 min		5 min		60 min	
	Corr	Coint	Corr	Coint	Corr	Coint
USD:CAD	0.978	(77.8%)	0.998	(100.0%)	0.965	(89.9%)
USD:GBP	0.916	(66.7%)	0.974	(94.1%)	0.908	(72.5%)
USD:EUR	0.997	(77.8%)	0.914	(72.5%)	0.994	(69.6%)
CAD:GBP	0.917	(85.2%)	0.991	(72.5%)	0.903	(81.2%)
CAD:EUR	0.978	(88.9%)	0.980	(94.1%)	0.969	(98.6%)
GBP:EUR	0.917	(88.9%)	0.921	(98.0%)	0.913	(91.3%)

This table provides average rolling correlation calculations and cointegration test results for each currency pair between 1 January 2018 and 1 January 2020.

The datasets we use are publicly available from the Kraken OHLCVT dataset<sup>2</sup>, which includes historical trading data from Kraken exchange for open price, highest price, lowest price, close price, trading volume, and number of trades at different sampling time intervals.

### Accounting for Transaction Fees

Transaction costs in quantitative trading are always a non-negligible factor. One of the incentives to introduce optimal trading is the benefit of reducing transaction costs. The buy-and-hold strategy is least affected by transaction costs.

Kraken Exchange uses a volume-based fee scheme, where higher trading volumes result in lower transaction fees. In practice, quantitative trading typically involves large trading volumes. For our analysis, we adopt a flat 0.1% transaction fee for both long and short orders.

### 5.2. Data Source Screening

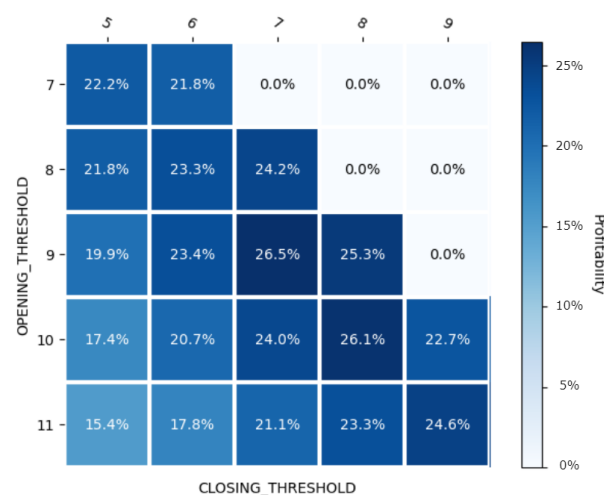
As mentioned previously (cf. Section 3.1), a preliminary requirement of the proposed OTT is that fiat currencies quoted for ETH are correlated and cointegrated. If not, it indicates that trading on the bucket of ETH/fiat pairs is unlikely to follow a mean-reversion process. We use the SPOT ETH price quoted in USD, CAD, GBP, and EUR from 1 January 2018 to 1 January 2020 for screening and formation (Algorithm 1 line 1). We test the stationarity of the pairs (spreads) not only for the complete formation period but also for shorter cycles corresponding to the time frames of our arbitrage. Therefore, we examine the data on a rolling 1-week basis to test for correlation and stationarity via cointegration.

The correlation and cointegration test results are presented in Table 2. USD:CAD under 1-week represents Pearson's correlation for USD:CAD between 1 January 2018 and 1 January 2020, divided into 105 weeks. The correlation presented is the simple weekly average for these 105 weeks. We expect our trading targets to be stationary, so we also expect the spreads we trade to be cointegrated during the period we hold the positions. Similar to Pearson's correlation test, the data are also divided into weeks for the Engle–Granger test. The passing percentages are presented in the table. We can see that the data generally have good correlations and cointegration relationships, allowing us to confidently conduct the proposed arbitrage.

### 5.3. Tuning Trading Parameters

Once the stationarity of the spreads is examined, we compute several parameters, including thresholds, which determine the boundaries to open and close positions based on the formation data (Algorithm 1 line 1).

The two parameters that determine the closing and opening boundaries are *open\_threshold* and *close\_threshold* (cf. Table 1). In our research, we conduct a grid search for the thresholds that are likely to generate high profitability. Figure 5 illustrates the grid search results for the 5min sample interval formation dataset. It is a heatmap for the indicators of profitability (calculation method in Section 5.3). A darker shade means better profit for the *open\_threshold*/*close\_threshold* combination. For the 5 min sampled dataset, opening at nine standard deviations and closing at seven standard deviations generates the best profitability. The grid search is extended to the trading intervals of 1 min and 60 min as well. Table 3 records the most lucrative combination for the trading intervals adopted in this research.



**Figure 5.** Heatmap for profitability tuning under 5min interval between 15 January 2020 and 1 October 2022.



**Table 3.** The most profitable trading thresholds.

Interval	OPEN_THRESHOLD	CLOSE_THRESHOLD
1 min	11 Standard Deviations	9 Standard Deviations
5 min	9 Standard Deviations	7 Standard Deviations
60 min	7 Standard Deviations	6 Standard Deviations

The thresholds are calculated from retrospective grid searches from 05 January 2020 to 1 October 2022 for the most lucrative combination of *open\_threshold* and *close\_threshold*.

Another tunable parameter is the risk-preference parameter  $\lambda$ , which is used to adjust the penalty factor for investment risk in Equation (2). We can tweak  $\lambda$  to cater to different risk preferences. A high  $\lambda$  represents a strong penalty factor for portfolio volatility, leading to a more conservative investment decision. By default,  $\lambda$  is set to 1, but we will also examine the profitability of arbitrage under a riskier  $\lambda = 0.5$  and a more conservative  $\lambda = 2$ .

