QUANTITATIVE METHODS OF STATISTICAL ARBITRAGE

by

Boming Ning

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Statistics

West Lafayette, Indiana

May 2024

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Kiseop Lee, Chair

Department of Statistics

Dr. Jianxi Su

Department of Statistics

Dr. Nianqiao Ju

Department of Statistics

Dr. Mengyi Xu

Department of Statistics

Approved by:

Dr. Jun Xie

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to all those who provided me the possibility to complete this dissertation. A doctoral journey is never a solitary effort, and this dissertation stands as a testament to the support and encouragement I have received from many.

First and foremost, my utmost appreciation to Kiseop Lee, my supervisor, for his unwavering support, insightful criticisms, and invaluable advice throughout this research. His mentorship was crucial in shaping both the research and my growth as a scholar.

I am deeply thankful to the members of my dissertation committee, Jianxi Su, Mengyi Xu, and Nianqiao Ju, for their constructive feedback and essential guidance on my dissertation work.

My sincere thanks also go to Department of Statistics for providing the resources and environment conducive to my research. Special thanks to the Data Science in Finance team and all staff members for their assistance and camaraderie.

To my peers and colleagues, thank you for the stimulating discussions, and for all the fun we have had in the last few years.

I want to express my heartfelt gratitude for the constant support and patience extended by my family throughout this journey. A special thanks to my wife, Jiayu Zhang, whose understanding and boundless love have been my stronghold during my studies. My deepest appreciation also goes to my parents, who have unfailingly supported every decision I've made, providing me with the strength to pursue my dreams.

And finally, to all those who were there when needed, both within and outside the journey of my doctoral research, whose names I may not have mentioned but whose support and influence have been invaluable, thank you.

Completing this dissertation was a monumental task, and it would not have been possible without the contributions and encouragement of these individuals and many others. To everyone who was a part of my journey, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AnnMDD	annual maximum drawdown
AnnPnL	annual cumulative profit and loss
AR	autoregressive
CIR	Cox–Ingersoll–Ross
CRSP	center ror Research in security prices
DailyRet	daily return
DailySR	daily Sharpe ratio
DailyStd	daily standard deviation
DM	distance method
EMRT	empirical mean reversion time
ETF	exchange-traded fund
GAN	generative adversarial networks
MA	moving average
MDP	Markov decision process
MLE	maximum likelihood estimation
MRB	mean reversion budgeting
MRR	mean reversion ranking
MVA	mean-variance analysis
OU	Ornstein–Uhlenbeck
PCA	principal component analysis
POS	positions
RL	reinforcement learning
SD	standard deviation
SOT	signature optimal trading

ABSTRACT

Statistical arbitrage is a prevalent trading strategy which takes advantage of mean reverse property of spreads constructed from pairs or portfolios of assets. Utilizing statistical models and algorithms, statistical arbitrage exploits and capitalizes on the pricing inefficiencies between securities or within asset portfolios.

In chapter 2, We propose a framework for constructing diversified portfolios with multiple pairs trading strategies. In our approach, several pairs of co-moving assets are traded simultaneously, and capital is dynamically allocated among different pairs based on the statistical characteristics of the historical spreads. This allows us to further consider various portfolio designs and rebalancing strategies. Working with empirical data, our experiments suggest the significant benefits of diversification within our proposed framework.

In chapter 3, we explore an optimal timing strategy for the trading of price spreads exhibiting mean-reverting characteristics. A sequential optimal stopping framework is formulated to analyze the optimal timings for both entering and subsequently liquidating positions, all while considering the impact of transaction costs. Then we leverages a refined signature optimal stopping method to resolve this sequential optimal stopping problem, thereby unveiling the precise entry and exit timings that maximize gains. Our framework operates without any predefined assumptions regarding the dynamics of the underlying mean-reverting spreads, offering adaptability to diverse scenarios. Numerical results are provided to demonstrate its superior performance when comparing with conventional mean reversion trading rules.

In chapter 4, we introduce an innovative model-free and reinforcement learning based framework for statistical arbitrage. For the construction of mean reversion spreads, we establish an empirical reversion time metric and optimize asset coefficients by minimizing this empirical mean reversion time. In the trading phase, we employ a reinforcement learning framework to identify the optimal mean reversion strategy. Diverging from traditional mean reversion strategies that primarily focus on price deviations from a long-term mean, our methodology creatively constructs the state space to encapsulate the recent trends in price

movements. Additionally, the reward function is carefully tailored to reflect the unique characteristics of mean reversion trading.

1. INTRODUCTION

1.1 Arbitrage and No Free Lunch

Arbitrage is the practice of taking advantage of price differences in different markets for the same asset, thereby securing a risk-free profit. This strategy is based on the predication of one price, which states that in an efficient market, all identical assets should carry the same price. Arbitrage plays a crucial role in financial markets by helping to ensure that prices do not deviate significantly from their fair value for long periods.

The concept of 'No Free Lunch' ties closely with the idea of market efficiency, which posits that it's impossible to consistently achieve higher returns than the average market return on a risk-adjusted basis, given that all available information is already priced into stocks. This principle suggests that arbitrage opportunities are typically short-lived, as the actions of arbitrageurs will quickly correct the price discrepancies.

Different from the general arbitrage strategy, statistical arbitrage is a more sophisticated approach that involves using statistical models and algorithms to identify and exploit pricing inefficiencies between a pair of securities or within a portfolio of assets. Unlike traditional arbitrage seeking to exploit price differences for the same asset across different markets, statistical arbitrage often involves complex strategies based on mean reverting spreads constructed from pairs or multiple assets.

This dissertation concentrates on the quantitative methodologies employed in statistical arbitrage, exploring the intricate strategies and mathematical models that enable traders to identify and exploit pricing inefficiencies across various financial markets.

1.2 Statistical Arbitrage

Statistical arbitrage, utilizing statistical models and algorithms, identifies and exploits pricing inefficiencies between securities or within asset portfolios. The approach involves complex strategies based on mean reverting spreads constructed from pairs or portfolios of assets. It capitalizes on the principle that the spread prices are always expected to revert to the long-term mean, when deviating from their historical average. Thus, this strategy is also frequently referred to as mean reversion trading or pairs trading, as highlighted in various research papers.

The essence of statistical arbitrage unfolds through a meticulously structured process, involving: (1) the identification of co-moving securities, (2) the construction of spreads that reflect the mean-reverting nature, and (3) the formulation of trading strategies based on the behavior of these spreads. The first two components are referred to as the formation phase, while the third is considered the trading phase.

The initial step in statistical arbitrage strategy is the identification of similar securities, which relies on recognizing the co-movement of risk-on assets in the same direction or the correlation between stocks within a shared sector. For a trading pair to be considered profitable, it is imperative that the two assets exhibit a certain degree of correlation. This strategy entails the simultaneous purchase of one asset and sale of another, thereby achieving market neutrality within the designated group or sector, with the ultimate goal of capitalizing on the subsequent price convergence between these two assets.

After identifying groups of similar stocks, the next step involves constructing statistical arbitrage portfolios or spreads, which are essential to possess a mean-reverting characteristic, as this property underpins the opportunities available to traders engaged in statistical arbitrage. To illustrate, when the price of a spread deviates below its long-term mean, traders adopt a long position in the spread and await its reversion to the mean, thereby securing a profitable outcome.

Within the trading period, investors confront the intricate task of deciding the trading rules for both entering and exiting a position, a determination contingent upon the dynamic fluctuations in the price of the spreads. Investors are presented with the choice of either entering the market immediately or exercising patience while monitoring the prevailing market prices, awaiting a more propitious moment. After executing the initial trade, the investor encounters the consequential decision of pinpointing the most advantageous instance for closing the position. This challenge motivates the central inquiry of our researchnamely, the investigation of the optimal timing of trades.

Finally, I'd like to briefly delve into the topic of multiple pairs trading within the realm of statistical arbitrage. Current research predominantly examines the performance of individual pairs, overlooking the combined performance of several pairs. Trading multiple pairs offers significant practical advantages. Firstly, a greater number of spreads can lead to more trading opportunities at any given moment. Secondly, spreads from various stocks tend to be largely uncorrelated, providing an excellent opportunity for diversification. Lastly, engaging in multiple pairs trading introduces a novel aspect of portfolio optimization that is absent in single pair trading. This approach not only broadens the scope of statistical arbitrage but also enhances the potential for strategic portfolio management.

1.3 Data Driven Methods in Statistical Arbitrage

Examples of mean-reverting spreads are well-documented in a variety of empirical studies spanning diverse asset classes. They have been observed in pairs of stocks and ETFs [1]–[4], futures contracts [5], [6], physical commodity and commodity stocks/ETFs [7], as well as cryptocurrencies [8], [9].

The work of Gatev, Goetzmann, and Rouwenhorst [1] stands as one of the pioneering studies in the realm of mean reversion trading, often referred to as pairs trading due to their method's reliance on pairs to construct a mean-reverting process. They introduce the widely adopted Distance Method (DM) and conduct empirical testing using CRSP stocks spanning from 1962 to 2002. The DM method initiates a trading position for a pair of assets when their prices deviate from each other by more than two historical standard deviations, subsequently closing the position when the prices converge again. Their findings reveal an excess return of 1.3 % for the top 5 pairs identified by the DM method and 1.4% for the top 20 pairs. Furthermore, Do and Faff [10] delve into the profitability of pairs trading while accounting for transaction costs. Their research extends the understanding of the practicality of pairs trading by incorporating the impact of transaction expenses, offering valuable insights into the real-world feasibility of this trading strategy. For further researches in Distance Method, see [11]–[19].

In addition to the Distance Method (DM), cointegration tests are commonly employed in various alternative methods for mean reversion trading. Vidyamurthy [20] elucidate a cointegration framework for mean reversion trading, drawing inspiration from Engle and

Granger's error correction model representation of cointegrated series as presented in the seminal work of Engle and Granger [21]. Galenko, Popova, and Popova [22] explore an active ETF trading strategy built upon cointegrated time series. Their research leverages the concept of cointegration to develop trading strategies within the context of exchange-traded funds. Leung and Nguyen [8] construct cointegrated portfolios of cryptocurrencies using the Engle-Granger two-step approach and the Johansen cointegration test, exemplifying the versatility of cointegration in portfolio construction across diverse asset classes. Huck and Afawubo [18] undertake a comparative analysis of the performance of the DM method and cointegration-based approaches using the components of the S&P 500 index. This study sheds light on the relative merits of these two methodologies in the context of mean reversion trading. For more comprehensive studies on the cointegration testing methods, consider consulting the following references [22]–[27].

Another popular approach in mean reversion trading is the stochastic spread method, which characterizes the path behavior of the spread using a stochastic process exhibiting mean-reverting properties. In this methodology, the construction of spreads and the extraction of trading signals typically rely on the analysis of parameters within the underlying stochastic model. For instance, Elliott, Van Der Hoek, and Malcolm [28] introduce a mean-reverting Gaussian Markov chain model to describe the dynamics of spreads. They utilize this model's estimates in comparison with observed spread data to make informed trading decisions. Expanding on this framework, Do, Faff, and Hamza [29] delve deeper into the methodology proposed by Elliott, Van Der Hoek, and Malcolm and present a generalized stochastic residual spread method. This method is designed to model relative mispricing, offering a broader perspective on the stochastic spread approach in mean reversion trading. In their comprehensive work, Leung and Li [4] provide insights into optimal mean reversion trading strategies founded on various stochastic modeling approaches. Their research encompasses a range of models, including the Ornstein-Uhlenbeck (OU) model, Exponential OU Model, and CIR Model, elucidating how these different models can be leveraged to optimize mean reversion trading strategies. This work contributes to the broader understanding of mean reversion trading by exploring the effectiveness of diverse stochastic frameworks. Further studies on stochastic spread methods can be found in [7], [28]–[37].

Furthermore, various alternative methods for mean reversion trading have emerged in recent years. Some of these methods leverage stochastic control approach [38]–[42], copulas [43]–[47], principal component Analysis (PCA) [2], [48], and machine learning techniques [49]–[54] to identify trading opportunities based on statistical patterns and relationships among assets. In addition to these methodologies, new optimization algorithms have been proposed to construct spreads with maximum in-sample mean reversion. Notably, some of these algorithms are designed for automation, enabling the simultaneous analysis of a large number of stocks [55]–[57]. These advancements reflect the ongoing innovation in the field of mean reversion trading, offering traders and researchers a diverse set of tools and techniques to explore.

In the domain of optimal timing strategies for trading mean-reverting spreads, several notable studies have contributed to the understanding of this field. Ekström, Lindberg, and Tysk [39] investigate the optimal single liquidation timing in the context of the Ornstein-Uhlenbeck (OU) model with zero long-run mean and no transaction costs. Song, Yin, and Zhang [58] introduce a numerical stochastic approximation scheme to solve for optimal buy-low-sell-high strategies over a finite time horizon. Leung and Li [59] address an optimal double stopping problem, providing optimal entry and exit decision rules. They derive analytical solutions for both the entry and exit problems under the assumption of an OU process. These studies collectively contribute to the advancement of optimal timing strategies for trading mean-reverting spreads, offering various perspectives and methodologies for tackling this challenging problem. The current paper is motivated by the model introduced by these prior researchers, proposing a sequential optimal stopping problem to determine optimal trading times without assuming specific dynamics for the spread.

2. DIVERSIFICATION FRAMEWORK FOR MULTIPLE PAIRS TRADING

To our best knowledge, existing studies on statistical arbitrage (pairs trading) analyze the performance of a single pair, rather than the aggregate performance of multiple pairs. There are several practical benefits of trading multiple pairs and considering them together in a trading program. Firstly, more spreads may give rise to more trading opportunities at any point in time. Secondly, the spreads generated from different stocks are likely to be mostly uncorrelated, which makes it an ideal setting to take advantage of diversification. Lastly, trading multiple pairs also opens up a new direction for portfolio optimization, which does not exist when trading a single pair.

In this chapter, we propose a novel framework for constructing diversified portfolios from multiple pairs trading strategies. Capital is allocated among different pairs based on the statistical characteristics, such as the speed of mean reversion and volatility, of the historical spreads. Moreover, our approach is adaptive as portfolio weights are adjusted periodically. Among our allocation methods, we introduce the novel Mean Reversion Budgeting (MRB) and Mean Reversion Ranking (MRR) methods. The MRB method determines the portfolio weights based on the speeds of mean reversion, volatilities, and estimated average log-likelihood scores together. In contrast, the MRR method ranks spreads based on their estimated likelihood scores and speeds of mean reversion and assigns prespecified portfolio weights based on the rankings of the spreads.

Working with the empirical price data of six stock pairs, our experiment suggests that the proposed framework offers some desirable return profiles and portfolio features. We compare several allocation methods to the benchmark equal-weight portfolio and illustrate how dynamic rebalancing can improve portfolio performance.

The rest of the chapter is organized as follows. In Section 2.1, we outline the steps of portfolio construction and trading rules for OU mean reversion trading. As part of Section 2.2, we demonstrate how the volatility of spreads and speed of mean reversion affect the performance of simulated trades. In Section 2.3, we describe the proposed diversification framework. Section 2.4 describes the asset pairs included in our study, and our data collection process, and compares the performance of the proposed framework to the baseline. Concluding remarks are provided in Section 2.5.

2.1 Pairs Construction and Trading Rules

We construct a long-short position with two highly correlated assets S_t^1 and S_t^2. The portfolio value is given by

$$X_t = S_t^1 - BS_t^2,$$

where the coefficient B is called the hedge ratio and can be optimized. Following the procedure detailed in [4] by Leung and Li, we determine the optimal pair ratio through maximum likelihood estimation (MLE), whereby the resulting historical spread time series best fits the Ornstein–Uhlenbeck (OU) model.

An OU process is a mean-reverting process, described by the following stochastic differential equation:

$$dX_t = \mu(\theta - X_t)dt + \sigma dW_t, \tag{2.1}$$

where $\mu \in \mathbb{R}$ represents the speed of mean reversion, $\theta \in \mathbb{R}$ is the long-term mean, and $\sigma > 0$ is the volatility parameter. Here, W_t is a standard Brownian motion under the historical measure $\mathbb{P}$.