#### Calculating P&L in Base Currency

Since we are trading with multiple fiat currencies, we need a unified method to measure profitability (or loss)—so-called P&L. We choose a base currency among the participating assets, which is USD in our case. We obtain the foreign currency exchange rates for the beginning of the arbitrage and the end. Note that our testing dataset does *not* overlap with the formation period (12 March 2016 to 31 December 2021). Hence, all our arbitrage tests are on out-of-sample (never seen before) datasets.

We start with USD, CAD, GBP, and EUR 10,000 each, totaling USD 42,340.72 on 15 January 2021. By 1 October 2022, if all the assets remain the same, their total value would be equivalent to USD 40,564.62 (Table 4). This calculation indicates a 4.19% loss due to currency exchange rate changes. However, this does not necessarily mean a 4.19% overall loss in action. Our technique tracks forex volatility as part of the trading signal. When USD is bullish, the strategy directs us to long ETH with USD against other currencies. Therefore, our profitability results from a compounded effect of cryptocurrency arbitrage and forex volatility arbitrage.

**Table 4.** Profit calculation from 15 January 2021 to 1 October 2022.

15 January 2021			1 October 2022	
USD	USD 10,000	(USD 10,000.00)	USD 12,000	(USD 12,000.00)
CAD	CAD 10,000	(USD 7896.06)	CAD 8000	(USD 5783.40)
GBP	GBP 10,000	(USD 13,186.23)	GBP 9000	(USD 10,035.00)
EUR	EUR 10,000	(USD 11,258.42)	EUR 13,000	(USD 12,746.22)
Total		USD 42,340.72		USD 40,564.62

#### 5.4. Profit and Loss Results

The proposed Optimal Trading Technique (OTT) experiments with out-of-sample datasets under different conditions to examine the validity of the technique, including full-cycle market, bull-run market, and bear-run market in this section. Further experiments with applications on BTC and SOL are included in Appendix A.

##### 5.4.1. Full-Cycle Market

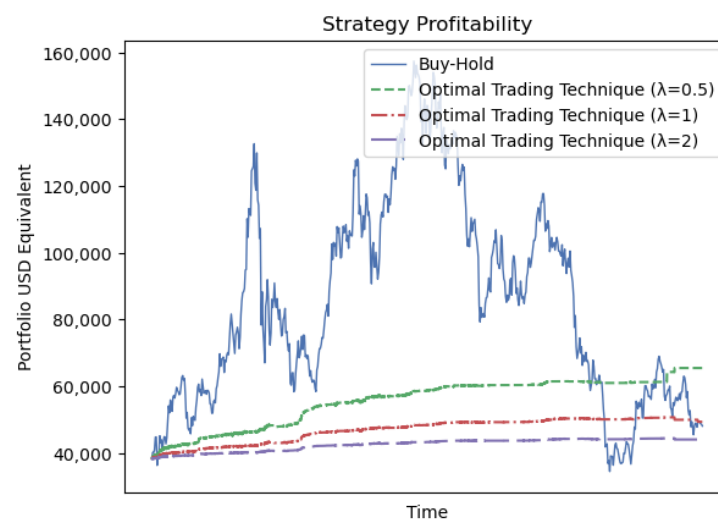
We use the method in Section 5.3 to calculate the profitability in USD, presented in Table 5. The numbers in the table for various currencies represent the profit (or loss) at the end of the trading period, starting with 10,000 units of each currency. All currencies are converted to USD based on the exchange rate at the start of the trading period on 15 January 2021. We convert the currencies back to USD for profitability calculation on 1 October 2022. During this period, cryptocurrency experienced both a bull run and a bear run, allowing us to test our technique through peak and trough. The results of our

technique are encouraging (Figure 6). We achieved an annualized return of 14.87% in USD for the 1min sampled dataset ( $\lambda = 1$ ). For the 5min and 60min sampled datasets, the annualized returns were 15.49% and 5.00%, respectively, which greatly exceed the Distance Method (DM) in all experimented cases.

**Table 5.** Annualized profitability results in full-cycle market.

Interval	BH	DM	OTT $_{\lambda=0.5}$	OTT $_{\lambda=1}$	OTT $_{\lambda=2}$
1 min (TC 0.1%)	12.91%	5.33%	35.54%	14.87%	7.44%
1 min (TC 0.0%)	13.14%	0.73%	79.98%	40.03%	19.97%
5 min (TC 0.1%)	12.91%	0.60%	37.74%	15.49%	7.74%
5 min (TC 0.0%)	13.14%	4.08%	90.03%	44.98%	22.49%
60 min (TC 0.1%)	12.91%	3.93%	9.93%	5.00%	2.51%
60 min (TC 0.0%)	13.14%	5.54%	37.43%	18.96%	9.49%

The arbitrage period is 15 January 2021 to 1 October 2022. Participating assets are USD, CAD, GBP, and EUR against ETH. The abbreviations used in the table are BH (Buy–Hold), DM (Distance Method), TC (Transaction Cost), and OTT (Optimal Trading Technique). *open\_threshold* and *close\_threshold* are tuned to the most lucrative combination.