2.1.1 Statistical Estimation for Optimized Mean Reversion

For any given hedge ratio b, consider the observed time series of the resulting spread $X_t = S_t^1 - bS_t^2$ up to time t_0. Then, we apply the method of maximum likelihood estimation to determine the optimal parameters for the OU model based on the historical data. Specifically, given $x_i = X_{t_i}$ with time increment $\Delta t = t_i - t_{i-1}$, define the average log-likelihood function of the past L observations as

$$
\begin{aligned}
l(\mu, \theta, \sigma, b) &= \frac{1}{L} \sum_{\mathrm{i}=t_0-L+1}^{t_0-1} \log f(x_{\mathrm{i}} \mid x_{\mathrm{i}-1}; \mu, \theta, \sigma) \\
&= -\frac{1}{2} \ln(2\boldsymbol{\pi}) - \ln(\tilde{\sigma}) - \frac{1}{2L\tilde{\sigma}^2} \sum_{\mathrm{i}=t_0-L+1}^{t_0-1} [x_{\mathrm{i}} - x_{\mathrm{i}-1}\mathrm{e}^{-\mu\Delta t} - \theta(1-\mathrm{e}^{-\mu\Delta t})]^2,
\end{aligned}
$$

where

$$\tilde{\sigma}^2 = \sigma^2 \frac{1-\mathrm{e}^{-2\mu\Delta t}}{2\mu}.$$

To express the OU parameter values that maximize the average log-likelihood in (2), we define the following:

$$X_x = \sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}-1},$$

$$X_y = \sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}},$$

$$X_{xx} = \sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}-1}^2,$$

$$X_{xy} = \sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}-1} x_{\mathrm{i}},$$

$$X_{yy} = \sum_{\mathrm{i}=1}^{n} x_{\mathrm{i}-1}^2.$$

In turn, the optimal parameter estimates under the OU model are given explicitly by

$$\theta^* = \frac{X_y X_{xx} - X_x X_{xy}}{n(X_{xx} - X_{xy}) - (X_x^2 - X_x X_y)},$$

$$\mu^* = -\frac{1}{\Delta t} \ln \frac{X_{xy} - \theta^* X_x - \theta^* X_y + n(\theta^*)^2}{X_{xx} - 2\theta^* X_x + n(\theta^*)^2},$$

$$
\begin{aligned}
(\sigma^*)^2 =& \frac{2\mu^*}{n(1-\mathrm{e}^{-2\mu^*\Delta t})} (X_{yy} - 2\mathrm{e}^{-2\mu^*\Delta t} X_{xy} + \mathrm{e}^{-2\mu^*\Delta t} X_{xx} \\
&- 2\theta^*(1-\mathrm{e}^{-\mu^*\Delta t})(X_y - \mathrm{e}^{-\mu^*\Delta t} X_x) + n(\theta^*)^2(1-\mathrm{e}^{-\mu^*\Delta t})^2).
\end{aligned}
$$

At this stage, we maximize the average log-likelihood function over the three OU model parameters and denote the maximized average log-likelihood by

$$l^*(b) = l(\mu^*, \theta^*, \sigma^*, b).$$

Then, the maximized log-likelihood function $l^*(b)$ is further optimized over the ratio b to give the optimal hedge ratio

$$B = \arg\max_b \ l^*(b).$$

In practice, the implementation of this optimization can be done via a grid search. For instance, suppose we limit the absolute value of the hedge ratio to be 2. Then, we compute the maximized log-likelihood function $l^*(b)$ for

$$b = \{-2, -1.99, -1.98, \cdots, 1.99, 2\}.$$

The maximizer is denoted by B.

2.1.2 Trading Rules

We now describe the mechanism of the trading strategy in this study. Denote by S_t^1 and S_t^2 the prices of two assets at time t. Let C_t be the capital available at time t. The initial investment amount is C_0.

A number of indicators are required in order to set up the trading signals. We consider the M-day moving average of the spread X_t, denoted by $MA(X_t)$. Similarly, we define $SD(X_t)$ to be the standard deviation of the spread over the past M-days. We use the notations $POS(X_t)$ and $POS(S_t^2)$ for positions of S_t^1 and S_t^2. Finally, K is the trading threshold, r is the chosen risk level for the stop-loss rule, L is the length of the lookback window, and T is the length of the entire trading period.

In turn, the trading strategy is described as follows.

1. When entering trading period at time 0, use L-days historical prices of S_t^1 and S_t^2 to calculate the optimal pair ratio, given by

$$B = \arg\max_b l^*(b).$$

2. At each time point t, construct the spread by $X_t = S_t^1 - BS_t^2$.

3. **Entry rule**: If the current position is zero and that $X_t < MA(X_t) - K * SD(X_t)$, then enter a long position as follows:

$$POS(X_t) = C_t/S_t^1, \quad C_t = C_t - X_t \cdot POS(X_t).$$

   If the current position is zero but $X_t > MA(X_t) + K * SD(X_t)$, then enter the short position:

$$POS(X_t) = -C_t/S_t^1, \quad C_t = C_t - X_t \cdot POS(X_t).$$

4. **Exit rule**: If $POS(X_t) > 0$ and $X_t > MA(X_t)$, then quit the long position:

$$C_t = C_t + X_t \cdot POS(X_t), \quad POS(X_t) = 0,$$

   If $POS(X_t) < 0$ and $X_t < MA(X_t)$, then quit the short position:

$$C_t = C_t + X_t \cdot POS(X_t), \quad POS(X_t) = 0,$$

5. **Stop-loss rule**: Let X_0 denote the value of spread at entry time. If $POS(X_t) > 0$ and $X_t < X_0 * (1 - r)$, then quit the long position:

$$C_t = C_t + X_t \cdot POS(X_t), \quad POS(X_t) = 0.$$

   If $POS(X_t) < 0$ and $X_t > X_0 * (1 + r)$, then quit the short position:

$$C_t = C_t + X_t \cdot POS(X_t), \quad POS(X_t) = 0.$$

6. If there is no signal indicating entry or liquidation, then stay put:

$$POS(X_t) = POS(X_t), \quad C_t = C_t.$$

In Step 3, a long position of X_t is established if no shares of X_t are held and the price of X_t drops below the past M-days average value minus K times standard deviation. The number of shares purchased is determined by the ratio of current cash C_t and price of S_t^1. That is, invest all the cash at hand to long S_t^1 and short the corresponding S_t^2 with the best pair ratio B.

As for the liquidation rule described in Step 4, the entire position is liquidated whenever the price of X_t rises beyond the past M-days average value. In short, we enter the long position if the price of spread drops sufficiently far from the long-term mean, and then wait for the price to revert and take profit. The trading rule for short positions follows a similar logic. For more details on this approach, we refer to Lee and Leung (2020) and the references therein.

The trading system includes four parameters: L is the length of look-back window length that determines how many data points are used to estimate the best ratio; T determines how long we will use the estimated best ratio; M is the number of days used to calculate the moving average and standard deviation of the spread, and K is the threshold that determines the trading boundary.

Intuitively, L and M control how much past data should be incorporated in the current trading decisions. The other parameter, K, will have direct influence on the timing and frequency of trades. For instance, a smaller K will result in more frequent trading.

2.2 Monte Carlo Simulation

In this section, we analyze how different degrees of mean reversion may impact trading performance via Monte Carlo simulation. This analysis will shed light on capital allocation across different spreads in our diversification framework.

Table 2.1. The simulated one-year mean reversion trading performance under different configurations of OU model.

μ	σ	θ	**Ret (%)**	**Std (%)**	**Sharpe**	**PnL (%)**	**No.Trade**
10	1.0	15	0.0136	0.2951	0.0467	3.4067	15.410
20	1.0	15	0.0229	0.2903	0.0794	5.8520	18.092
30	1.0	15	0.0299	0.2881	0.1043	7.7184	20.992
40	1.0	15	0.0353	0.2841	0.1248	9.2014	23.563
50	1.0	15	**0.0400**	**0.2797**	**0.1431**	**10.4706**	**26.018**
30	0.1	15	0.0029	0.0286	0.1031	0.7385	20.826
30	0.5	15	0.0148	0.1433	0.1032	3.7566	20.879
30	1.0	15	0.0300	0.2872	0.1048	7.7462	20.899
30	1.5	15	0.0456	0.4295	**0.1064**	11.9390	**20.922**
30	2.0	15	**0.0600**	**0.5780**	0.1041	**15.7092**	20.698
30	1.0	5	0.0292	0.2779	0.1051	**7.9878**	20.960
30	1.0	10	0.0290	0.2821	0.1028	7.5805	**21.055**
30	1.0	15	0.0295	**0.2871**	0.1032	7.6169	20.699
30	1.0	20	**0.0304**	0.2888	**0.1053**	7.6606	20.706
30	1.0	25	0.0298	0.2834	0.1052	7.7975	20.787

We begin by considering trading a single spread, whose values are simulated according to the OU model. There are three parameters in the model. The speed of mean reversion is denoted by μ, the volatility of the process is denoted by σ, and the long-term mean is θ. We generate 1000 sample paths with different combinations of parameters and simulate the trading transactions over one year ($T = 252$) for each path by following the rule described in the previous section.

We examine the trading performance through the Sharpe ratio, average daily return, volatility, total trading numbers, and cumulative return, which are calculated based on the average over 1000 paths. The other parameters of the trading system are $K = 1$ and $M = 30$ in this experiment. Since we generate the spread time series directly, there is no need to compute the optimal pair ratio.

The results are summarized in Table 2.1. Among the parameters, we observe that the speed of mean reversion μ has a significant effect on the performance. Raising μ with σ

and θ constant, the daily returns, Sharpe ratio, and number of trades all increase while the standard deviation decreases slightly. This is intuitive since a fast mean-reverting spread offers more trading opportunities and reduces the risk associated with the trading strategy.

Next, we examine the effect of the spread volatility σ. For different σ with $\mu = 30$ and $\theta = 15$, the Sharpe ratio is almost unchanged since both daily returns and volatility increase simultaneously as σ increases. Intuitively, higher volatility may offer more trading opportunities as the spread fluctuates more rapidly. However, higher volatility also leads to a wider trading band, thus increasing the risk exposure. Lastly, the long-term mean θ does not have a significant impact on trading performance as expected, given that trades are triggered by the deviation of the spread from the long-term mean.

Based on the simulation results, the speed of mean reversion has a significantly positive effect on performance as measured by average return and Sharpe ratio. Moreover, a faster spread of mean reversion also results in more frequent trades. On the other hand, the spread volatility increases the mean and standard deviation of returns simultaneously, so the Sharpe ratio is lower for very high volatility. In contrast, the long-term mean θ does not affect the trading performance materially. Hence, we consider the mean reversion rate as a critical factor when designing the diversification framework for trading multiple pairs.

2.3 Diversification Framework

We consider a diversified portfolio approach to trading multiple pairs simultaneously. This leads to the analysis of a number of different methods to allocate portfolio weights to the traded pairs.

The first step is to divide the entire trading period into stages with a defined schedule for adjusting the allocation. During the first trading stage, we apply equal weights to trade all the spreads in the portfolio. After that, empirical spreads and returns are recorded and analyzed to yield an optimal allocation for the next trading stage. Moving forward, we will periodically obtain updates on the portfolio weights and re-allocate the capital for the next trading period accordingly.

2.3.1 Portfolio Weights

In our framework, capital is allocated across different pairs in order to form a diversified portfolio. There are various methods to determine the portfolio weights dynamically over time. To that end, we consider several methods and examine their effects on portfolio performance.

Mean-Variance Analysis

Inspired by Mean-Variance Analysis (MVA), one method for determining the portfolio weights is by maximizing the Sharpe ratio. For any given lookback period, let $R \in \mathbb{R}^N$ be the vector of average returns of each pair and $\Sigma \in \mathbb{R}^{N \times N}$ be the corresponding co-variance matrix of daily returns of each pair. Then, the optimal weights are those that give the best historical in-sample Sharpe ratio. Precisely, we have

$$\omega^* = \arg\max_{\omega} \ \frac{w^T R}{w^T \Sigma w}, \quad s.t. \quad \|w\|_1 = 1. \tag{2.2}$$

This MVA-based method assumes that the weights that have produced the best Sharpe ratio in the recent past will lead to a good performance in the next trading period. This depends on the stability of the returns over time, which may not be a reliable assumption. To address this issue, we propose several alternative allocation methods based on the characteristics of the spreads, rather than returns.

Mean Reversion Budgeting

We propose an allocation method, called Mean Reversion Budgeting (MRB), which is based on the mean-reverting properties of the spreads. In essence, we seek to measure the degree of mean reversion for each spread and assign more weights to those that are considered highly mean-reverting.

The first index is the estimated maximum average log-likelihood function. To measure the goodness of fit to an OU process for each spread, we consider the corresponding log-likelihood score. For each spread i, we denote by $\hat{l}_{\mathrm{i}}$ the estimated maximum average log-likelihood score from the previous stage. Next, we do min-max normalization to make them all positive. Precisely, for each spread i, we define the index:

$$\tilde{l}_{\mathrm{i}} = \frac{\hat{l}_{\mathrm{i}} - \min\{\hat{l}_{\mathrm{i}}\}}{\max\{\hat{l}_{\mathrm{i}}\} - \min\{\hat{l}_{\mathrm{i}}\}}.$$

In addition, we consider the speed of mean reversion μ. Intuitively, a more rapidly mean-reverting spread offers more trading opportunities, so it should be allocated with more portfolio weights. In order to account for the fluctuations due to volatility, we scale the speed of mean reversion by the volatility parameter. Precisely, we write

$$\hat{\mu}_{\mathrm{i}}^{r} = \frac{\hat{\mu}_{\mathrm{i}}}{\hat{\sigma}_{\mathrm{i}}},$$

where $\hat{\mu}_{\mathrm{i}}$ and $\hat{\sigma}_{\mathrm{i}}$ are, respectively, the estimated speed of mean reversion and volatility for each pair. In turn, comparing across all spreads, we define the normalized relative speed of mean reversion for spread i as

$$\tilde{\mu}_{\mathrm{i}} = \frac{\hat{\mu}_{\mathrm{i}}^{r} - \min\{\hat{\mu}_{\mathrm{i}}^{r}\}}{\max\{\hat{\mu}_{\mathrm{i}}^{r}\} - \min\{\hat{\mu}_{\mathrm{i}}^{r}\}}.$$

Lastly, we incorporate both $\tilde{l}_{\mathrm{i}}$ and $\tilde{\mu}_{\mathrm{i}}$ into the portfolio weight allocation. This results in the formula

$$w_{\mathrm{i}} = \frac{\tilde{\mu}_{\mathrm{i}}\tilde{l}_{\mathrm{i}}}{\sum_{\mathrm{i}=1}^{N} \tilde{\mu}_{\mathrm{i}}\tilde{l}_{\mathrm{i}}}, \ \mathrm{i} = 1, \cdots, N. \tag{2.3}$$

In summary, this means that more capital is allocated into pairs that are more OU-like and mean-revert more rapidly.

Mean Reversion Ranking

The MRB allocation method above may sometimes lead to extreme portfolio weights. For instance, if spread i has a likelihood score l_{i} that is significantly larger than the others, the method will allocate most of the capital to that pair based on (2.3), leading to portfolio concentration. Furthermore, the MRB method is susceptible to estimation errors since the portfolio weights are functions of the estimates.

These observations motivate us to propose an alternative allocation method, called Mean Reversion Ranking (MRR), which assigns fixed values based on the rankings of the product of the estimated likelihood score and speed of mean reversion. Assuming that n pairs are currently traded, we sort the n pairs in ascending order based on the product, $\tilde{l}_{\mathrm{i}}$ and $\tilde{\mu}_{\mathrm{i}}$, where $\mathrm{i} = 1, \cdots, n$. Then, we calculate the weights of sorted pairs by

$$\left(\frac{n-1}{2n(n-1)}, \frac{n+1}{2n(n-1)}, \cdots, \frac{3(n-1)}{2n(n-1)} \right). \tag{2.4}$$

This method suggests that we allocate the fraction $\frac{n-1}{2n(n-1)}$ of the capital to the pair with the lowest likelihood value and speed of mean reversion, then another fraction $\frac{n+1}{2n(n-1)}$ to the pair with the second-ranked spread, and so on.

The intuition behind this approach is as follows. If we start with equal weights and then reduce by half the weight of the pair with the lowest likelihood value and speed of mean reversion, then we invest the saved capital into the pair with the highest score. This yields the fractions $\frac{1}{2n}$ and $\frac{3(n-1)}{2n(n-1)}$ for the smallest and largest weights, respectively. The weights in the middle then increase uniformly based on their ranking indices. The ranking method offers a more moderate weight distribution than the MRB method. As our experiment shows, this method produces the highest Sharpe ratio for a given set of spreads.

2.3.2 Trading Rules under Diversification Framework

Next, we present all the stages within our trading framework.

Let $\{S_t^{11}, S_t^{12}, \ldots, S_t^{N1}, S_t^{N2}\}$ be the prices of N pairs of co-moved stocks at time t. With the initial investment C_0, we denote C_t to be the total cash at time t.

For each spread, we write $MA(X_t)$ and $SD(X_t)$ to represent the moving average and standard deviation of past M-days spread, respectively. At each time t, the position of X_t is given by $POS(X_t)$. The weight vector for all pairs is defined as $\omega = (\omega_1, \omega_2, \cdots, \omega_N)$.

The length of the whole trading period is T and the length of lookback window is L. L_R represents the window length of each small stage, also referred to as rebalancing window size. Note that the trading period is divided into T/L_R small stages. Then, there are two types of transactions within the trading program: stage transition and intra-stage trading.

The complete trading procedure is described as followed:

1. **Pairs ratio calculation**. Entering trading period at time 0, use L-days historical prices of each pair $\{S_t^{\mathrm{i}1}, S_t^{\mathrm{i}2}\}$ to calculate the optimal pair ratio, given by

$$B^{\mathrm{i}} = \arg\max_b l_{\mathrm{i}}^*(b),$$

   where $l_{\mathrm{i}}^*(\cdot)$ is the maximized average likelihood function for the i-th pair and $1 \le \mathrm{i} \le N$.

2. **Spreads construction**. At each time point t, construct N spreads by $X_t^{\mathrm{i}} = S_t^{\mathrm{i}1} - B^{\mathrm{i}} S_t^{\mathrm{i}2}$, $1 \le \mathrm{i} \le N$.

3. **Stage transition**. Entering a new stage, liquidate all the positions and use past L_R-days trading information to update the weights vector ω by equation (3), (4) or (5). Allocate the total cash at hand based on weights:

$$C_t^{\mathrm{i}} = \omega_{\mathrm{i}} \cdot C_t.$$

4. **Entry of intra-stage trading**. Set a trading threshold K (a positive constant, e.g., 0.5, 1, or 1.25). Check the entry condition for each pair. For $\mathrm{i} \in \{1, \cdots, N\}$, if $POS(X_t^{\mathrm{i}}) = 0$ and $X_t^{\mathrm{i}} < MA(X_t^{\mathrm{i}}) - K * SD(X_t^{\mathrm{i}})$, then enter the long position:

$$POS(X_t^{\mathrm{i}}) = C_t^{\mathrm{i}}/S_t^{\mathrm{i}1}, \quad C_t^{\mathrm{i}} = C_t^{\mathrm{i}} - X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}).$$

   If $POS(X_t^{\mathrm{i}}) = 0$ and $X_t^{\mathrm{i}} > MA(X_t^{\mathrm{i}}) + K * SD(X_t^{\mathrm{i}})$, then enter the short position:

$$POS(X_t^{\mathrm{i}}) = -C_t^{\mathrm{i}}/S_t^{\mathrm{i}1}, \quad C_t^{\mathrm{i}} = C_t^{\mathrm{i}} - X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}).$$

5. **Liquidation of intra-stage trading.** Check the liquidation for each pair. For $\mathrm{i} \in \{1, \cdots, N\}$, if $POS(X_t^{\mathrm{i}}) > 0$ and $X_t^{\mathrm{i}} > MA(X_t^{\mathrm{i}})$, then quit the long position:

$$C_t^{\mathrm{i}} = C_t^{\mathrm{i}} + X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}), \quad POS(X_t^{\mathrm{i}}) = 0.$$

If $POS(X_t^{\mathrm{i}}) < 0$ and $X_t^{\mathrm{i}} < MA(X_t^{\mathrm{i}})$, then quit the short position:

$$C_t^{\mathrm{i}} = C_t^{\mathrm{i}} + X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}), \quad POS(X_t^{\mathrm{i}}) = 0.$$

6. **Stop-loss rule of intra-stage trading.** Set a risk level r. For $\mathrm{i} \in 1, \cdots, N$, let X_0^{i} denote the value of each spread at entry time. If $POS(X_t^{\mathrm{i}}) > 0$ and $X_t^{\mathrm{i}} < X_0^{\mathrm{i}} * (1 - r)$, then quit the long position:

$$C_t^{\mathrm{i}} = C_t^{\mathrm{i}} + X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}), \quad POS(X_t^{\mathrm{i}}) = 0.$$

If $POS(X_t^{\mathrm{i}}) < 0$ and $X_t^{\mathrm{i}} > X_0^{\mathrm{i}} * (1 + r)$, then quit the short position:

$$C_t^{\mathrm{i}} = C_t^{\mathrm{i}} + X_t^{\mathrm{i}} \cdot POS(X_t^{\mathrm{i}}), \quad POS(X_t^{\mathrm{i}}) = 0.$$

7. **No-trade scenario.** If there is no signal of entry or liquidation, then keep the positions unchanged:

$$POS(X_t^{\mathrm{i}}) = POS(X_t^{\mathrm{i}}), \quad C_t^{\mathrm{i}} = C_t^{\mathrm{i}}.$$

8. Conduct intra-stage mean reversion trading for L_R days and go back to step 3.

The entire framework can be briefly summarized as follows: (1) divide the trading period into several smaller stages; (2) when beginning a new trading stage, update the weights based on the speed of mean reversion and volatility of spreads from the previous stage; (3) liquidate all holding positions and re-allocate the total capital based on updated weights; (4) conduct mean reversion trading on each pair separately with the allocated capital until the end of this stage; (5) record the historical spreads at the end and then move onto the next stage.

2.4 Backtesting

In this section, we examine the performance of the proposed diversified trading framework through a backtest. Different allocation methods are implemented and their results are compared. It is worth noting that we do not apply the stop-loss rule in our experiments.

2.4.1 Traded Assets and Price Data

To begin, we introduce the assets used in our experiment. Six pairs of stocks are selected from six different sectors in the US market. Each pair consists of two stocks from the same industry, ranging from airlines to banking. Table 2.2 provides a description of the companies, along with the ticker symbols of all the pairs: WM-RSG, UAL-DAL, V-MA, MS-GS, NVDA-AMD, and CVX-XOM.

Table 2.2. The stock pairs used for testing diversification framework and the description of the company.

Pairs Symbols	Description
WM-RSG	Waste Management and Republic Services provide waste management and environmental services
UAL-DAL	United and Delta are two of the largest American airlines
V-MA	Visa and Mastercard are two dominant players in the payment industry
MS-GS	Morgan Stanley and Goldman Sachs are two large-cap stocks in the banking industry
NVDA-AMD	Nvidia and AMD are two American semiconductor companies that develop computer processors
CVX-XOM	Chevron and Exxon Mobil are major companies in the energy sector

For these six pairs, we collect the daily close prices for these stocks from 1/1/2021 to 12/31/2022 using the Yahoo! Finance API. Figure 2.1 shows the price movements of each pair. Each price time series has been divided by its initial values for normalization. As we can see, the stock prices for each pair exhibit persistent comovement.

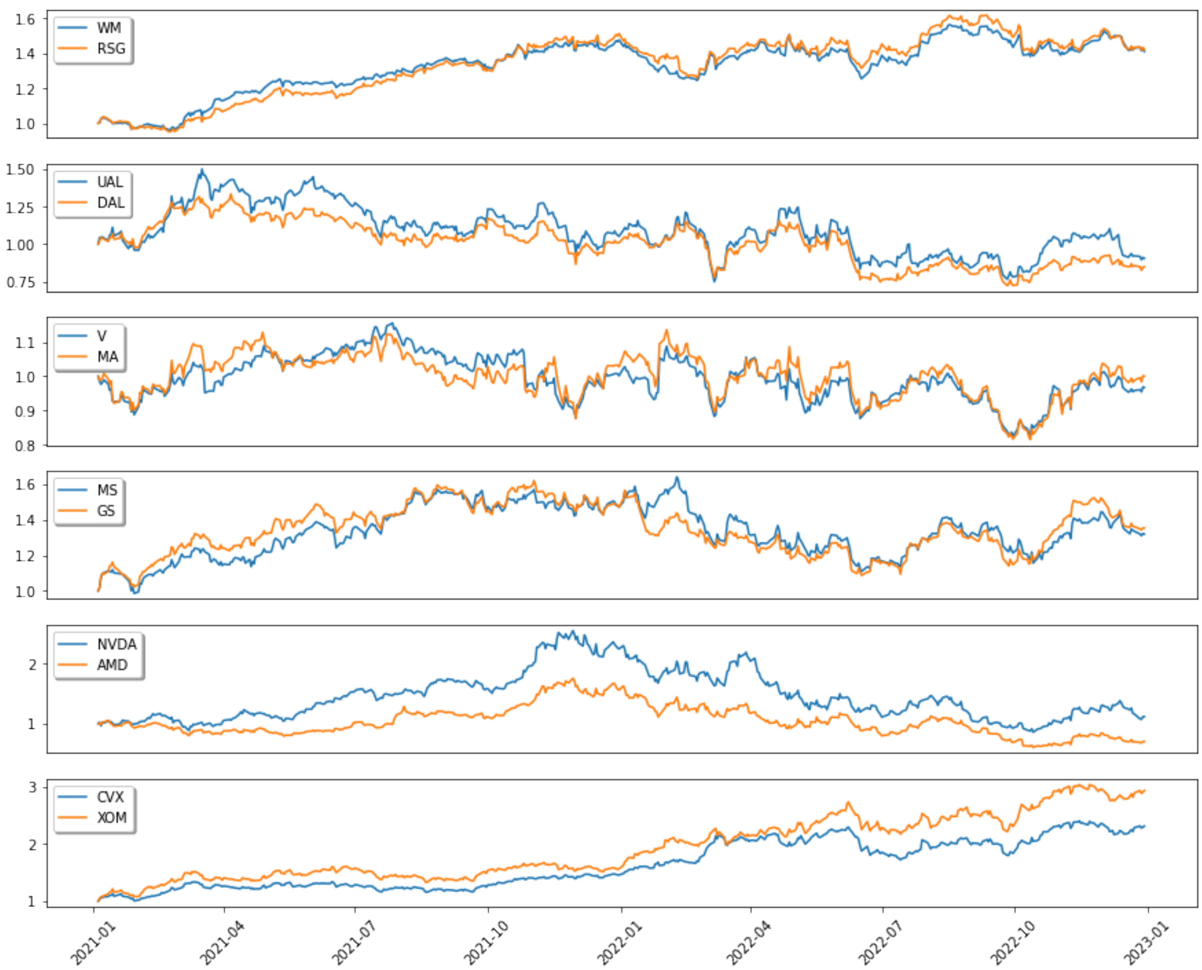


Figure 2.1. Normalized historical daily close prices of six target pairs.

Pairs are constructed using the method detailed in Section 2.2 and the daily adjusted close prices from the first year. In the experiment, the subsequent year is the trading period. In each quarter, data are used to generate new trading signals. With quarterly rebalancing, we test several trading thresholds and analyze annual performance statistics.

2.4.2 Trading on Each Pair

To demonstrate the efficacy of the proposed mean-reversion trading strategy, we first apply the strategy to six selected pairs separately. In the formation period, we use the data from 01/01/2021 to 12/31/2021 to compute the optimal ratio for each pair. Table 2.3 lists the estimated best ratio and likelihood score of each pair. The parameters are selected as $K = 1$, $L = 252$, $T = 252$ and $M = 63$.

Table 2.3. Optimal ratio and likelihood score for each pair estimated from 01/01/2021 to 12/31/2021.

S_t^1	S_t^2	Optimal Ratio b	Likelihood Score l
WM	RSG	0.98	5.7101
UAL	DAL	1.02	5.3507
V	MA	0.99	5.8283
MS	GS	1.00	6.1103
NVDA	AMD	1.31	2.8395
CVX	XOM	0.95	4.6977

Then, we construct spreads from each pair based on the best ratio. In turn, each spread is traded in the trading period. Figure 2.2 illustrates separately the spread, trading positions, and returns for the six pairs in 2022. From top to bottom in each panel, we record the movement of spread X_t, the position change $POS(X_t)$, and the cumulative return curve for each pair. As we can see, almost all of the pairs, except CVX-XOM, have positive annual profits, which demonstrates the effectiveness of the pairs trading strategy. As for the pair CVX-XOM, the main reason for failure is that the spreads constructed from the pairs do not display good mean reversion properties, as we can clearly observe a long-term downward trend in the spread. With this in mind, based on the spread behaviors and trading performance, it is intuitively optimal to allocate more weight to the first five pairs and less weight to CVX-XOM during the trading period. This echoes our motivation to propose a diversification framework to dynamically allocate weights on multiple pairs. It would be interesting to see whether any of the allocation methods can assign the weights accordingly.

As shown in Figure 2.3, the correlation coefficients between different spreads are generally small, with the most positive correlation coefficient being approximately 0.3. These low correlations suggest that there is a potential for diversification benefits in combining these spreads in a trading portfolio. By diversifying across multiple spreads, traders can potentially reduce portfolio volatility and increase the expected return. This motivates us to investigate the effectiveness of our proposed diversified trading framework.

2.4.3 Equal-Weight Portfolio

The baseline diversification is to trade all pairs equally, which means we implement the trading strategies for all pairs with equally allocated initial cash during the entire trading period. This equal-weight portfolio is conceptually the most straightforward way to create a diversified portfolio. It is included here to compare against nondiversified trading of single pairs and against other diversification methods.

To start, we first compare the performance of the equal-weight portfolio to the performance of every single pair in this section. The trading period is the year 2022, and the average annual and daily performance statistics of the trading strategies are presented in Table 2.4. We consider different trading thresholds $K \in \{0.5, 0.75, 1.0, 1.25\}$ and $L = 252$, $T = 252$ and $M = 63$. With a higher K, the spread needs to deviate further from its moving average in order to trigger a trade. The criteria presented include daily return (DailyRet), daily standard deviation (DailyStd), daily Sharpe ratio (DailySR), annual maximum drawdown (AnnMDD), and annual cumulative PnL (AnnPnL). In this experiment, we observe that the equal-weight portfolio generally outperforms single pairs in terms of daily Sharpe ratio, daily standard deviation, and annual maximum drawdown. This indicates the effectiveness of diversification for mean reversion trading.

Moreover, we observe that the standard deviations of individual pairs range from 0.4 to 1, while the standard deviation of our portfolio is lower than 0.3. This indicates that the equal-weight portfolio strategy has a lower level of risk compared to the individual pairs. This can be attributed to the diversification effect, as the portfolio strategy involves trading multiple pairs simultaneously, which helps reduce the overall risk. Therefore, our results demonstrate that an equal-weight portfolio strategy can generate a higher return and achieve a lower level of risk through diversification.

2.4.4 Diversification Framework

We implement multiple mean reversion trading under the diversification framework using different allocation methods, and compare the performance with the baseline equal-weight portfolio. Since the trading period is one year (1/1/2022–12/31/2022) and the rebalancing

Table 2.4. Performance statistics of the equal-weight portfolio and trading of single pairs.

Index	Equal Weight	WM/ RSG	UAL/ DAL	V/ MA	MS/ GS	NVDA/ AMD	CVX/ XOM
				$K = 0.5$			
DailyRet (%)	0.0260	0.0461	**0.0741**	0.0266	0.0060	0.0251	−0.0181
DailyStd (%)	**0.3020**	0.4616	0.8667	0.5882	0.5935	1.0802	1.0132
DailySR	0.0861	**0.0999**	0.0855	0.0452	0.0100	0.0233	−0.0179
AnnMDD (%)	−1.7527	−4.0295	**−1.5036**	−4.7033	−6.5959	−3.6869	−10.1602
AnnPnL (%)	6.6218	11.9691	**19.3084**	6.4325	1.0577	4.9606	−5.6825
				$K = 0.75$			
DailyRet (%)	0.0405	0.0337	**0.0996**	0.0377	0.0140	0.0435	−0.0168
DailyStd (%)	**0.2702**	0.4425	0.8010	0.5480	0.5233	1.0543	0.9531
DailySR	**0.1499**	0.0762	0.1244	0.0689	0.0267	0.0412	−0.0177
AnnMDD (%)	−1.0363	−4.0295	−1.5036	−4.7033	−4.7938	**−1.0167**	−9.4947
AnnPnL (%)	10.5989	8.5626	**27.383**	9.5192	3.2131	9.9788	−5.2313
				$K = 1$			
DailyRet (%)	**0.0444**	0.0362	0.1075	0.0442	0.0191	0.0393	0.0005
DailyStd (%)	**0.2603**	0.4406	0.7962	0.5410	0.5197	1.0602	0.9026
DailySR	**0.1707**	0.0823	0.1350	0.0821	0.0368	0.0371	0.0005
AnnMDD (%)	−0.9259	−4.0295	−1.5036	−4.7033	−3.5482	**−0.6721**	−7.5687
AnnPnL (%)	**11.7003**	9.2573	29.9381	11.3813	4.5634	8.8227	0.9027
				$K = 1.25$			
DailyRet (%)	0.0339	0.0199	**0.0646**	0.0190	0.0240	−0.0156	0.0336
DailyStd (%)	**0.2377**	0.3816	0.6833	0.4912	0.4816	0.9437	0.8126
DailySR	**0.1428**	0.0523	0.0945	0.0386	0.0499	−0.0165	0.0413
AnnMDD (%)	**−2.1982**	−4.0295	−5.3303	−3.5172	−2.3069	−13.0758	−6.6102
AnnPnL (%)	8.8155	4.9415	**16.9198**	4.5624	5.9091	−4.9065	7.8850

window is set to be three months, the trading period is divided into four periods. As before, trading parameters are as follows: $K \in \{0.5, 0.75, 1.0, 1.25\}$, $L = T = 252$ days, $M = 63$ days.

Figure 2.4 displays the cumulative return curves with different allocation methods in the year 2022. Here, "MRB" stands for the Mean Reversion Budgeting method, "MRR" represents the Mean Reversion Ranking method, and "MVA" refers to the portfolio generated by the mean-variance analysis. The equal-weight portfolio serves as a baseline portfolio for comparison.

In terms of cumulative returns, the MRB method outperforms the baseline portfolio by a significant margin. On the other hand, the MRR method's returns are only slightly higher than the baseline. This is not surprising since the rank-based method tends to spread the weights more gradually from higher to lower-ranked pairs. In other words, the MRR method does not penalize the poorly fitted spreads severely and does not assign oversized weights to the spreads with the best mean-reverting properties.

However, taking volatility into account, subsequent experiments reveal that the ranking method achieves a higher Sharpe ratio than the baseline. In contrast, the MVA portfolio lags far behind the baseline. This is perhaps intuitive since the traditional mean-variance analysis does not take into account the mean-reverting properties of the spreads.

To provide a comparative analysis of different portfolio allocation methods versus the equal-weight portfolio, we summarize the statistics of trading performance with different thresholds in Table 2.5. The statistics include daily returns (DailyRet), daily standard deviation (DailyStd), daily Sharpe ratio (DailySR), annual maximum drawdown (AnnMDD), and annual cumulative profit and loss (AnnPnL).