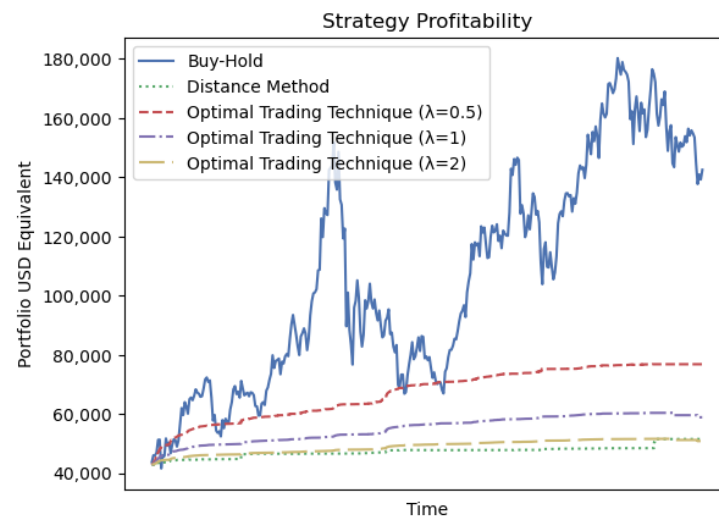
**Figure 6.** The profitability of full-cycle market between 15 January 2021 and 1 October 2022.

We also conduct experiments on conditions where investors have a higher or lower tolerance for portfolio volatility. We observe a clear relationship: the lower the  $\lambda$ , the higher the profitability. This relationship meets our expectation that investors receive a risk premium in exchange for a more volatile portfolio.

Another comparison includes transaction costs. We can see that even a 0.1% transaction cost significantly impacts our returns. On the surface, transaction cost charges a flat percentage of trading capital and influences the optimization process's expectation. Our technique trades on a frequent basis; hence, we can see the profitability is almost halved in three examined  $\lambda$  values after including transaction costs. The Distance Method, on the other hand, does not have a significant impact due to transaction cost, because it trades less frequently (Table in Section 5.5).

#### 5.4.2. Bull Market

Financial assets are subject to fluctuations, with periods of both growth and decline. However, the duration of a trading period may be one-sided. The year 2021 was particularly noteworthy for cryptocurrencies, with Ethereum (ETH) appreciating fourfold across all major fiat currencies. Although such growth is unlikely to recur frequently, it is possible to evaluate trading techniques in extreme market conditions to determine if any specific trading traits emerge.



**Figure 7.** The profitability of bull market between 1 January 2021 and 1 January 2022.

Table 6 presents the trading results for 2021. A significant growth trend is evident from the Buy–Hold strategy, which yielded a return of 412.09%. Such a circumstance is rare and occurs only in exceptional cases. Compared to full-cycle market trading, both the Distance Method (DM) and our Optimal Trading Technique (OTT) yielded more profit in the bull market (Figure 7). For a standard scenario of  $\lambda = 1$  and  $TC = 0.1\%$  over a five-minute period, the OTT achieved an annual return of 36.75% (compared to 19.89% for the DM under the same conditions). Single-sided markets are easier to estimate, and the expected\_return and expected\_risk are therefore closer to reality. The profitability trend is consistent with that of the full-cycle market. Arbitrages that do not entail transaction fees generate better returns than those with transaction fees. A lower value of  $\lambda$  implies a higher acceptance of risk, resulting in a better overall arbitrage return.

**Table 6.** Annualized profitability results in bull market.

Interval	BH	DM	OTT $_{\lambda=0.5}$	OTT $_{\lambda=1}$	OTT $_{\lambda=2}$
1 min (TC 0.1%)	412.08%	22.17%	82.13%	39.85%	19.93%
1 min (TC 0.0%)	412.90%	20.74%	170.04%	85.02%	42.51%
5 min (TC 0.1%)	412.08%	19.89%	77.78%	36.75%	18.38%
5 min (TC 0.0%)	412.90%	21.19%	165.34%	82.67%	41.34%
60 min (TC 0.1%)	412.08%	6.50%	25.23%	14.18%	6.40%
60 min (TC 0.0%)	412.90%	7.46%	45.25%	22.75%	11.37%

The arbitrage period is 1 January 2021 to 1 January 2022. Participating assets are USD, CAD, GBP, and EUR against ETH. The abbreviations used in the table are BH (Buy–Hold), DM (Distance Method), TC (Transaction Cost) and OTT (Optimal Trading Technique). *open\_threshold* and *close\_threshold* are tuned to the most lucrative combination.

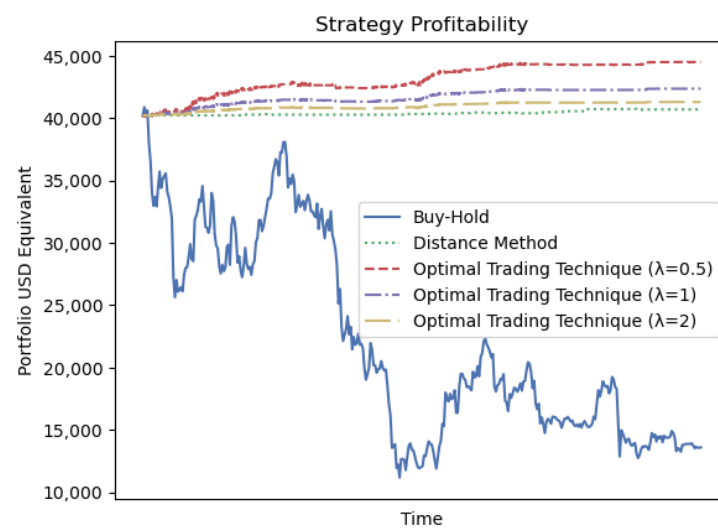
#### 5.4.3. Bear Market

We also have an interest in examining how well our technique performs in a bear market (Figure 8). Our optimization process calculates weights to avoid holding ETH in our portfolio. Therefore, in theory, the process should perform as well as in a bull market. Therefore, we experiment with the technique under the bear market in 2022, from 1 January 2022 to 31 December 2022. During this period, the price of ETH dropped by  $-61.77\%$ . We examine the profitability with data for this price dip in Table 7.