When compared to the baseline equal-weight portfolio, the MRB method significantly improves the annual return without increasing the standard deviation considerably. As a result, MRB yields a higher Sharpe ratio than the baseline portfolio. The MRR method significantly decreases volatility and also achieves a high Sharpe ratio. These observations support the notion that allocation based on mean reversion characteristics has the potential to improve portfolio performance.

Regarding the standard deviations of the differently weighted portfolios, it is notable that the MRR method exhibits the lowest standard deviation among all the portfolios, indicating that this particular combination of pairs is the least volatile. In addition, the MRB method also appears to decrease the standard deviation when compared to the equal weights and MVA methods. These findings provide further evidence for the advantages of employing a portfolio weights allocation approach to mitigate risk and enhance the overall performance of the portfolio.

Table 2.5. The comparison of different portfolio allocation methods vs. the baseline (equal-weight (EW)).

Index	EW	MRB	MRR	MVA
	$K = 0.5$			
DailyRet (%)	0.0260	**0.0410**	0.0289	0.0230
DailyStd (%)	0.3020	0.2755	**0.2636**	0.3624
DailySR	0.0861	**0.1488**	0.1097	0.0633
AnnMDD (%)	$-$**1.7527**	-1.9555	-1.7722	-3.5034
AnnPnL (%)	6.6218	**10.7271**	7.4317	5.7548
	$K = 0.75$			
DailyRet (%)	0.0405	**0.0431**	0.0405	0.0272
DailyStd (%)	0.2702	0.2485	**0.2349**	0.4414
DailySR	0.1499	**0.1734**	0.1723	0.0616
AnnMDD (%)	$-$**1.0363**	-1.9116	-1.3602	-3.3486
AnnPnL (%)	10.5989	**11.3333**	10.6107	6.7990
	$K = 1$			
DailyRet (%)	0.0444	**0.0479**	0.0438	0.0277
DailyStd (%)	0.2603	0.2614	**0.2310**	0.4581
DailySR	0.1707	0.1834	**0.1897**	0.0604
AnnMDD (%)	$-$**0.9259**	-1.9116	-1.3052	-3.3865
AnnPnL (%)	11.7003	**12.6890**	11.5503	6.9090
	$K = 1.25$			
DailyRet (%)	0.0339	**0.0351**	0.0328	0.0216
DailyStd (%)	0.2377	0.2402	**0.2123**	0.2618
DailySR	0.1428	0.1461	**0.1543**	0.0823
AnnMDD (%)	$-$**2.1982**	-2.4710	-2.2249	-2.5295
AnnPnL (%)	8.8155	**9.1281**	8.5101	5.4669

2.4.5 MRB Portfolio Weights

In our experiment, the MRB method appears to be the best allocation method. In this section, we provide more details on how the weights vary under the MRB method during the trading period. Let us focus on the case where $K = 1$. Table 2.6 shows the estimated parameters using the data from 1/1/2022–12/31/2022. As we can see, the NVDA-AMD and

CVX-XOM pairs have the lowest likelihood scores. Table 2.7 shows the changes in portfolio weights for each quarter during that year. Notice that the weights for the NVDA-AMD and CVX-XOM pairs are close to zero. This is consistent with the MRB method where weights are proportional to the likelihood scores.

Table 2.6. Parameters estimated from 01/01/2022 to 12/31/2022.

S_t^1	S_t^2	μ	σ	l
WM	RSG	0.1415	0.0193	5.7888
UAL	DAL	0.7284	0.0316	5.2948
V	MA	0.0108	0.0175	5.8850
MS	GS	0.0357	0.0244	5.5570
NVDA	AMD	2.7227	0.2594	3.1927
CVX	XOM	2.3461	0.3394	2.9276

Table 2.7. MRB portfolio weights for all six pairs over four quarters from 01/01/2022 to 12/31/2022.

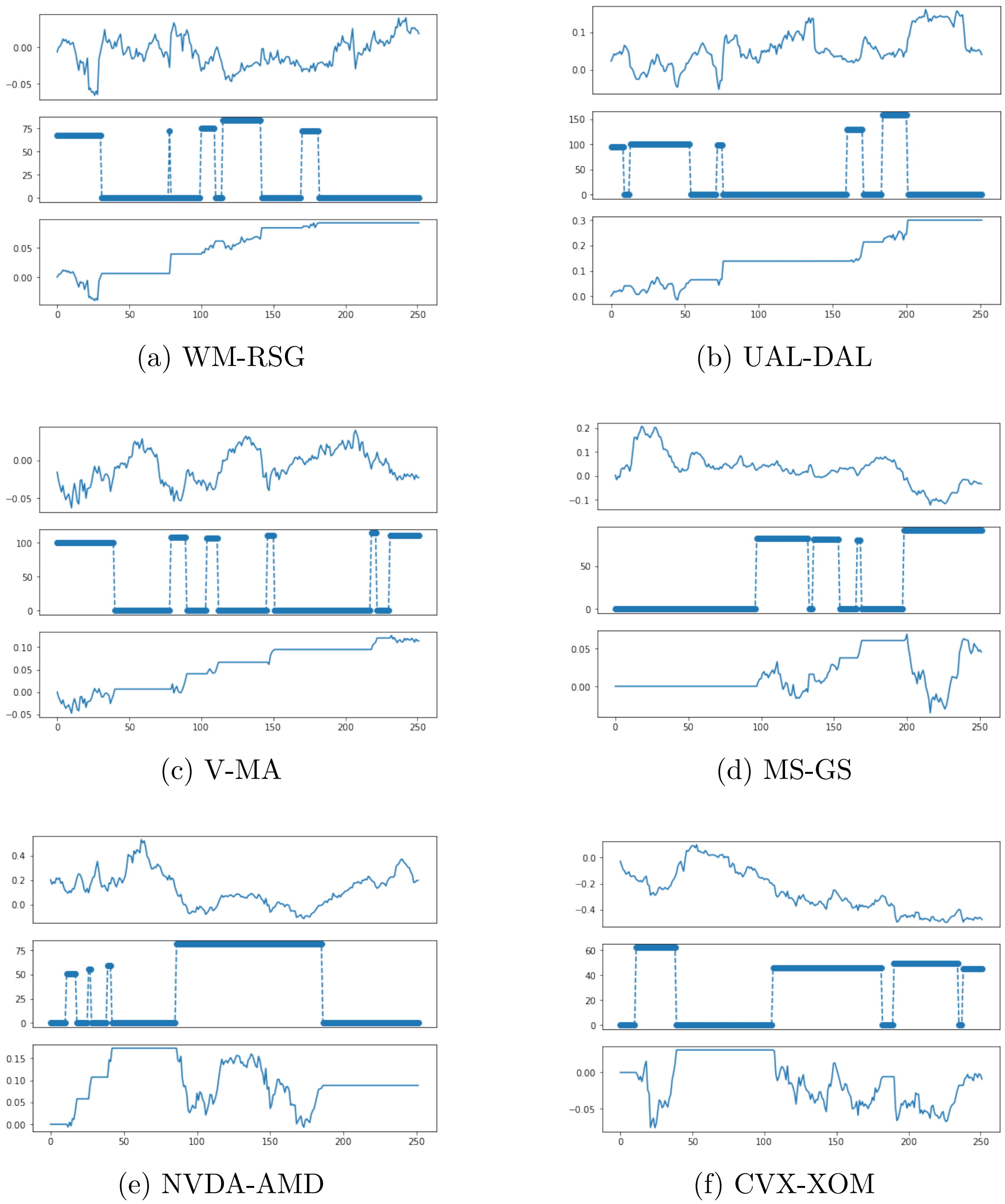
S_t^1	S_t^2	**Period 1**	**Period 2**	**Period 3**	**Period 4**
WM	RSG	0.481	0.455	0.259	0.780
UAL	DAL	0.176	0.161	0.069	0.030
V	MA	0.096	0.309	0.249	0.142
MS	GS	0.220	0.065	0.422	0.048
NVDA	AMD	0.000	0.000	0.000	0.000
CVX	XOM	0.027	0.010	0.000	0.000

2.5 Conclusions

We have presented a trading program that dynamically allocates capital to multiple mean reversion trading strategies. The approach is designed for trading multiple pairs in order to achieve diversification effects. Moreover, the dynamic rebalancing is adaptive to the current model estimates and based on the relative performance or path behaviors of the pairs in the portfolio. Our empirical experiments have shown that, for a given set of pairs traded, the allocation method plays a significant role in the success of the diversification framework.

This chapter provides portfolio managers and traders with a useful and flexible framework to test trading strategies and build portfolios involving multiple pairs. The approach discussed herein can also be applied to multiple futures spread trading that is common in other asset classes, such as interest rates, commodities, and currencies. There are plenty of examples of mean-reverting spreads between futures contracts in different markets, including Brent vs. WTI crude oil, soybean vs. soybean meal, gold vs. silver, and many more. Spread trading is very common in managed futures portfolios.

Future research directions include considering a risk-sensitive approach and incorporating additional risk controls into the trading problem. The effects on trading performance would be of practical interest. Another direction is to consider a variation of the mean-reverting model. In particular, regime-switching mean-reverting models have been used for futures trading (see [61]) and they can be suitable here for multiple spread trading.



(a) WM-RSG
(b) UAL-DAL
(c) V-MA
(d) MS-GS
(e) NVDA-AMD
(f) CVX-XOM

Figure 2.2. The separate trading performance on six pairs in 2022.

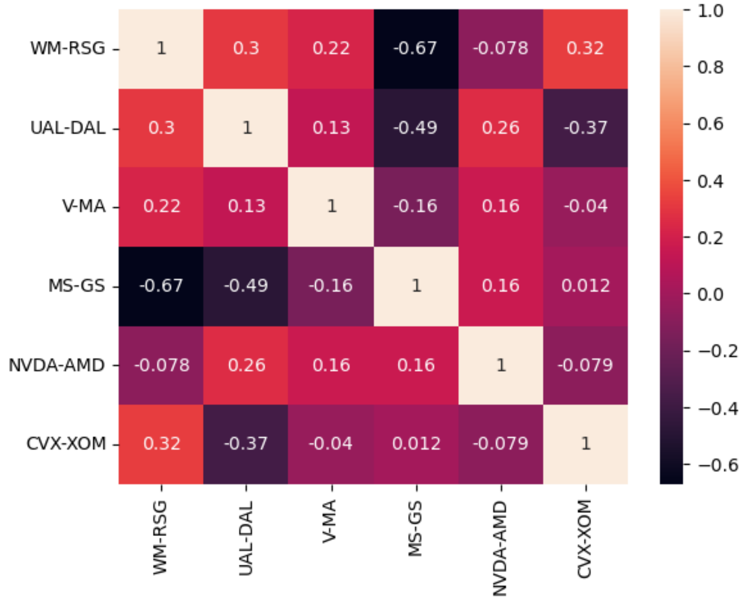
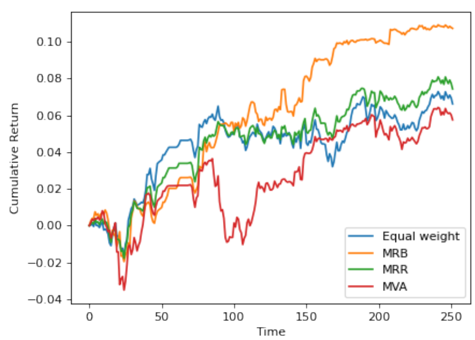
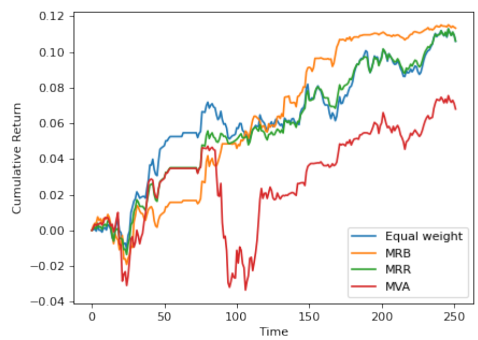


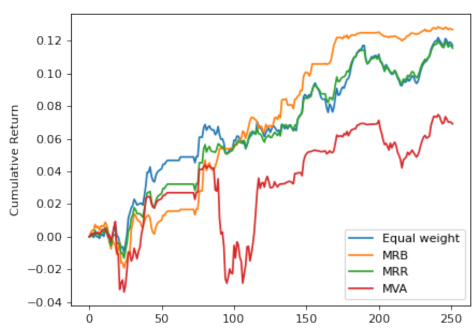
Figure 2.3. Correlation matrix of the constructed six spreads.



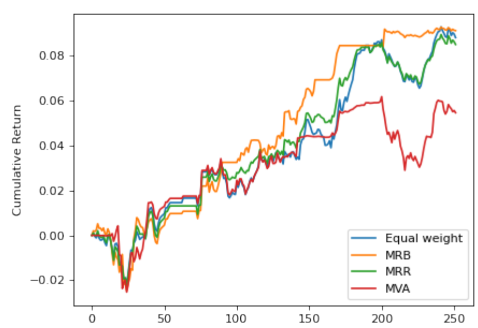
(a) k=0.5



(b) k=0.75



(c) k=1



(d) k=1.25

Figure 2.4. The accumulative returns for portfolios with different weights allocation methods.

3. OPTIMAL ENTRY AND EXIT WITH SIGNATURE IN STATISTICAL ARBITRAGE

Within the trading period of statistical arbitrage, investors confront the intricate task of deciding the optimal trading rules for both entering and exiting a position, a determination contingent upon the dynamic fluctuations in the price of the risky asset. Investors are presented with the choice of either entering the market immediately or exercising patience while monitoring the prevailing market prices, awaiting a more propitious moment. After executing the initial trade, the investor encounters the consequential decision of pinpointing the most advantageous instance for closing the position. This challenge motivates the central inquiry of our researchnamely, the investigation of the optimal timing of trades.

In response to this challenge, we introduce a optimal trading timing framework designed to systematically assess the optimal moments for both initiating and subsequently liquidating positions. Our methodology formulates a sequential optimal stopping problem to analyze the optimal trading timings, and capitalizes on the signature optimal stopping method, a powerful tool for solving this sequential optimal stopping problem, consequently revealing the precise timings for entry and exit that yield maximal gains. Importantly, our framework is adapted to diverse scenarios without any predefined assumptions regarding the dynamics of the underlying mean-reverting spreads.

In the domain of optimal timing strategies for trading mean-reverting spreads, several notable studies have contributed to the understanding of this field. Ekström, Lindberg, and Tysk [39] investigate the optimal single liquidation timing in the context of the Ornstein-Uhlenbeck (OU) model with zero long-run mean and no transaction costs. Song, Yin, and Zhang [58] introduce a numerical stochastic approximation scheme to solve for optimal buy-low-sell-high strategies over a finite time horizon. Leung and Li [59] address an optimal double stopping problem, providing optimal entry and exit decision rules. They derive analytical solutions for both the entry and exit problems under the assumption of an OU process. These studies collectively contribute to the advancement of optimal timing strategies for trading mean-reverting spreads, offering various perspectives and methodologies for tackling this challenging problem. The current paper is motivated by the model introduced by these prior researchers, proposing a sequential optimal stopping problem to determine optimal trading times without assuming specific dynamics for the spread.

The utilization of the signature method in the paper to address the optimal stopping problem aligns with recent developments in mathematical finance. Bayer, Hager, Riedel, *et al.* [62] have introduced a novel approach to solving optimal stopping problems based on signature theory. Importantly, this method does not require specific assumptions about the underlying stochastic process other than it being a continuous (geometric) random rough path. The core concept revolves around considering classic stopping times approximated as signature optimal stopping times, which can be decided through a functional of the signature associated with the stochastic process. The authors demonstrate that optimizing over these classes of signature stopping times effectively solves the original optimal stopping problem. This transformation allows for the problem to be reformulated as an optimization task dependent solely on the truncated signature. They illustrate the practicality of this method in the context of optimal stopping for fractional Brownian motions.

Indeed, motivated by the innovative methodology introduced by Bayer, Hager, Riedel, *et al.* [62], we make a modification of signature stopping framework and explore the potential application of the signature framework to tackle the optimal entry and exit problem in mean reversion trading. This extension highlights the remarkable versatility and promising capabilities of signature-based methods in effectively addressing a wide spectrum of challenges within financial contexts.

The remainder of this chapter is structured as follows. Section 3.1 lays the foundation by introducing the fundamental concepts pertaining to the signature optimal stopping method. In Section 3.2, we elaborate on our proposed framework for optimizing entry and exit decisions, presenting the associated solutions. The empirical validation of our methodology through numerical experiments, encompassing both simulated and real-world data, is detailed in Section 3.3. Finally, Section 3.4 provides concluding remarks.

3.1 Preliminaries of Signature

Signature methods play an important role in this paper. In this section, we will introduce the theory of signatures that will serve as a foundational framework for the article.

3.1.1 Tensor algebra

A signature is an infinite series that takes values in a specific graded space known as the tensor algebra. To provide a solid foundation for understanding signatures, we will begin by introducing the fundamentals of tensor algebra.

The tensor product of two vectors $\mathbf{u} = (u_{\mathrm{i}})$ and $\mathbf{v} = (v_{\mathrm{j}})$ is given by:

$$(\mathbf{u} \otimes \mathbf{v})_{\mathrm{i,j}} = u_{\mathrm{i}} v_{\mathrm{j}}$$

For example, if $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$ and $\mathbf{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$, then their tensor product is given by

$$\mathbf{u} \otimes \mathbf{v} = \begin{pmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

The concept of the tensor product can indeed be extended to accommodate multiple vectors or tensors. To illustrate, consider the example of taking the 3-way tensor product of three vectors $\mathbf{u} = [1, 2]$, $\mathbf{v} = [3, 4]$, $\mathbf{w} = [5, 6]$:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 3 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 5 \\ 6 \end{pmatrix} \\ 2 \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 5 \\ 6 \end{pmatrix} \end{pmatrix} = \left[\begin{array}{cc} 15 & 18 \\ 20 & 24 \\ \hline 30 & 36 \\ 40 & 48 \end{array}\right]$$

In this context, the resulting tensor $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ is a $2 \times 2 \times 2$ array, where the $(\mathrm{i}, \mathrm{j}, k)$-th entry is the product of the i-th entry of u, the j-th entry of v, and the k-th entry of w.

Let U and V represent two vector spaces with bases $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, respectively. The tensor product space $U \otimes V$ is defined as the vector space formed by the linear combinations of all possible products $\mathbf{u}_{\mathrm{i}} \otimes \mathbf{v}_{\mathrm{j}}$, where i ranges from 1 to m and j ranges from 1 to n. Formally, we express this as follows:

$$U \otimes V = span\{\mathbf{u}_{\mathrm{i}} \otimes \mathbf{v}_{\mathrm{j}} \mid 1 \leq \mathrm{i} \leq m, 1 \leq \mathrm{j} \leq n\}.$$

The tensor product space $U \otimes V$ possesses a natural basis $\{\mathbf{u}_{\mathrm{i}} \otimes \mathbf{v}_{\mathrm{j}} \mid 1 \leq \mathrm{i} \leq m, 1 \leq \mathrm{j} \leq n\}$. As an illustrative example, consider $U = \mathbb{R}^2$ and $V = \mathbb{R}^3$. In this case, the tensor product space $U \otimes V$ has a dimension of $2 \times 3 = 6$, and a possible basis can be represented as follows:

$$\begin{aligned}
\mathbf{u}_1 \otimes \mathbf{v}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\mathbf{u}_1 \otimes \mathbf{v}_2 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \\
&\vdots \\
\mathbf{u}_2 \otimes \mathbf{v}_3 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

The concept of tensor product for two spaces can be readily extended to multiple spaces. Let $\mathbb{R}^d$, where $d \geq 1$, be a finite-dimensional real vector space with a basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_d\}$. Then, $(\mathbb{R}^d)^{\otimes n}$ is the tensor product of n copies of $\mathbb{R}^d$, resulting in a vector space with d^n dimensions. Each element in this tensor product space is represented as a tensor formed by taking the tensor product of n vectors, each belonging to $\mathbb{R}^d$. We can define a basis for $(\mathbb{R}^d)^{\otimes n}$ as follows:

$$\{\mathbf{e}_{\mathrm{i}_1} \otimes \cdots \otimes \mathbf{e}_{\mathrm{i}_n} \mid \mathrm{i}_{\mathrm{j}} \in \{1, \cdots, d\} \text{ for } \mathrm{j} = 1, \cdots, n\}.$$

Based on $(\mathbb{R}^d)^{\otimes n}$, we can provide the definition for the extended tensor algebra.

Definition 3.1.1. *(Extended tensor algebra) The extended tensor algebra over* $\mathbb{R}^d$ *is defined by*

$$T((\mathbb{R}^d)) := \{\boldsymbol{a} = (a_n)_{n=0}^{\infty} \;\mid\; a_n \in (\mathbb{R}^d)^{\otimes n}\}.$$

Given $\boldsymbol{a} = (a_{\mathrm{i}})_{\mathrm{i}=0}^{\infty}$, $\boldsymbol{b} = (b_{\mathrm{i}})_{\mathrm{i}=0}^{\infty} \in T((\mathbb{R}^d))$, *define the sum and product by*

$$\boldsymbol{a} + \boldsymbol{b} := (a_{\mathrm{i}} + b_{\mathrm{i}})_{\mathrm{i}=0}^{\infty}$$

$$\boldsymbol{a} \otimes \boldsymbol{b} := \left(\sum_{k=0}^{\mathrm{i}} a_k \otimes b_{\mathrm{i}-k}\right)_{\mathrm{i}=0}^{\infty}.$$

We also define the action on $\mathbb{R}$ *given by* $\lambda\boldsymbol{a} := (\lambda a_{\mathrm{i}})_{\mathrm{i}=0}^{\infty}$ *for* $\lambda \in \mathbb{R}$.

Likewise, we can define the tensor algebra as the space of all finite sequences of tensors. Additionally, we can define the truncated tensor algebra as the space of sequences of tensors of a specified, fixed length. These definitions allow us to work with sequences of tensors, accommodating various mathematical operations and analyses in this paper.

Definition 3.1.2. *(Tensor algebra) The tensor algebra over* $\mathbb{R}^d$, *denoted by* $T(\mathbb{R}^d) \subset T((\mathbb{R}^d))$, *is given by*

$$T(\mathbb{R}^d) := \{\boldsymbol{a} = (a_n)_{n=0}^{\infty} \;\mid\; a_n \in (\mathbb{R}^d)^{\otimes n} \text{ and } \exists N \in \mathbb{N} \text{ such that } a_n = 0 \;\forall n \geq N\}.$$

Definition 3.1.3. *(Truncated tensor algebra) The truncated tensor algebra of order* $n \in \mathbb{N}$ *over* $\mathbb{R}^d$ *is defined by*

$$T^{(N)}(\mathbb{R}^d) := \{\boldsymbol{a} = (a_n)_{n=0}^{\infty} \;\mid\; a_n \in (\mathbb{R}^d)^{\otimes n} \text{ and } a_n = 0 \;\forall n \geq N\}.$$

In the context of the dual space and linear functions on the tensor algebra, let $\{\mathbf{e}_1^*, \cdots, \mathbf{e}_d^*\}$ be a dual basis for $(\mathbb{R}^d)^*$, where $(\mathbb{R}^d)^*$ represents the dual space of $\mathbb{R}^d$. This dual space consists of all continuous linear functions mapping from $\mathbb{R}^d$ to $\mathbb{R}$. We can define a basis for $((\mathbb{R}^d)^*)^{\otimes n}$ as follows:

$$\{\mathbf{e}_{\mathrm{i}_1}^* \otimes \cdots \otimes \mathbf{e}_{\mathrm{i}_n}^* \;\mid\; \mathrm{i}_{\mathrm{j}} \in \{1, \cdots, d\} \text{ for } \mathrm{j} = 1, \cdots, n\}.$$

We denote the dual space of $T((\mathbb{R}^d))$ as $T((\mathbb{R}^d)^*) := ((\mathbb{R}^d)^*)^{\otimes n}$.

It's important to note that there exists a natural pairing between $T((\mathbb{R}^d)^*)$ and $T((\mathbb{R}^d))$, which we denote as $\langle \cdot, \cdot \rangle : T((\mathbb{R}^d)^*) \times T((\mathbb{R}^d)) \to \mathbb{R}$. Consequently, the dual space $T((\mathbb{R}^d)^*)$ encompasses all continuous linear functions defined on the extended tensor algebra $T((\mathbb{R}^d))$.

3.1.2 Signature

We will now introduce a pivotal concept in this paper: the signature of a path. For a given path $X : [0, T] \mapsto \mathbb{R}^d$, we denote the coordinate paths as $(X_t^1, \cdots, X_t^d)$, where each $X^{\mathrm{i}} : [0, T] \mapsto \mathbb{R}$ represents a real-valued path. For any single index $\mathrm{i} \in \{1, \cdots, d\}$ and $0 \leq a \leq b \leq T$, we define the quantity:

$$S(X)_{a,b}^{\mathrm{i}} = \int_{a<s<b} dX_s^{\mathrm{i}} = X_b^{\mathrm{i}} - X_a^{\mathrm{i}},$$

which signifies the increment of the i-th coordinate of the path within the time interval $t \in [a, b]$.

For any pair of indices $\mathrm{i}, \mathrm{j} \in \{1, \cdots, d\}$, we define the second-order path integral:

$$S(X)_{a,b}^{\mathrm{i},\mathrm{j}} = \int_{a<r<s<b} dX_r^{\mathrm{i}} dX_s^{\mathrm{j}},$$

and this concept can be generalized to arbitrary orders. For any integer $n \geq 1$ and a collection of indices $\mathrm{i}_1, \cdots, \mathrm{i}_n \in \{1, \cdots, d\}$, we define:

$$S(X)_{a,b}^{\mathrm{i}_1,\cdots,\mathrm{i}_n} = \int_{a<t_1<\cdots<t_n<b} dX_{t_1}^{\mathrm{i}_1} \cdots dX_{t_n}^{\mathrm{i}_n},$$

referred to as the n-fold iterated integral of X along the indices $\mathrm{i}_1, \cdots, \mathrm{i}_n$. The vector comprising all the n-fold iterated integrals of X is defined as:

$$\mathbb{X}_{a,b}^n := \left(S(X)_{a,b}^{\mathrm{i}_1,\cdots,\mathrm{i}_n} \,\middle|\, \mathrm{i}_1, \cdots, \mathrm{i}_n \in \{1, \cdots, d\} \right).$$

In fact, $\mathbb{X}^n_{a,b}$ belongs to $(\mathbb{R}^d)^{\otimes n}$, and we can express it concisely using the tensor product symbol:

$$\mathbb{X}^n_{a,b} = \int_{a<t_1<\cdots<t_n<b} dX_{t_1} \otimes \cdots \otimes dX_{t_n} \in (\mathbb{R}^d)^{\otimes n}.$$

To illustrate, I shall provide a specific example of n-fold iterated integrals. Consider a path $X : [0,T] \mapsto \mathbb{R}^3$, then

$$\mathbb{X}^1_{a,b} = \left(S(X)^1_{a,b}, S(X)^2_{a,b}, S(X)^3_{a,b}\right),$$

$$\mathbb{X}^2_{a,b} = \left(S(X)^{1,1}_{a,b}, S(X)^{1,2}_{a,b}, S(X)^{1,3}_{a,b}, S(X)^{2,1}_{a,b}, S(X)^{2,2}_{a,b}, S(X)^{2,3}_{a,b}, S(X)^{3,1}_{a,b}, S(X)^{3,2}_{a,b}, S(X)^{3,3}_{a,b}\right).$$

Now, we can introduce the definition of a signature.

Definition 3.1.4. *(Signature) Let $0 \leq a < b \leq T$. For a path $X : [0,T] \to \mathbb{R}^d$, we define the signature of X over $[a,b]$ by*

$$\mathbb{X}^{<\infty}_{a,b} := (1, \mathbb{X}^1_{a,b}, \cdots, \mathbb{X}^n_{a,b}, \cdots) \in T((\mathbb{R}^d)),$$

where

$$\mathbb{X}^n_{a,b} := \int_{a<t_1<\cdots<t_n<b} dX_{t_1} \otimes \cdots \otimes dX_{t_n} \in (\mathbb{R}^d)^{\otimes n}.$$

Similarly, the truncated signature of order N is defined by

$$\mathbb{X}^{\leq N}_{a,b} := (1, \mathbb{X}^1_{a,b}, \cdots, \mathbb{X}^N_{a,b}) \in T^{(N)}((\mathbb{R}^d)).$$

If we refer to the signature of X without specifying the interval over which the signature is taken, we will implicitly refer to $\mathbb{X}^{<\infty}_{0,T}$, i.e., $\mathbb{X} = \mathbb{X}^{<\infty}_{0,T}$.

In practical applications of the signature method, we typically transform the original path into an augmented path and then compute the truncated signature of the augmented path as the input features for any algorithm. Specifically, given a d-dimensional path $X_t : [0,T] \to \mathbb{R}^d$, we denote its augmentation by $\widehat{X_t} = (t, X_t) \in \mathbb{R}^{1+d}$, and define the augmented signature $\widehat{\mathbb{X}}^{\leq\infty}$ and $\widehat{\mathbb{X}}^{\leq\lfloor p\rfloor}$ as the signature and $\lfloor p\rfloor$-truncated signature of $\widehat{X_t}$, respectively.

Due to the first dimension of the augmented path (time) being monotonically increasing, the signature $\widehat{\mathbb{X}}^{\leq\infty}$ provides a complete characterization of $\widehat{X}$ (and therefore X). See [63]. However, computing the entire signature with infinite length is not feasible, so the truncated signature with a fixed order p is used in the practice of the signature method. As a result, the time series to be analyzed in this paper will first be transformed to its truncated augmented signature for analysis.

3.1.3 Signature Optimal Stopping Method

In this section, we introduce the signature optimal stopping method proposed by [62], which will be applied in our approach to solve the proposed sequential signature optimal stopping problems.

Signature Optimal Stopping Time

The general optimal stopping problem involves determining an optimal stopping time that maximizes the value of the underlying process. Fix the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose X_t is a stochastic process and denote the filtration generated by X as $(\mathcal{F}_t)$, $\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t)$. Consider a payoff process $Y : [0,T] \times \Omega \to \mathbb{R}$ being a real-valued continuous stochastic process adapted to the filtration $(\mathcal{F}_t)$, then the optimal stopping problem for Y_t is

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}[Y_{\tau \wedge T}], \tag{3.1}$$

where τ is a stopping time adapted to $(\mathcal{F}_t)$ and $\mathcal{T}$ is the space of all $(\mathcal{F}_t)$-stopping times.

Similarly, the signature optimal stopping problem also involves determining an optimal stopping time to maximize the expected payoff of a stochastic process before a fixed time. However, the key difference between the two problems lies in the filtration used to define stopping times, which fundamentally affects the information available for making stopping decisions. Specifically, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Given a d-dimensional path $X_t : [0,T] \to \mathbb{R}^d$ and $\widehat{\mathbb{X}}$ denotes its augmented signature in $\widehat{\Omega}^p_T$. The filtration generated by the signatures are defined as $(\mathcal{G}_t)$, $\mathcal{G}_t = \sigma(\widehat{\mathbb{X}}_{0,s} : 0 \leq s \leq t)$ and denote by $\mathcal{S}$ the space of all $(\mathcal{G}_t)$-stopping times. $\mathcal{S}$ is referred as the space of signature stopping times. Consider a payoff process $Y : [0,T] \times \Omega \to \mathbb{R}$ being a real-valued continuous stochastic process adapted to the filtration $(\mathcal{G}_t)$, then the signature optimal stopping problem for Y_t is:

$$\sup_{\tau \in \mathcal{S}} \mathbb{E}[Y_{\tau \wedge T}] \tag{3.2}$$

where τ is a signature stopping time adapted to $(\mathcal{G}_t)$.

In summary, the general optimal stopping problem utilizes $(\mathcal{F}_t)$, a filtration based on the observable trajectory of the stochastic process X_t, to determine stopping times τ using data up to time t. Conversely, the signature optimal stopping problem employs $(\mathcal{G}_t)$, generated from X_t's augmented signatures $\widehat{\mathbb{X}}$, which encapsulate both positions and dynamic patterns over time, offering richer information for stopping decisions. This enriched filtration potentially provides a more detailed insight into X_t's behavior, guiding the derivation of stopping times.

Randomized Stopping Time

The fundamental concept for addressing the signature optimal stopping problem lies in the utilization of randomized stopping times. As demonstrated by Bayer, Hager, Riedel, *et al.* [62], randomized stopping times are, in fact, equivalent to signature stopping times. Consequently, the signature optimal stopping problem, as depicted in Equation (3.2), can be effectively transformed into an equivalent problem known as the randomized optimal stopping problem.

Firstly, we introduce the concept of randomized stopping time. Define $\widehat{\Omega}_T^p$ as the space of all the p-truncated augmented signatures, that is, elements of $\widehat{\Omega}_T^p$ are p-truncated signatures of some augmented paths (t, X_t). For a given time horizon $T > 0$, we define $\Lambda_T := \bigcup_{t \in [0,T]} \widehat{\Omega}_t^p$, referring to it as the space of stopped signatures. Then, we denote a continuous stopping policy θ as a Borel measurable mapping $\theta : \Lambda_T \to \mathbb{R}$, and the space of continuous stopping policies as $\Theta := C(\Lambda_T, \mathbb{R})$. θ is called a stopping policy because it can be employed to define the randomized stopping time based on its mapping value.

Definition 3.1.5. *(Randomized stopping time) Let Z be a non-negative random variable independent of augmented signature $\widehat{\mathbb{X}}$ and such that $P(Z = 0) = 0$. For a continuous stopping policy $\theta \in \Theta$, we define the randomized stopping time by*

$$\tau_\theta^r := \inf\left\{t \geq 0 : \int_0^{t\wedge T} \theta(\widehat{\mathbb{X}} \mid_{[0,s]})^2 ds \geq Z\right\} \tag{3.3}$$

where $\inf \phi = +\infty$.

Randomized stopping times, which are derived from continuous stopping policies, possess the capability to approximate signature stopping times, as demonstrated in the following theorem [62, Proposition 4.2].