**Table 7.** Annualized profitability results in bear market

Interval	BH	DM	OTT <sub>λ=0.5</sub>	OTT <sub>λ=1</sub>	OTT <sub>λ=2</sub>
1 min (TC 0.1%)	−61.84%	1.54%	6.58%	9.57%	1.64%
1 min (TC 0.0%)	−61.77%	1.71%	19.14%	3.29%	4.78%
5 min (TC 0.1%)	−61.84%	1.86%	9.98%	4.99%	2.49%
5 min (TC 0.0%)	−61.77%	2.08%	36.41%	18.20%	9.10%
60 min (TC 0.1%)	−61.84%	5.25%	18.93%	9.01%	4.92%
60 min (TC 0.0%)	−61.77%	5.53%	33.08%	13.93%	8.52%

The arbitrage period is 1 January 2022 to 1 January 2023. Participating assets are USD, CAD, GBP, and EUR against ETH. The abbreviations used in the table are BH (Buy–Hold), DM (Distance Method), TC (Transaction Cost) and OTT (Optimal Trading Technique). *open\_threshold* and *close\_threshold* are tuned to the most lucrative combination.

**Figure 8.** The profitability of bear market between 1 January 2022 and 1 January 2023.

During the bear market in 2022, ETH witnessed a significant drop in all major currencies. The traditional Buy–Hold strategy and most non-market-neutral techniques would experience a loss from the bear market. By not holding ETH at all, both the Distance Method and our Optimal Trading Technique (OTT) achieved positive returns in all experimented cases. We can see from Table 7 that the patterns we found in a full-cycle market also apply here.

#### 5.4.4. Risk and Sharpe Ratio

Profit and risk are two sides of the same coin. We have shown that the lower the  $\lambda$  parameter, the higher the profit. Risk is measured with the Sharpe Ratio [Sharpe \(1966\)](#). The Sharpe Ratio  $S_p$  is calculated as follows:  $S_p = \frac{E[R_p] - R_f}{\sigma}$ , where  $R_p$  is the portfolio's annualized return,  $R_f$  is the annual risk-free rate<sup>3</sup>, and  $\sigma$  is the annualized standard deviation of the portfolio returns.

The results in Table 8 provide comparisons of different techniques. From the annualized standard deviation  $\sigma$ , we can see that Distance Method (DM) and Optimal Trading Technique (OTT) under all conditions have significantly lower volatility than Buy–Hold. DM is less volatile than OTT because it trades less. For OTT, there is an expected negative relation between  $\lambda$  and  $\sigma$ . A greater value of  $\lambda$  effectively reduces portfolio volatility. When measuring the overall portfolio performance with the Sharpe Ratio, lower  $\lambda$  brings higher profitability and, therefore, a higher Sharpe Ratio.

**Table 8.** Position movement for different trading strategies in markets.

(Bull)	Full-Cycle		Bull-Run		Bear-Run	
	$S_p$	( $\sigma$ )	$S_p$	( $\sigma$ )	$S_p$	( $\sigma$ )
Buy–Hold	0.12	(0.961)	1.69	(1.313)	−0.61	(0.870)
DM	−0.41	(0.082)	0.60	(0.273)	−1.51	(0.019)
OTT $_{\lambda=2}$	0.44	(0.104)	2.11	(0.070)	−0.32	(0.042)
OTT $_{\lambda=1}$	0.68	(0.193)	2.66	(0.126)	0.16	(0.082)
OTT $_{\lambda=0.5}$	1.11	(0.343)	4.34	(0.174)	0.41	(0.161)

The abbreviations used in the table are BH (Buy–Hold), DM (Distance Method) and OTT (Optimal Trading Technique).  $S_p$  is the Sharpe Ratio of the investment portfolio calculated based on a risk-free annual rate of 4%.  $\sigma$  is the standard deviation of the investment portfolio. And  $\lambda$  represents the risk control factor. The full-cycle market is from 15 January 2021 to 1 October 2022. The bull market is from 1 January 2021 to 1 January 2022. The bear market is from 1 January 2022 to 1 January 2023.

### 5.5. Strategy Indicators

To have an overview of the trading strategy, we also look into the trading metrics, not just the profitability and risk. For consistency of the counting metrics, we set the following: one arbitrage equals two trades (position open/close) equals four orders (buy/sell).

Table 9 includes some common trading strategy indicators. It consists of three sets of experiments: full-cycle, bull, and bear markets based on Section 5.4.

- *number of trades*: the total number of trades happened during the whole arbitrage period;
- *% winning*: the percentage of winning trades among all trades;
- *% losing*: the percentage of losing trades among all trades;
- *% long wins*: the percentage of winning trades among all long trades;
- *% short wins*: the percentage of winning trades among all short trades;
- *win/loss ratio*: the ratio of winning trades against losing trades;
- *average loss*: the average USD loss among losing trades;
- *average win*: the average USD win among winning trades;
- *largest loss*: the largest USD loss among losing trades;
- *largest win*: the largest USD win among winning trades;
- *avg holding hrs*: on average how many hours an arbitrage is open.