Theorem 3.1.1. *For every stopping time $\tau \in \mathcal{S}$, there exists a sequence $\theta_n \in \Theta$ such that the randomized stopping times $\tau_{\theta_n}^r$ satisfy $\tau_{\theta_n}^r \to \tau$ almost surely as $n \to \infty$. In particular, if $\mathbb{E}[\|Y\|_\infty] < \infty$, then*

$$\sup_{\theta\in\Theta} \mathbb{E}[Y_{\tau_\theta^r \wedge T}] = \sup_{\tau\in\mathcal{S}} \mathbb{E}[Y_{\tau\wedge T}]. \tag{3.4}$$

An astonishing result is that the general stopping time can be approximated by using stopping policies that are linear functions of the signature. The space of linear signature stopping policies $\Theta_l \subset \Theta$ is defined as

$$\Theta_l = \{\theta \in \Theta : \exists l \in T((\mathbb{R}^{1+d})^*) \text{ such that } \theta(\widehat{\mathbb{X}} \mid_{[0,t]}) = \langle l, \widehat{\mathbb{X}}_{0,t}^{<\infty}\rangle \ \forall \widehat{\mathbb{X}} \mid_{[0,t]} \in \Lambda_T\}.$$

Note that every $l \in T((\mathbb{R}^{1+d})^*)$ defines a stopping policy $\theta_l \in \Theta$ by setting $\theta(\widehat{\mathbb{X}} \mid_{[0,t]}) = \langle l, \widehat{\mathbb{X}}_{0,t}^{<\infty}\rangle$. Then we introduce the following notation for randomized stopping times associated to linear signature stopping policies

$$\tau_l^r := \tau_{\theta_l}^r = \inf\left\{t \geq 0 : \int_0^{t\wedge T} \langle l, \widehat{\mathbb{X}}_{0,s}^{<\infty}\rangle^2 ds \geq Z\right\}. \tag{3.5}$$

The subsequent theorem, detailed in [62, Proposition 5.3], demonstrates that the randomized stopping times linked to linear signature stopping policies are able to approximate general randomized stopping times.

Theorem 3.1.2. *Assume that Z has a continuous density and $\mathbb{E}[\|Y\|_\infty] < \infty$. Then*

$$\sup_{\theta\in\Theta} \mathbb{E}[Y_{\tau^r_\theta \wedge T}] = \sup_{\theta\in\Theta_l} \mathbb{E}[Y_{\tau^r_\theta \wedge T}].$$

It follows that

$$\sup_{\theta\in\Theta} \mathbb{E}[Y_{\tau^r_\theta \wedge T}] = \sup_{l\in T((\mathbb{R}^{1+d})^*)} \mathbb{E}[Y_{\tau^r_l \wedge T}]. \tag{3.6}$$

Continuing our exploration, we turn our attention to the evaluation of the optimal stopping problem with regard to a specific stopping policy θ. This evaluation plays a crucial role as it serves as a pivotal loss function, which will be utilized in the process of identifying the optimal linear policy during training. [62, Proposition 4.4] offers comprehensive details into this matter.

Theorem 3.1.3. *Let F_Z denote the cumulative distribution function of Z. Then*

$$\mathbb{E}[Y_{\tau^r_\theta \wedge T} \mid \widehat{\mathbb{X}}] = \int_0^T Y_t d\widetilde{F}(t) + Y_T(1-\widetilde{F}(T)) = Y_0 + \int_0^T (1-\widetilde{F}(t))dY_t \tag{3.7}$$

where the second integral is implicitly defined by integration by parts and

$$\widetilde{F}(t) := F_Z\left(\int_0^t \theta(\widehat{\mathbb{X}}\mid_{[0,s]})^2 ds\right).$$

In particular, if Z has a density f_Z,

$$\mathbb{E}[Y_{\tau^r_\theta \wedge T}] = \mathbb{E}\left[\int_0^T Y_t\ \theta(\widehat{\mathbb{X}}\mid_{[0,t]})^2\ f_Z\left(\int_0^t \theta(\widehat{\mathbb{X}}\mid_{[0,s]})^2 ds\right) dt + Y_T(1-\widetilde{F}(T))\right]. \tag{3.8}$$

Implementation of Signature Optimal Stopping Method

In this section, we detail the implementation of solving the optimal signature stopping problem, utilizing the previously introduced theorems.

The initial step involves discretizing the paths, where we establish a time grid comprising equidistant points denoted as $\{t_\mathrm{j}\}_{\mathrm{j}=0,1,\dots,n}$, with $t_\mathrm{j} = \frac{\mathrm{j}}{n}T$. After discretization, we compute the signature of the time-augmented path $(t_\mathrm{j}, X_{t_\mathrm{j}})_{0\le \mathrm{j}\le n}$, denoted by $\widehat{\mathbb{X}}_{0,t_\mathrm{j}}$. Then, we adapt the definition of randomized stopping times to the discrete-time setting. Given a positive random variable Z and a continuous stopping rule $\theta \in \Theta$, we define the discrete version of the randomized stopping time in (3.3) as follows:

$$\nu_\theta^r := \inf \left\{ 0 \leq \mathrm{j} \leq n : \sum_{\mathrm{i}=0}^{\mathrm{j}} \theta(\widehat{\mathbb{X}} \mid_{[0,t_{\mathrm{i}}]})^2 \geq Z \right\} \tag{3.9}$$

The discrete version of randomized stopping times associated with linear signature stopping policies is defined as:

$$\nu_l^r := \inf \left\{ 0 \leq \mathrm{j} \leq n : \sum_{\mathrm{i}=0}^{\mathrm{j}} \langle l, \widehat{\mathbb{X}}_{0,t_{\mathrm{i}}}^{<\infty} \rangle^2 \geq Z \right\} \tag{3.10}$$

The similar results established in Theorems 3.1.1 to 3.1.3 also extend to discrete stopping times. Building upon Theorem 3.1.3, we can represent the expected payoff in (3.7) as follows:

$$\mathbb{E}[Y_{\nu_\theta^r} \mid \widehat{\mathbb{X}}] = Y_0 + \sum_{\mathrm{j}=0}^{n-1} G_Z \left(\sum_{\mathrm{i}=0}^{\mathrm{j}} \theta(\widehat{\mathbb{X}} \mid_{[0,t_{\mathrm{i}}]})^2 \right) (Y_{\mathrm{j}+1} - Y_{\mathrm{j}}), \tag{3.11}$$

where $G_Z = 1 - F_Z$.

We can further elaborate on the loss function for linear stopping policies based on the regular form of the expected payoff provided above. Given a data-set comprising M samples in the format $(X_{\mathrm{j}}^{(m)}, Y_{\mathrm{j}}^{(m)})_{1 \leq m \leq M, 0 \leq \mathrm{j} \leq n}$ and a predetermined truncation level $N \geq 1$, we compute the N-truncated augmented signature of the path on each sub-interval $[0, t_{\mathrm{i}}]$ for each series of samples, define as $\widehat{\mathbb{X}}_{0,t_{\mathrm{i}}}^{(m)}$. Subsequently, we establish the following loss function for the linear stopping policy θ_l based on Equation (3.11):

$$\mathrm{loss}(l) = -\frac{1}{M} \sum_{\mathrm{i}=1}^{M} \left\{ Y_0^{(m)} + \sum_{\mathrm{j}=0}^{n-1} G_Z \left(\sum_{\mathrm{i}=0}^{\mathrm{j}} \langle l, \widehat{\mathbb{X}}_{0,t_{\mathrm{i}}}^{(m)} \rangle^2 \right) (Y_{\mathrm{j}+1}^{(m)} - Y_{\mathrm{j}}^{(m)}) \right\}. \tag{3.12}$$

Equations (3.10) and (3.12) will serve as our foundational framework for addressing the signature optimal stopping problem. Initially, we utilize Equation (3.12) to derive the optimal linear policy l^* from the training samples. Subsequently, we determine the optimal stopping time for the target time series by applying Equation (3.10).

3.2 Signature Optimal Mean Reversion Trading

The main contribution of this paper is the development of an innovative trading strategy tailored for mean reversion markets. Our methodology hinges on the formulation of a sequence of optimal stopping problems, aimed at identifying the most advantageous moments for both market entry and exit.

In addition, we propose a solution to these optimal stopping problems through a sequential signature-based optimal stopping framework. This serves as a robust approximation to the broader, more general optimal stopping challenges commonly found in financial mathematics. Notably, our approach is highly adaptable, accommodating any form of mean-reverting dynamics. This framework empowers us to derive a sequence of optimal timing solutions for trading activities, thereby elevating the existing mean-reversion trading rules in both theoretical constructs and real-world applications.

3.2.1 Sequential Signature Optimal Stopping Framework

Given that a price process or portfolio value following a mean-reverting process, our primary goal is to investigate the optimal timing for opening a long position and subsequently closing the position. To achieve this, we first introduce a framework that involves solving a sequence of signature optimal stopping problems.

Long Strategy

Suppose that the investor trades on a spread or portfolio whose value process $(X_t)_{t\geq 0}$ follows a mean reverting process. Recall that $\widehat{\mathbb{X}}$ denotes the augmented signature of X_t and the filtration generated by signatures is $(\mathcal{G}_t)$, where $\mathcal{G}_t = \sigma(\widehat{\mathbb{X}}_{0,s} : 0 \leq s \leq t)$. Denote by $\mathcal{S}$ the space of all $(\mathcal{G}_t)$-stopping times.

Let us define the shorthand notation for the expectation conditional on the starting value x as $E_x\{\cdot\} := E\{\cdot \mid X_0 = x\}$. If the long position is opened at some time $\tau_0 > 0$, then the investor will pay the value X_{τ_0} plus a constant transaction cost $c > 0$. Assuming that $r > 0$ is the investors subjective constant discount rate, the total cost of entering a long position is $e^{-r\tau_0}(X_{\tau_0}+c)$. Our goal is to minimize the expectation of the entering cost. Therefore, we obtain the following optimal stopping problem for finding the optimal time of entering a long position:

$$\sup_{\tau_0\in\mathcal{S}} E_{x_0}\{e^{-r\tau_0}(-X_{\tau_0}-c)\}, \tag{3.13}$$

where $\mathcal{S}$ is the space of all $(\mathcal{F}_t)$-stopping times. In a word, we are minimizing the expected cost of entering a long position, given the initial point x_0.

After obtaining the optimal time to enter a long position, denoted by τ_0^*, the next question is when to exit the position. If the position is closed at time ν_0, the investor will receive the value X_{ν_0} and pay a constant transaction cost $\hat{c}>0$. To maximize the expected discounted value, the investor solves the optimal stopping problem:

$$\sup_{\nu_0\in\mathcal{S}} E_{x_{\tau_0^*}}\{e^{-\hat{r}\nu_0}(X_{\nu_0}-\hat{c})\}. \tag{3.14}$$

Here, we aim to maximize the expectation of discounted earnings conditional on the starting point $x_{\tau_0^*}$. Note that the pre-entry and post-entry discount rates, $\hat{r}$ and r, as well as the entry and exit transaction costs $\hat{c}$ and c, can differ in our analysis.

Then we can solve a sequence of optimal stopping problems iteratively to obtain the optimal entry and exit times. That is, for $\mathrm{i}=1,2,3,\ldots$, we solve the following equations iteratively:

$$\begin{aligned}&\sup_{\tau_\mathrm{i}\in\mathcal{S}} E_{x_{\nu_{\mathrm{i}-1}^*}}\{e^{-r\tau_\mathrm{i}}(-X_{\tau_\mathrm{i}}-c)\},\\ &\sup_{\nu_\mathrm{i}\in\mathcal{S}} E_{x_{\tau_\mathrm{i}^*}}\{e^{-\hat{r}\nu_\mathrm{i}}(X_{\nu_\mathrm{i}}-\hat{c})\}.\end{aligned} \tag{3.15}$$

Thus, we get the optimal times of trading $\{\tau_0^*,\nu_0^*,\tau_1^*,\nu_1^*,\tau_2^*,\nu_2^*,\cdots\}$.

Short Strategy

In the short strategy, our main objective is to determine the optimal timing for opening and subsequently closing a short position, similar to the long strategy. We only need to change the order of equations for the optimal stopping problems in (3.14). Specifically, we obtain the first optimal entry time ν_0^* and exit time τ_0^* by solving:

$$\begin{aligned}&\sup_{\nu_0\in\mathcal{S}} E_{x_0}\{\mathrm{e}^{-r\nu_0}(X_{\nu_0}-c)\},\\&\sup_{\tau_0\in\mathcal{S}} E_{x_{\nu_0^*}}\{\mathrm{e}^{-\hat{r}\tau_0}(-X_{\tau_0}-\hat{c})\}.\end{aligned} \tag{3.16}$$

Next, we solve the following equations iteratively for $\mathrm{i}=1,2,3,\cdots$:

$$\begin{aligned}&\sup_{\nu_\mathrm{i}\in\mathcal{S}} E_{x_{\tau^*_{\mathrm{i}-1}}}\{\mathrm{e}^{-r\nu_\mathrm{i}}(X_{\nu_\mathrm{i}}-c)\},\\&\sup_{\tau_\mathrm{i}\in\mathcal{S}} E_{x_{\nu_\mathrm{i}^*}}\{\mathrm{e}^{-\hat{r}\tau_\mathrm{i}}(-X_{\tau_\mathrm{i}}-\hat{c})\}\end{aligned} \tag{3.17}$$

to obtain the optimal trading times $\{\nu_0^*,\tau_0^*,\nu_1^*,\tau_1^*,\nu_2^*,\tau_2^*,\cdots\}$. The intuition behind is opening a short position when the spread price reaches a high value, and closing the position as the price drops significantly.

3.2.2 Non-Randomization of Signature Stopping Times

Following the signature optimal stopping framework, we attempt to deploy linear signature stopping times, denoted ν_l^r, as specified in Equation (3.10), to ascertain the optimal stopping times for a given stochastic process. The optimization of the linear policy l^* is conducted via a minimization of the loss function presented in Equation (3.12), and subsequently, the policy is applied to determine the signature optimal stopping time for the observed time series. However, the randomization introduced by the random variable Z in Equation (3.10) typically precludes a deterministic solution of stopping time. To address this challenge, we propose a novel approach by replacing the random variable Z with a constant k to provide a deterministic solution for the optimal stopping time, while preserving the applicability of Equations (3.10) and (3.12).

Specifically, the non-randomized stopping time associated with linear signature stopping policies is then defined as follows:

$$\nu_l := \inf\left\{0 \leq \mathrm{j} \leq n : \sum_{\mathrm{i}=0}^{\mathrm{j}} \langle l, \widehat{\mathbb{X}}^{<\infty}_{0,t_\mathrm{i}} \rangle^2 \geq k\right\}, \tag{3.18}$$

where k is a positive constant. This approach allows us to compute the stopping time without relying on randomness, offering a deterministic solution. Therefore, in our practical implementation, we calculate the optimal stopping time based on Equation (3.18) once we have determined the optimal linear stopping policy l^*.

Substituting the random variable Z with a constant k transforms the cumulative distribution function of Z into an indicator function. Thus, this substitution introduces non-differentiability into the loss function. Such non-differentiability poses computational challenges, particularly in the application of gradient-based optimization methods which are foundational to our algorithm's practical implementation. To overcome this challenge, we propose a refinement to the loss function as originally defined in Equation (3.12). Initially, we substitute G_Z with G_k, defined as $G_k = 1 - F_k$. Here, $F_k(x) = \mathbb{I}\{x > k\}$, with $\mathbb{I}$ as the indicator function. To facilitate gradient-based optimization, we approximate F_k using a sigmoid function centered at k. The approximation $\widehat{F}_k(x)$ is given by:

$$\widehat{F}_k(x) = \frac{1}{1 + \mathrm{e}^{-\mu\cdot(x-k)}}$$

where μ is a large positive constant that sharpens the sigmoid function. Figure 3.1 illustrates the approximation of the indicator function using a sigmoid function, providing a visual representation of how the sigmoid curve serves as a smoothed substitute for the step-like behavior of the indicator function. In the subsequent experiments, we employ a μ value of 20.

Subsequently, we employ the following modified loss function in our practical implementation:

$$\mathrm{loss}(l) = -\frac{1}{M}\sum_{\mathrm{i}=1}^{M}\left\{Y_0^{(m)} + \sum_{\mathrm{j}=0}^{n-1}\widehat{G}_k\left(\sum_{\mathrm{i}=0}^{\mathrm{j}}\langle l, \widehat{\mathbb{X}}^{(m)}_{0,t_\mathrm{i}}\rangle^2\right)(Y^{(m)}_{\mathrm{j}+1} - Y^{(m)}_{\mathrm{j}})\right\}, \tag{3.19}$$

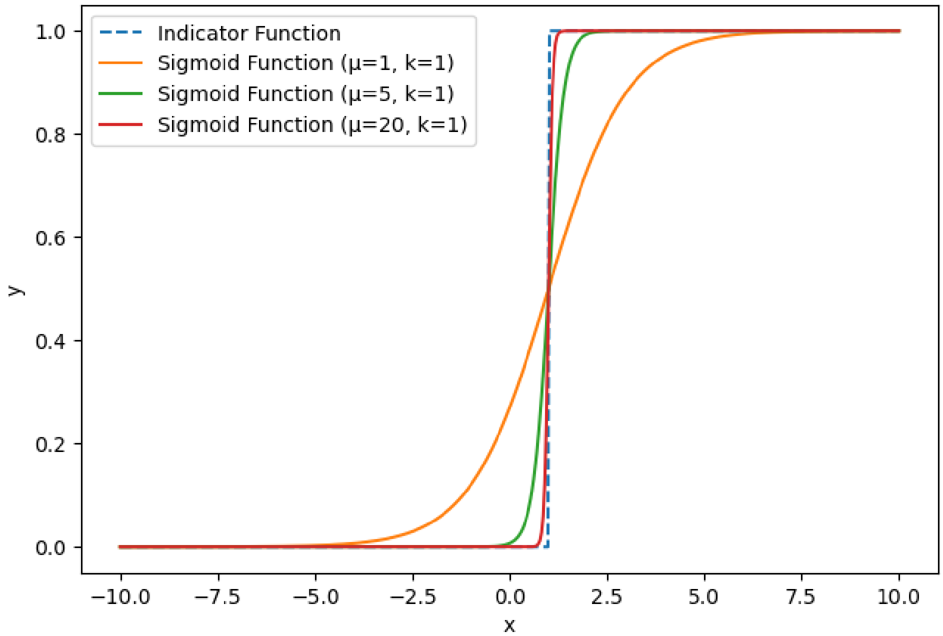


Figure 3.1. Approximating indicator function using sigmoid function.

where $\widehat{G}_k = 1 - \widehat{F}_k$. This formulation allows for gradient-based methods to be effectively employed for optimization.