We can clearly see that our technique conducts trades much more frequently than the Distance Method. This is mainly because our spread focuses on the deviation of forex movements anchored to ETH, while the Distance Method concentrates on the distance. Our six experimental setups have a win/loss ratio greater than 1 (Table 9), forming the foundation of pair trading profitability. It is reasonable that the longer the interval, the fewer trades occur, resulting in longer average holding hours. We are also pleased to see that there are no particularly large values in the largest win or largest loss that could impact the overall revenue. This indicates the stability of our arbitrage strategy.

As for the comparison between the bull market and bear market, there is a significant decrease in the number of trades and a slight drop in the average profit (or loss) in the bear market. This indicates that although our strategy remains effective in a bear market, the quality and quantity of arbitrage opportunities are impacted.

**Table 9.** Strategy Indicators.

	Optimal Trading Technique			Distance Method		
Indicators ( <i>Full Cycle</i> )	1 min	5 min	60 min	1 min	5 min	60 min
number of trades	13,790	12,166	1906	112	56	30
% of winning	53.2%	54.4%	52.6%	63.3%	53.6%	56.7%
% of losing	46.8%	45.6%	47.3%	36.7%	46.4%	43.3%
% of long win	54.8%	55.5%	52.4%	78.6%	71.4%	73.3%
% of short win	51.6%	53.2%	53.0%	48.2%	35.7%	40.0%
win/loss ratio	1.14	1.19	1.13	1.73	1.15	1.31
average loss (US\$)	−\$6.09	−\$8.67	−\$37.31	−\$50.72	−\$174.64	−\$657.00
average win (US\$)	\$7.00	\$9.21	\$37.43	\$59.04	\$210.41	\$518.33
largest loss (US\$)	−\$347.31	−\$475.24	−\$1078.51	−\$299.56	−\$864.43	−\$2806.60
largest win (US\$)	\$386.01	\$540.01	\$976.62	\$366.06	\$862.34	\$2690.23
avg holding hrs	4.52	5.09	16.10	236.88	565.50	933.07
Indicators ( <i>Bull Run</i> )	1 min	5 min	60 min	1 min	5 min	60 min
number of trades	13,490	10,456	1620	400	290	122
% of winning	54.0%	55.2%	53.6%	62.5%	60.0%	58.2%
% of losing	46.0%	44.8%	46.4%	37.5%	40.0%	41.8%
% of long win	56.8%	57.1%	54.9%	74.0%	75.9%	70.5%
% of short win	51.3%	53.4%	52.3%	51.0%	44.1%	45.9%
win/loss ratio	1.18	1.23	1.16	1.67	1.5	1.39
average loss (US\$)	−\$6.38	−\$9.61	−\$48.33	−\$53.67	−\$78.53	−\$179.77
average win (US\$)	\$7.97	\$10.84	\$48.89	\$55.93	\$84.80	\$169.11
largest loss (US\$)	−\$335.44	−\$804.18	−\$1502.54	−\$605.46	−\$792.19	−\$1193.86
largest win (US\$)	\$359.88	\$794.02	\$1554.55	\$635.35	\$1480.52	\$1325.57
avg holding hrs	2.88	4.03	19.83	41.99	56.79	114.67
Indicators ( <i>Bear Run</i> )	1 min	5 min	60 min	1 min	5 min	60 min
number of trades	2146	3306	814	48	40	32
% of winning	52.6%	52.8%	54.3%	60.4%	55.0%	53.1%
% of losing	47.4%	47.2%	45.7%	39.6%	45.0%	46.9%
% of long win	47.2%	52.8%	60.2%	50.0%	50.0%	43.8%
% of short win	58.0%	52.9%	48.4%	70.1%	60.0%	62.5%
win/loss ratio	1.11	1.12	1.19	1.53	1.22	1.13
average loss (US\$)	−\$10.11	−\$10.40	−\$40.85	−\$85.21	−\$107.74	−\$153.77
average win (US\$)	\$10.25	\$10.51	\$39.00	\$73.48	\$109.60	\$172.19
largest loss (US\$)	−\$420.11	−\$737.76	−\$1342.21	−\$249.60	−\$382.87	−\$350.41
largest win (US\$)	\$440.61	\$736.69	\$1387.96	\$294.02	\$435.66	\$402.41
avg holding hrs	22.69	15.15	49.80	433.01	433.23	447.44

The full-cycle market is from 15 January 2021 to 1 October 2022. The bull market is from 1 January 2021 to 1 January 2022. The bear market is from 1 January 2022 to 1 January 2023. All the data are collected under  $\lambda$  as 1 and transaction cost as 0.1%

## 6. Conclusions

In this paper, we examined the feasibility and profitability of multivariate pair trading. We used fiat currencies quoted for ETH as the bucket of our pairs. We constructed the spread focusing on anomalies in forex movements due to independent trades against cryptocurrencies. The trading threshold is a tunable parameter based on the standard deviation of spreads. We demonstrated a grid search result of profitability under untuned conditions. In addition to the conventional pair trading strategy, we also introduced an optimization process for resource allocation in multivariate pair trading, ensuring market neutrality and utilizing most of the holding assets without leverage.