In this section, we have provided a comprehensive overview of our implementation process of signature optimal stopping method, with a particular emphasis on the non-randomization of signature stopping times and the refined loss function. This loss function in (3.19) serves as a crucial component in our training procedure. Upon successfully obtaining the optimal linear policy, we proceed to compute the corresponding optimal signature stopping time using Equation (3.18). This comprehensive methodology help us solve the complex sequential signature optimal stopping problems in section 4.1.

3.2.3 Training Samples Generation

To implement the tailored signature optimal stopping method introduced in Section 4.2, this section delves into the training procedures employed in our methodology. The whole process can be applied to both long and short strategies, but we will focus on the long strategy in this discussion.

Recall that the long strategy framework is grounded in the formulations presented by equations (3.13), (3.14), and (3.15). To solve the first signature optimal stopping problem (3.13), we need to obtain the optimal linear function l_0^* by minimizing the loss function given in equation (3.19). However, we can only observe one path $X_t, 0 \leq t \leq T$, which is actually our target time series. So we first generate training samples of X_t starting at a given initial value x, and the corresponding payoff process $Y_t = \mathrm{e}^{-rt}(-X_t - c)$. We then use these samples to minimize the loss function to obtain the optimal linear function $l_{0,1}^*$. Once we obtain the optimal linear stopping policy $l_{0,1}^*$, we can use it to approximate the optimal entry time τ_0^* by the linear non-randomized optimal stopping time defined in equation (3.18). Specifically, the optimal entry time is given by:

$$\tau_0^* = \inf\left\{0 \leq \mathrm{j} \leq n : \sum_{\mathrm{i}=0}^{\mathrm{j}} \langle l_{0,1}^*, \widehat{\mathbb{X}}_{0,t_\mathrm{i}}^{\leq N}\rangle^2 \geq k\right\}$$

where n is the number of discretization intervals, $\widehat{\mathbb{X}}_{0,t_\mathrm{i}}^{\leq N}$ is the N-truncated augmented signature of the path X_t from time 0 to time t_i, and k is a positive threshold value.

To address the optimal exit problem specified in Equation (3.14), it is necessary to simulate sample paths of X_t starting from the new initial value $X_{\tau_0^*}$ over the time interval $[\tau_0^*, T]$. The payoff process is calculated as $Y_t = \mathrm{e}^{-\hat{r}t}(X_t - \hat{c})$. We then can obtain another optimal linear function $l_{0,2}^*$ using the generated samples, and the optimal exit time is given by:

$$\nu_0^* = \inf\left\{0 \leq \mathrm{j} \leq n : \sum_{\mathrm{i}=0}^{\mathrm{j}} \langle l_{0,2}^*, \widehat{\mathbb{X}}_{\tau_0^*,\tau_0^*+t_\mathrm{i}}^{\leq N}\rangle^2 \geq k\right\}$$

where $\widehat{\mathbb{X}}_{\tau_0^*,\tau_0^*+t_\mathrm{i}}^{\leq N}$ is the N-truncated augmented signature of the path of X_t from time τ_0^* to time $\tau_0^*+t_\mathrm{i}$. We can continue this computation process to solve the sequence of problems in (3.15) and thus obtain $\{\tau_0^*, \nu_0^*, \tau_1^*, \nu_1^*, \tau_2^*, \nu_2^*, \cdots\}$, where each pair of values (τ_n^*, ν_n^*) corresponds to the optimal entry and exit times for the n-th trading. To be more specific, for $n \geq 1$, we have:

Generate sample paths of X_t starting from $X_{\nu^*_{n-1}}$ over $[\nu^*_{n-1}, T]$ and $Y_t = \mathrm{e}^{-rt}(-X_t - c)$;

Use the signature optimal stopping method to obtain $l^*_{n,1}$ and τ^*_n;

Generate sample paths of X_t starting from $X_{\tau^*_n}$ over $[\tau^*_n, T]$ and $Y_t = \mathrm{e}^{-\hat{r}t}(X_t - \hat{c})$;

Use the signature optimal stopping method to obtain $l^*_{n,2}$ and ν^*_n.

We continue this process until we reach the end of the trading period. At each step, we obtain the optimal linear function and the corresponding optimal stopping time, which can be used to guide our trading decisions.

As for the methods for generating training sample paths from the observed time series X_t, we have several choices. The initial approach is block bootstrapping [31], a widely utilized method in the bootstrapping of time series data. This technique involves resampling contiguous blocks of data points, thereby preserving the data's inherent structures. An alternative strategy involves fitting the observed path to a parametric statistical or stochastic model, subsequently generating new samples from this model. For instance, the path could be modeled using an Autoregressive (AR) model or an Ornstein-Uhlenbeck (OU) process. Additionally, generative machine learning approaches, such as time-series Generative Adversarial Networks (GAN) [64], present a modern method for sample path generation. Researchers have the flexibility to select the most suitable method to fulfill the requirements of the algorithm, considering the specific characteristics of the observed data.

3.3 Numerical Experiments

In this section, we undertake numerical experiments to rigorously assess the efficacy of our proposed methodologies. We initiate our evaluation by scrutinizing the performance of the proposed optimal trading timing approach through its application to simulated mean reversion spreads. Following this, we extend our analysis to real-world scenarios by evaluating the mean reversion trading performance of our method across four carefully chosen, highly correlated asset pairs in the U.S. stock markets.

3.3.1 Experiments on Simulated Data

In this section, we conduct simulated experiments on the effectiveness of the proposed signature optimal mean reversion trading method.

Test on Signature Optimal Stopping Method

In this part, we conduct simulations to gauge the effectiveness of the signature-based optimal stopping method in the context of mean reversion spreads. We employ the Ornstein-Uhlenbeck (OU) process to simulate a mean-reverting spread, governed by the following stochastic differential equation:

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t.$$

Our primary aim is to identify the supremum of the expected value of the OU process at the signature optimal stopping time, formally expressed as:

$$\sup_{\tau \in \mathcal{S}} \mathbb{E}[X_{\tau \wedge T}].$$

By leveraging this simulation framework, we intend to rigorously evaluate the capability of the signature optimal stopping method to accurately identify optimal stopping values in mean-reverting contexts.

In applying the signature optimal stopping method to solve the optimal stopping problem for an Ornstein-Uhlenbeck (OU) process, our approach consists of two distinct phases. In the first phase, we synthesize a dataset consisting of 100 training samples. These samples are used to compute the optimal linear function l^* that minimizes the loss function, as delineated in Equation (3.19). In the second phase, we generate a supplementary set of 10 testing samples. Utilizing the previously computed optimal linear function l^*, we ascertain the optimal stopping times for these test cases in accordance with Equation (3.18) and record the corresponding values of path at that time. This two-step procedure enables a comprehensive evaluation of the method's performance, both in terms of its optimization capabilities and its generalizability to new instances.

In this specific experiment, the parameters that define the Ornstein-Uhlenbeck (OU) process are set as follows: $\mu = 10$, $\theta = 10$, and $\sigma = 1$. As depicted in Figure 3.2, the graph displays the trajectories of the 10 testing samples. The calculated average optimal stopping value is found to be 10.23. Importantly, the maxima for the majority of these paths are observed to cluster around 10.4, which lends credence to the computed optimal stopping value of 10.23. This result compellingly validates the effectiveness of the signature optimal stopping methodology, as it consistently produces high optimal stopping values for the simulated OU processes.

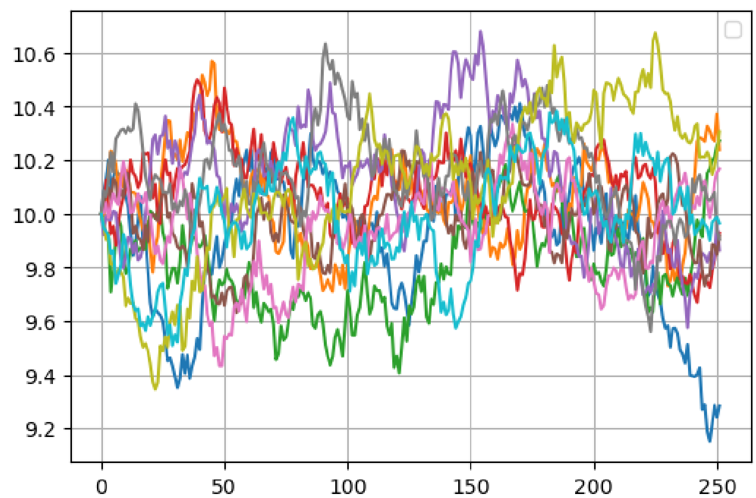


Figure 3.2. The paths of the 10 testing mean-reverting samples.

We have also executed a series of tests to examine the optimal stopping values generated by the signature method across a range of parameter combinations within the Ornstein-Uhlenbeck (OU) process. The aggregate findings of these tests are consolidated in Table 3.1. Significantly, all the optimal stopping values obtained exceeded the long-term mean of 10. A particularly insightful observation is the positive correlation between the optimal stopping value and the volatility parameter σ. This observed relationship underscores the sensitivity of the optimal stopping decision to the volatility of the underlying process, highlighting the nuanced interplay between process parameters and stopping outcomes.

Table 3.1. The optimal stopping value estimated by signature method on simulated mean-reversion process.

μ	σ	θ	Optimal Stopping Value
1	1.0	10	10.1284
5	1.0	10	10.2311
10	1.0	10	10.2332
15	1.0	10	10.2005
20	1.0	10	10.1703
10	0.1	10	10.0025
10	0.5	10	10.0713
10	1.0	10	10.2264
10	1.5	10	10.3599
10	2.0	10	10.5173

Simulation on Signature Optimal Mean Reversion Trading

In this section, we conduct simulations using the Ornstein-Uhlenbeck (OU) process to test the effectiveness of our proposed optimal trading timing method for identifying both entry and exit times in mean-reverting markets. We continue to simulate the mean-reversion paths by an OU model. This simulation study serves to validate the utility of our approach in generating actionable trading decisions based on the underlying stochastic process.

Figure 3.3 displays the performance of our simulation, where we have marked optimal entry times using red points and optimal exit times with green points. This color-coded representation provides an intuitive snapshot of how adeptly our method identifies strategic moments to both enter and exit trades in a mean-reverting market. As is evident from the graph, our method proves highly effective, consistently point times that align well with the natural peaks and troughs of the mean-reverting process. These results lend robust empirical support to the theoretical analysis of our proposed trading timing methodology. Moreover, they offer compelling evidence that our approach can serve as a valuable tool for traders and portfolio managers seeking to exploit mean-reversion opportunities in financial markets. The simulation outcomes underscore the practical applicability and accuracy of our algorithm, confirming its utility for making well-informed, timely trading decisions based on the dynamics of the underlying stochastic process.

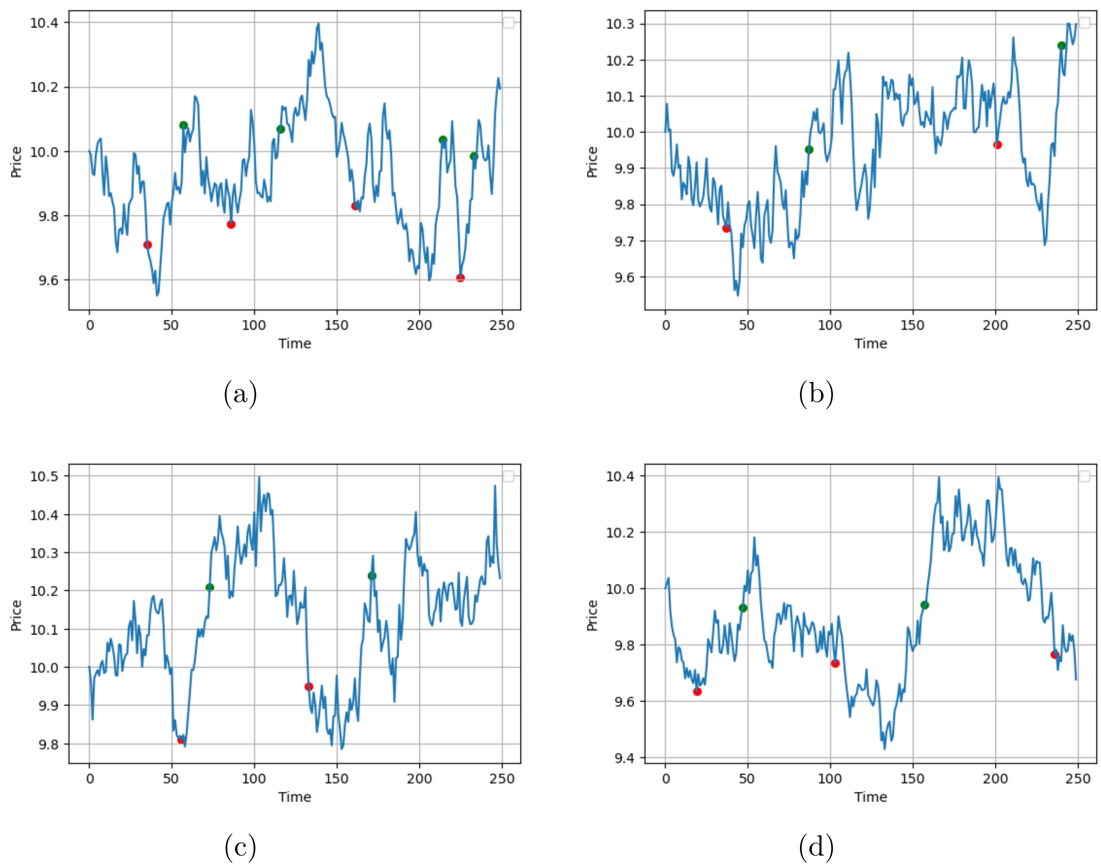


Figure 3.3. The sequential optimal entry and exit times for four simulated mean reversion processes based on optimal trade timing framework.

3.3.2 Experiments in Real Market

In this section, we broaden the application of our signature optimal mean reversion trading methodology to a real-world trading context. It is important to emphasize that our method operates without making any assumptions about the dynamics of the underlying traded spreads in real market.

Initially, we handpick four highly co-integrated stock pairs from diverse sectors within the U.S. market to create their corresponding spreads: WM-RSG from waste management, UAL-DAL from the airline industry, V-MA from financial services, and GS-MS from investment banking. A comprehensive description of these companies is provided in Table 3.2. Data pertaining to the daily closing prices for these stocks was collected for the period spanning January 1, 2021, to December 31, 2022. The data was sourced via the Yahoo! Finance API[1].

Table 3.2. The stock pairs used for testing and the description of the company.

Pairs Symbols	**Description**
WM-RSG	Waste Management and Republic Services provide waste management and environmental services
UAL-DAL	United Airlines and Delta Air Lines are two major American airline
V-MA	Visa and Mastercard are two large-cap stocks in the payment industry
GS-MS	Goldman Sachs and Morgan Stanley are American multinational investment bank

Figure 3.4 depicts the price trajectories of each stock pair, revealing a notable similarity in the price paths within each pair. For the sake of clarity and comparability, we have normalized each price series by dividing it by its respective initial value. This normalization allows for a more straightforward visual assessment of the relative movements between the stocks in each pair, thereby setting the stage for the forthcoming trading experiments.

In the experiments, we simulated one full year of trading activity, as visually presented in Figure 3.5. Specifically, we created asset pairs based on the daily adjusted closing prices from the first year, covering a period of 252 trading days. These pairs were then actively traded over the course of the subsequent year. By doing so, we aim to evaluate the real-world applicability and effectiveness of our signature-based optimal stopping method within a time frame that captures a meaningful range of market conditions. This setup allows us to assess how well the trading rules, derived from the initial year's data, generalize to unseen data in the following year. In the setup of the proposed signature optimal mean reversion trading method, the training samples generation procedure employed OU process and the threshold value $k = 0.05$.