The results are encouraging. The profitability of the proposed Optimal Trading Technique (OTT) is on average 2.72 times<sup>4</sup> more profitable compared to the baseline approach—the Distance Method. This implies that trading spread is more effective than distance in the volatile cryptocurrency market. Under the assumption of a 4% risk-free



interest rate, the Sharpe Ratio of the OTT achieved is a strong 1.17. For all the tested cases, the OTT achieved a Win/Loss ratio greater than 1, which drives steady capital growth.

Our bi-criteria optimization objective can be tuned with the risk tolerance parameter  $\lambda$ , which impacts the optimization process for how much risk a trader tolerates. For the full-cycle market, 5min sampled dataset, the high ( $\lambda = 2$ ), intermediate ( $\lambda = 1$ ), and low ( $\lambda = 0.5$ ) risk aversion parameter values have annualized returns of 7.74%, 15.49%, and 37.74%, respectively. The corresponding Sharpe Ratios of 0.44, 0.68, and 1.11 show a negative correlation with  $\lambda$ , indicating that strong profitability compensates for the volatility brought by riskier investment choices.

Our Optimal Trading Technique produces positive returns under both bull and bear market trends. In the bull market of 2021, the OTT conducted 10,456 trades, while in the bear market of 2022, the OTT conducted only 3306 trades. Clearly, the quantity and quality of arbitrage opportunities are better in a bull market.

The applicability of the OTT is not limited to ETH. We extended our analysis to include popular cryptocurrencies BTC and SOL as intermediates in our pair trading strategy, as detailed in Appendix A. This additional analysis validates the robustness and profitability of the OTT method across different cryptocurrencies and market conditions.

Despite the promising results, several limitations should be noted. Cryptocurrency carries significant risks for investors, as highlighted by several sources. For instance, [Aiyar \(2022\)](#) likens the cryptocurrency market to the Dutch Tulip Bulb Bubble of the 17th century, where tulip bulb prices soared to extraordinary levels before collapsing dramatically. Additionally, many have documented the risks associated with cryptocurrency investments ([Angerer et al. 2021](#); [LEE et al. 2018](#)). We use cryptocurrency merely as an intermediate for pair trading fiat currencies due to its liquidity, 24/7 trading time, easy access to multiple fiat currencies in one market, and low barrier for data access. We do not promote investment in cryptocurrencies. Additionally, we selected two years of high-frequency data as our test data source, covering a full-cycle market, including a bull run and a bear run. However, this timeframe coincides with the COVID-19 period, during which fiat currencies were also volatile, contributing to spread movement. Profitability might differ in less volatile market conditions. Future work could extend the technique to model the spread from the perspective of the Ornstein–Uhlenbeck process to deepen the understanding of the market stochastic process.

**Author Contributions:** Conceptualization, A.M.; methodology, H.Y.; software, H.Y.; validation, H.Y.; formal analysis, H.Y.; investigation, H.Y.; resources, A.M. and H.Y.; data curation, H.Y.; writing—original draft preparation, H.Y.; writing—review and editing, A.M.; visualization, H.Y.; supervision, A.M.; project administration, A.M.. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original data presented in the study are openly available from Kraken Exchange at Kraken Support (<https://support.kraken.com/hc/en-us/articles/360047124832-Downloadable-historical-OHLCVT-Open-High-Low-Close-Volume-Trades-data>), accessed on 15 Apr 2024.

**Acknowledgments:** The code implementing our methodology is available in a public repository: (<https://github.com/Hongshen-Yang/optimal-trading-technique>).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A. Sensitivity Analysis

To further examine the robustness of our Optimal Trading Technique (OTT) method, we expanded the analysis to include additional cryptocurrencies and an extended time

period. We included BTC (Bitcoin) and SOL (Solana) as intermediates for pair trading based on their market capitalization. Given that SOL is not directly tradeable against CAD, we focused on USD, GBP, and EUR as the common denominators. Additionally, since the selected full-cycle, bull-run, and bear-run markets coincide with the COVID-19 period<sup>5</sup>, we extended the experiment to the post-COVID period using the latest publicly available data from Kraken up to 1 April 2024. Major cryptocurrencies have generally experienced an exponential growth, contributing to a strong Buy–Hold profitability.

**Table A1.** Correlation and cointegration among currency pairs.

Intervals	15 min		60 min		720 min	
BTC	Corr	Coint	Corr	Coint	Corr	Coint
USD:GBP	0.9530	(83.7%)	0.9504	(87.5%)	0.9480	(55.8%)
USD:EUR	0.9687	(76.0%)	0.9672	(70.2%)	0.9659	(31.7%)
GBP:EUR	0.9806	(94.2%)	0.9792	(98.1%)	0.9776	(73.1%)
ETH	Corr	Coint	Corr	Coint	Corr	Coint
USD:GBP	0.9773	(69.2%)	0.9745	(81.7%)	0.9717	(79.8%)
USD:EUR	0.9839	(56.7%)	0.9822	(70.2%)	0.9807	(37.5%)
GBP:EUR	0.9885	(88.5%)	0.9873	(95.2%)	0.9857	(94.2%)
SOL	Corr	Coint	Corr	Coint	Corr	Coint
USD:GBP	0.9896	(76.9%)	0.9892	(86.5%)	0.9865	(95.2%)
USD:EUR	0.9932	(60.6%)	0.9929	(79.8%)	0.9918	(84.6%)
GBP:EUR	0.9958	(89.4%)	0.9956	(98.1%)	0.9936	(99.0%)

This table provides average rolling correlation calculations and cointegration test results for each currency pair between 1 January 2022 and 1 January 2024.

Table A1 demonstrates a relatively strong correlation among fiat currencies and a significant level of cointegration relationships. This period is characterized by a massive bull-run market, wherein the volatility among fiat currencies is negligible compared to the substantial appreciation of cryptocurrencies. Based on these observations, we will implement our OTT techniques on all three cryptocurrencies to assess their effectiveness across different market conditions.

**Table A2.** Annualized profitability results for popular cryptocurrencies.

Interval	ETH		BTC		SOL	
	BH	OTT	BH	OTT	BH	OTT
15 min (TC 0.1%)	357.88%	22.66%	246.99%	18.24%	1856.17%	7.42%
15 min (TC 0.0%)	358.80%	44.76%	247.68%	41.92%	1860.08%	68.42%
60 min (TC 0.1%)	357.88%	19.01%	246.99%	38.78%	1856.17%	3.70%
60 min (TC 0.0%)	358.80%	36.75%	247.68%	43.48%	1860.08%	33.67%
720 min (TC 0.1%)	357.88%	28.01%	246.99%	28.21%	1856.17%	−0.80%
720 min (TC 0.0%)	358.80%	15.72%	247.68%	31.04%	1860.08%	3.26%

The arbitrage period is 1 October 2020 to 1 April 2024 (547 days). Participating assets are USD, GBP, and EUR. The abbreviations used in the table are BH (Buy–Hold), TC (Transaction Cost) and OTT (Optimal Trading Technique). Risk-adjust  $\lambda$  is set to 1. *open\_threshold* and *close\_threshold* are tuned to the most lucrative combination.

The implementation of the OTT on ETH, BTC, and SOL traded under different time intervals is presented in Table A2. Overall, the results demonstrated strong profitability not only for ETH but also for BTC. ETH and BTC, which dominate the cryptocurrency market, accounting for approximately 65% of the total market cap (Coingecko 2024), showed robust results in pair trading. While SOL achieved profitability in most trading scenarios, it did not match the performance of BTC and ETH. This is likely due to SOL's lower trading volume as it has, on average, less than two trades per minute against the GBP in the examined period and abnormal eighteen-fold annual growth. Despite the significant

bull run experienced by major cryptocurrencies, our OTT method consistently produced positive returns, further validating its effectiveness across different market conditions. We measured the risk indicators of our implementation with  $\lambda = 1$  and transaction costs:

- ETH has a standard deviation of 0.151 and a Sharpe Ratio of 1.24.
- BTC has a standard deviation of 0.159 and a Sharpe Ratio of 0.90.
- SOL has a standard deviation of 0.032 and a Sharpe Ratio of 1.07.

The volatility results display a similar property compared to the previously examined volatility in Table 8, indicating that the technique maintained desirable market neutrality under a growth market. The Sharpe Ratio of ETH and BTC denotes the ability of OTT to generate stable yet profitable actions. Similar to profitability, SOL's low profit comes with a low standard deviation as well. With SOL's exponential growth in market value, the discrepancies in fiat currencies are negligible compared to the significantly volatile growth. We consider SOL as an exceptional case as it is extremely rare for a financial instrument to grow 44 times in just 18 months.

In conclusion, extending our experiment to include the post-COVID period provided additional validation for the robustness of our OTT method. Despite the unique market conditions during the pandemic, the method maintained its effectiveness, demonstrating its adaptability and resilience across different market environments. The application of multiple cryptocurrencies demonstrates the universality of the OTT, which is adoptable for not only the ETH but other financial instruments that meet the instrument assumptions. SOL, on the other hand, constitutes an outlier example of pair trading under extreme market conditions. This adaptability is crucial for traders seeking consistent returns amidst varying market dynamics.

## Notes

- <sup>1</sup> Kraken exchange (<https://www.kraken.com/>).
- <sup>2</sup> Please refer to the Data Availability Statement for data access.
- <sup>3</sup> We use 4% based on the U.S. 10-Year Treasury Note at the time of writing.
- <sup>4</sup> The average calculation in the conclusion section is based on 1 min, 5 min, and 60 min sampled datasets for bull, bear, and full-cycle markets under a 0.1% transaction fee.
- <sup>5</sup> COVID-19, a pandemic since 2019, has had a significant global impact (Shi et al. 2020).

## References

- Aiyar, Swaminathan Anklesaria. 2022. *Crypto Craze Can Go the Same Way as Tulip Mania—Times of India*. Noida: The Times of India.
- Angerer, Martin, Christian Hugo Hoffmann, Florian Neitzert, and Sascha Kraus. 2021. Objective and subjective risks of investing into cryptocurrencies. *Finance Research Letters* 40: 101737. <https://doi.org/10.1016/j.frl.2020.101737>.
- Bai, Qianlan, Chao Zhang, Yuedong Xu, Xiaowei Chen, and Xin Wang. 2020. Evolution of Ethereum: A Temporal Graph Perspective. *arXiv*, arXiv:2001.05251.
- Brandvold, Morten, Peter Molnár, Kristian Vagstad, and Ole Christian Andreas Valstad. 2015. Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money* 36: 18–35. <https://doi.org/10.1016/j.intfin.2015.02.010>.
- Burgess, Andrew Neil. 2000. *A Computational Methodology for Modelling the Dynamics of Statistical Arbitrage*. London: University of London.
- Buterin, Vitalik. 2014. A next-generation smart contract and decentralized application platform. *White Paper* 3: 1–36.
- Cheah, Eng-Tuck, and John Fry. 2015. Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters* 130: 32–36. <https://doi.org/10.1016/j.econlet.2015.02.029>.
- Chen, Cathy YihHsuan, Wolfgang K. HHrdle, Ai Jun Hou, and Weining Wang. 2018. Pricing Cryptocurrency Options: The Case of CRIX and Bitcoin. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3159130>.
- Cohen, Israel, Yiteng Huang, Jingdong Chen, Jacob Benesty, Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*. Cham: Springer, pp. 1–4.
- Coingecko. 2024. *Crypto Market Cap Charts*. Singapore: Coingecko.
- Dale, Richard S., Johnnie E. V. Johnson, and Leilei Tang. 2005. Financial markets can go mad: Evidence of irrational behaviour during the South Sea Bubble1. *The Economic History Review* 58: 233–71. <https://doi.org/10.1111/j.1468-0289.2005.00304.x>.
- Dunis, Christian L., and Richard Ho. 2005. Cointegration portfolios of European equities for index tracking and market neutral strategies. *Journal of Asset Management* 6: 33–52. <https://doi.org/10.1057/palgrave.jam.2240164>.

- Elliott, Robert J., John Van Der Hoek, and William P. Malcolm. 2005. Pairs trading. *Quantitative Finance* 5: 271–76. <https://doi.org/10.1080/14697680500149370>.
- Engle, Robert F., and Clive W. J. Granger. 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55: 251–76. <https://doi.org/10.2307/1913236>.
- Engle, Robert F., and Clive W. J. Granger. 1991. *Long-Run Economic Relationships: Readings in Cointegration*. Oxford: OUP Catalogue, Oxford University Press. ISBN: 9780198283393.
- Fama, Eugene F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25: 383–417. <https://doi.org/10.2307/2325486>.
- Fang, Fan, Carmine Ventre, Michail Basios, Leslie Kanthan, David Martinez-Rego, Fan Wu, and Lingbo Li. 2022. Cryptocurrency trading: A comprehensive survey. *Financial Innovation* 8: 13. <https://doi.org/10.1186/s40854-021-00321-6>.
- Fil, Miroslav, and Ladislav Kristoufek. 2020. Pairs Trading in Cryptocurrency Markets. *IEEE Access* 8: 172644–51. <https://doi.org/10.1109/ACCESS.2020.3024619>.
- Galenko, Alexander, Elmira Popova, and Ivilina Popova. 2012. Trading in the Presence of Cointegration. *The Journal of Alternative Investments* 15: 85–97. <https://doi.org/10.3905/jai.2012.15.1.085>.
- Gatev, Evan, William N. Goetzmann, and K. Geert Rouwenhorst. 2006. Pairs Trading: Performance of a Relative Value Arbitrage Rule. *The Review of Financial Studies* 19: 797–827. <https://doi.org/10.2139/ssrn.141615>.
- Isard, Peter. 1977. How Far Can We Push the “Law of One Price”? *The American Economic Review* 67: 942–48.
- Kerr, David S., Karen A. Loveland, Katherine Taken Smith, and Lawrence Murphy Smith. 2023. Cryptocurrency Risks, Fraud Cases, and Financial Performance. *Risks* 11: 51. <https://doi.org/10.3390/risks11030051>.
- Lee, David Kuo Chuen, Li Guo, and Yu Wang. 2018. Cryptocurrency: A new investment opportunity? *Journal of Alternative Investments* 20: 16–40. <https://doi.org/10.3905/jai.2018.20.3.016>.
- Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7: 77–91. <https://doi.org/10.2307/2975974>.
- Miljkovic, Dragan. 1999. The Law of One Price in International Trade: A Critical Review. *Applied Economic Perspectives and Policy* 21: 126–39. <https://doi.org/10.2307/1349976>.
- Nakamoto, Satoshi. 2008. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review* 21260: 1–9.
- Pakhnenko, Olena, Pavlo Rubanov, Olga Girzheva, Larysa Ivashko, Igor Britchenko, and Liliia Kozachenko. 2023. Cryptocurrency: Value Formation Factors and Investment Risks. *Financial Engineering* 1: 1–14. <https://doi.org/10.37394/232032.2023.1.1>.
- Perlin, Marcelo. 2007. M of a Kind: A Multivariate Approach at Pairs Trading. *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.952782>.
- Shah, Devavrat, and Kang Zhang. 2014. Bayesian regression and Bitcoin. Paper presented at 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), Urbana, IL, USA, September 30–October 3; pp. 409–14. <https://doi.org/10.1109/ALLERTON.2014.7028484>.
- Sharpe, William F. 1966. Mutual Fund Performance. *The Journal of Business* 39: 119–38.
- Shi, Yu, Gang Wang, Xiao-peng Cai, Jing-wen Deng, Lin Zheng, Hai-hong Zhu, Min Zheng, Bo Yang, and Zhi Chen. 2020, May. An overview of COVID-19. *Journal of Zhejiang University. Science. B* 21: 343–60. <https://doi.org/10.1631/jzus.B2000083>.
- Slivinski, Stephen. 2009. Too Interconnected to Fail? The rescue of Long-Term Capital Management. *Federal Reserve Bank of Richmond, Region Focus*, 34: 34–36.
- van den Broek, L. 2018. Cointegration-based pairs trading framework with application to the Cryptocurrency market. Bachelor thesis, Erasmus University Rotterdam, Rotterdam, The Netherlands.
- Vidyamurthy, Ganapathy. 2004. *Pairs Trading : Quantitative Methods and Analysis*/G. Vidyamurthy. Hoboken, John Wiley & Sons.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.