As illustrated in Figure 3.5, our algorithm excels in identifying the opportunities of trading, consistently entering positions at low prices and exiting at higher prices. This

[1] ↑https://pypi.org/project/yfinance/

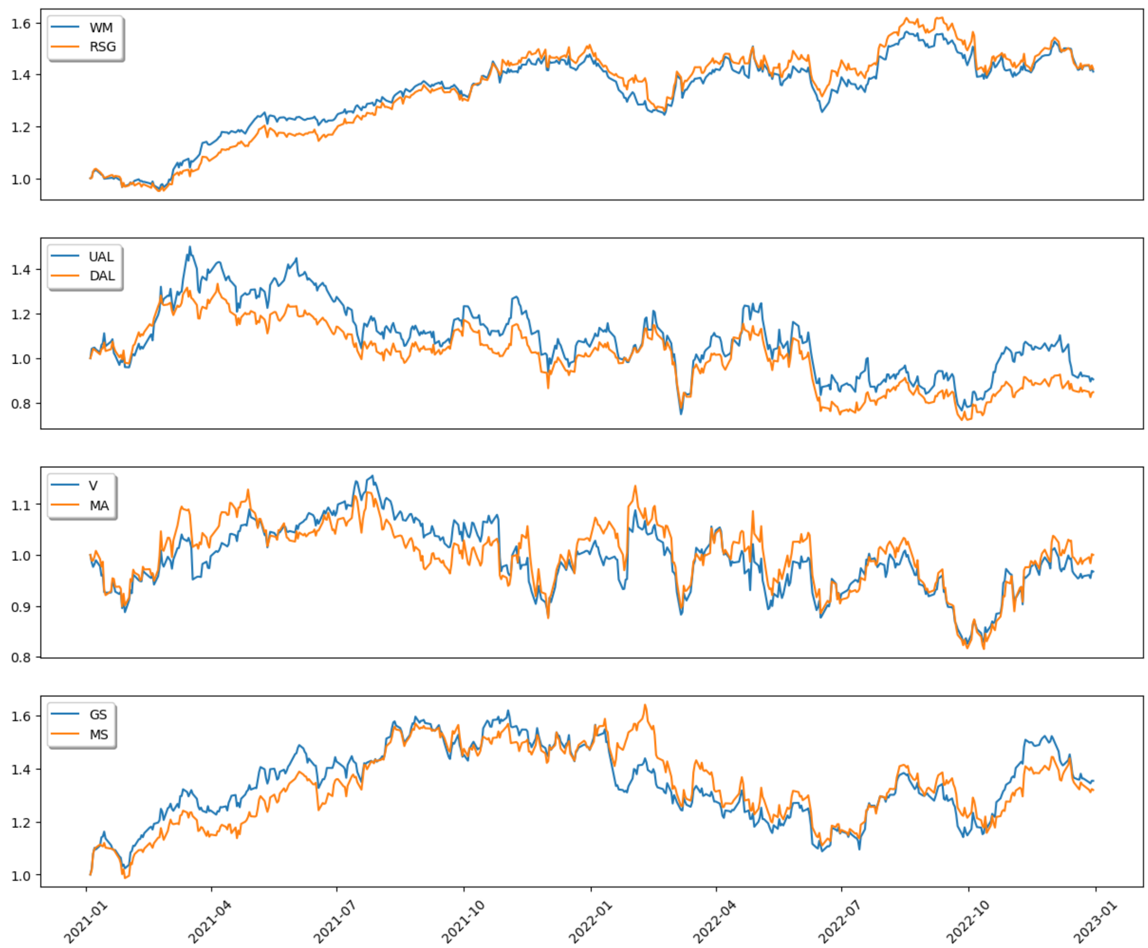
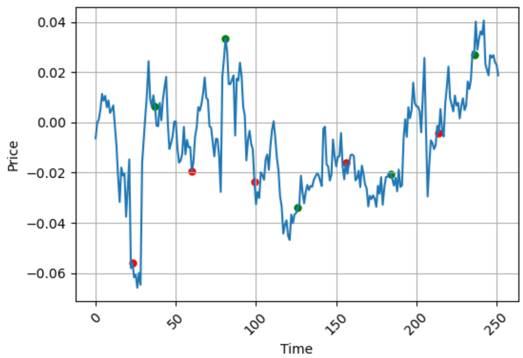


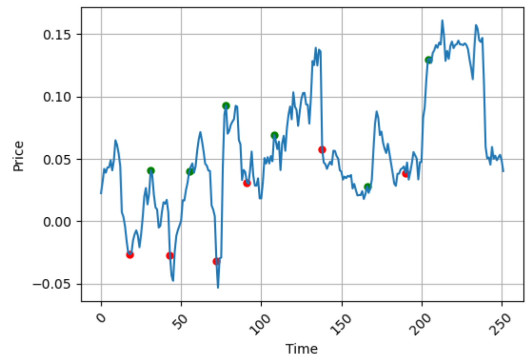
Figure 3.4. Normalized historical daily close prices of four target pairs.

behavior serves as a compelling demonstration to the efficacy of our signature-based optimal trading strategy. This strong performance underscores the algorithm's utility as a valuable tool for market participants seeking to exploit mean-reversion opportunities with optimized entry and exit timings.

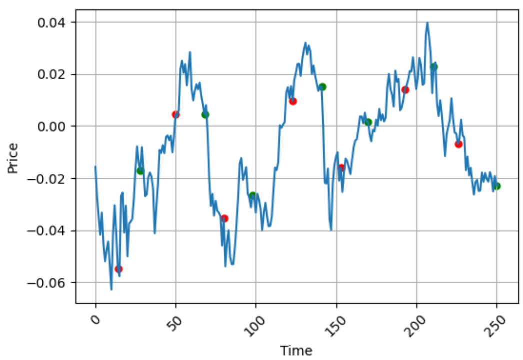
In the final phase of our analysis, we contrast the performance of our signature-based optimal trading strategy with a baseline strategy across all four asset pairs. The baseline strategy follows a simple rule: go long on X_t when its price falls below $MA(X_t) - k \times Std(X_t)$, and liquidate the position when the price rises above $MA(X_t) + k \times Std(X_t)$. Here we write



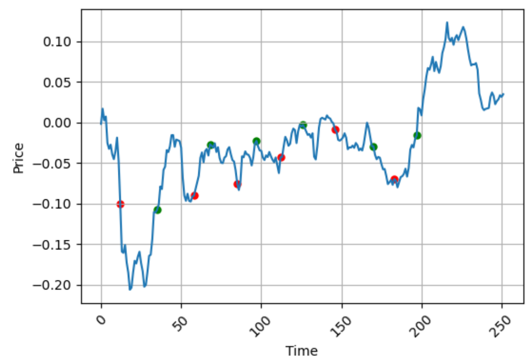
(a) WM-RSG



(b) UAL-DAL



(c) V-MA



(d) GS-MS

Figure 3.5. The constructed spreads and optimal entry and exit time points. Red spot is entry point and Green spot is exit point.

$MA(X_t)$ and $Std(X_t)$ to represent the moving average and standard deviation of past spread, respectively.

Table 3.3 presents a thorough comparison of performance metrics for both the baseline and the signature optimal trading (SOT) strategies. In this evaluation, the baseline strategy employs a constant k value of 0.1 for its trading rules and leverages the preceding 100 samples to compute both the moving average and standard deviation. Figure 3.6 graphically illustrates the cumulative returns generated by each strategy over a one-year time frame.

As revealed by Table 3.3, our proposed signature optimal trading strategy substantially outperforms the baseline model in key performance metrics, including higher cumulative returns and a more favorable daily Sharpe ratio. This head-to-head comparison is specifically

Table 3.3. Performance summary of testing portfolios for baseline and the signature optimal trading (SOT).

	WM/RSG		**UAL/DAL**		**V/MA**		**GS/MS**	
	Baseline	SOT	Baseline	SOT	Baseline	SOT	Baseline	SOT
DailyRet (%)	0.0110	0.0390	0.0401	0.1443	0.0107	0.0275	0.0119	0.0604
DailyStd (%)	0.3513	0.4321	0.8244	0.8122	0.4177	0.5306	0.3362	0.5986
Sharpe	0.0313	0.0903	0.0487	0.1777	0.0255	0.0519	0.0354	0.1012
MaxDD (%)	-1.9309	-0.7674	0	0	-0.7057	-0.2954	-2.2522	-7.9006
CumPnL (%)	2.6363	10.0345	9.6431	42.4559	2.4857	6.7732	2.8829	15.8486
TradeNum	9	5	4	6	3	7	4	6

tailored to underscore the superior efficacy of the SOT approach. It compellingly attests to the enhanced performance and incremental value that our method brings to the table in trading scenarios.

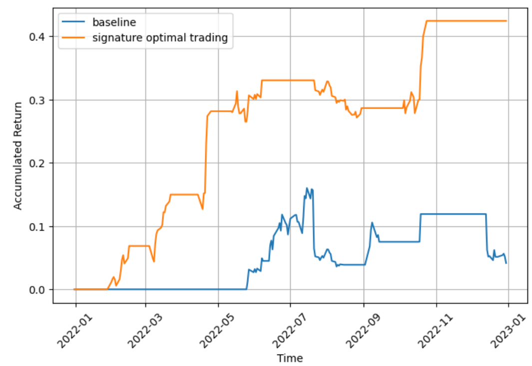
The visual representation of cumulative returns in Figure 3.6 serves as another pivotal element in our comparative assessment. It provides an intuitive snapshot of how the respective trading strategy compound over a one-year period. By offering this view, we further reinforce the idea that the SOT strategy is not only theoretically robust but also practically effective and adaptable to real-world market conditions. This comprehensive evaluation aims to highlight the enduring viability and real-world utility of the SOT methodology.

3.4 Conclusions

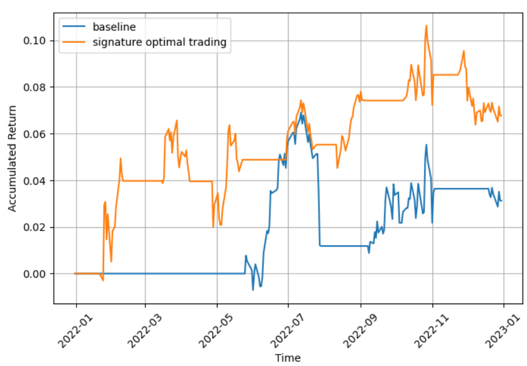
In this chapter, we have presented an innovative approach to identifying optimal timing strategies for trading price spreads with mean-reverting properties. Specifically, we formulated a sequential optimal stopping problem that accounts for the timing of both position entry and liquidation, while also incorporating transaction costs. To tackle this challenging problem, we employed the signature optimal stopping method, a powerful tool for determining the optimal entry and exit times that maximize trading returns.

Our approach is distinguished by its versatility: it is designed to operate without any assumptions of mean reversion dynamics, making it broadly applicable to a variety of trading scenarios. To validate the efficacy of our methodology, we carried out a comprehensive

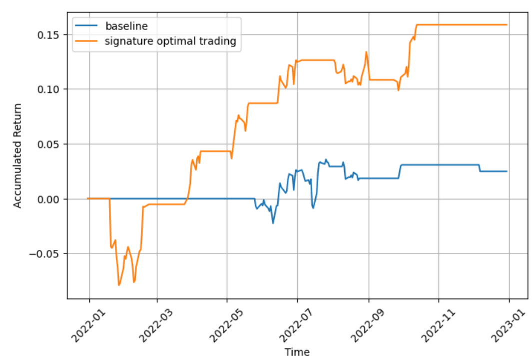
(a) WM-RSG



(b) UAL-DAL



(c) V-MA



(d) GS-MS

Figure 3.6. The accumulative returns for each pairs obtained from baseline and SOT.

set of numerical experiments. These experiments served to compare the performance of our approach against conventional mean-reversion trading rules. The results consistently demonstrated the superior performance of our signature-based optimal trading strategy in terms of key metrics such as cumulative returns and Sharpe ratios.

In summary, this chapter makes a significant contribution to the field of quantitative finance by introducing a robust and adaptable method for optimizing trading decisions in mean-reverting markets. Our findings affirm that the signature optimal stopping method offers not only theoretical rigor but also practical utility, presenting a compelling solution to a complex problem in finance. While our focus has been on specific asset pairs and trading conditions, the general principles and techniques introduced here have the potential for broader applications, opening up avenues for future research and practical implementations.

4. ADVANCED STATISTICAL ARBITRAGE WITH REINFORCEMENT LEARNING

After identifying groups of similar stocks, the next step in statistical arbitrage involves constructing an statistical arbitrage portfolio or spread with mean-reverting property. A foundational approach, as suggested by Gatev, Goetzmann, and Rouwenhorst [1], involves taking a long position in one security of the pair and a short position in the other, that is, trading on the spread $S_1 - S_2$ for two similar stocks S_1 and S_2. Although straightforward, this method often cannot create an optimal portfolio that exhibits high mean-reverting characteristics. A more sophisticated strategy frequently used is Ornstein-Uhlenbeck (OU) mean reversion trading [4]. For a pair of similar stocks S_1 and S_2, the goal is to find a coefficient B such that the spread $S_1 - B \cdot S_2$ mimics an OU process as closely as possible, with B typically determined through maximum likelihood estimation based on the OU process distribution. However, the real-world application of this strategy faces challenges due to the fact that financial markets may not always align with the assumptions of the Ornstein-Uhlenbeck process, which can compromise the effectiveness of strategies based on this model.

To overcome the limitations inherent in these assumption-dependent methods, we introduce a novel approach that utilizes the proposed empirical mean reversion time of any time series as a measure of reversion speed. This allows for the construction of a mean reversion spread without relying on any theoretical assumptions. By employing a grid search method, we can systematically explore different combinations to identify an optimal spread that possesses the least empirical mean reversion time. This technique offers a more flexible and potentially more robust framework for arbitrage portfolio construction.

The final phase of statistical arbitrage involves formulating a trading strategy based on the constructed mean-reverting spread. Traditional strategies heavily rely on model parameter estimations. For instance, Gatev, Goetzmann, and Rouwenhorst [1] initiate a trade when a spread's price deviation exceeds two historical standard deviations from the mean, calculated during the pair formation phase, and exit the trade upon the next convergence of prices to the historical mean. In the context of OU mean reversion trading, parameter estimations for the long-term mean and volatility of the OU model are typically employed to define trading criteria [4], [59]. These estimations depend on historical data from the formation period, with an underlying assumption that parameters remain constant in the following trading phasean assumption that may not hold due to market fluctuations. Additionally, the selection of hyper-parameters significantly impacts trading performance, yet a robust method for the optimal hyper-parameters selection remains absent. For example, the determination of an appropriate threshold for price deviation from the mean lacks a clear consensus.

To address these challenges, we introduce a reinforcement learning (RL) algorithm designed to dynamically optimize trading decisions over time, replacing the need for predefined rules. This approach models the task within a reinforcement learning framework, aimed at enabling agents to take actions that maximize cumulative rewards in an environment. We design the state space to capture the recent movements of the spread price, thus moving away from a dependence on historical mean and standard deviation estimates. This approach enables the agent to make informed decisions about future actions by leveraging insights into current market trends, rather than depending on parameter estimations from the formation period. We get the rid of the hyper-parameters choice at the same time. Simultaneously, we remove the necessity for hyper-parameter selection by not incorporating universal hyper-parameters, such as thresholds, which can significantly impact trading performance. This approach streamlines the trading process, focusing on dynamic adaptation without the constraints of fixed parameters.

The structure of the chapter is organized as follows. Section 4.1 details the definition of empirical reversion time for spreads and outlines the methodology for identifying optimal asset coefficients by minimizing mean reversion time. Section 4.2 presents a reinforcement learning framework designed for the development of optimal trading strategies. Experimental results, based on simulated data and real-world applications in the US stock market, are discussed in Section 4.3. Finally, Section 4.4 offers conclusions and outlines future research directions.

4.1 Empirical Mean Reversion Time: Spread Construction

Once we identify groups of similar stocks, we need to form an arbitrage portfolio with mean reversion property. In the traditional OU pairs trading, for two similar stocks S_1 and S_2, we choose B such that the spread $S_1 - BS_2$ follows an OU process as closely as possible. A reasonable way to find B is to use the maximum likelihood estimator, using the distribution of the OU process. However, since we do not have a model assumption, we cannot use MLE anymore.

We extend paired trading to a multi-asset portfolio. In other words, given n similar stocks $S_{\mathrm{i}}, \mathrm{i} = 1, 2, ..., n$, we form a spread $X = \sum_{\mathrm{i}=1}^{n} a_{\mathrm{i}} S_{\mathrm{i}}$. Our goal is to find a portfolio $(a_1, a_2, ..., a_n)$ such that the spread X has a mean reverting property as much as possible.

Let us consider a popular OU process

$$dX_t = \mu(\theta - X_t)dt + \sigma dW_t,$$

where W_t is a standard Brownian motion. An empirical result shows that μ has the biggest impact on the profit among three parameters μ, θ and σ. In general, a larger μ gives higher return. Intuitively, it should be the case, since a larger μ implies a faster mean reversion. Therefore, a trader can make a quick profit by taking advantage of the deviation from the mean.

Therefore, we want to make the spread $X = \sum_{\mathrm{i}=1}^{n} a_{\mathrm{i}} S_{\mathrm{i}}$ to have a faster mean reversion property. Consider a trading strategy to buy at $X_t = \theta - a$ and sell later at $X_t = \theta$ for $a > 0$ and a given long-term mean θ. Define the stopping time

$$\tau_t = \inf\{s > t : X_s = \theta \ \mid \ X_t = \theta - a\}. \tag{4.1}$$

A faster mean reversion corresponds to a smaller τ_t. Based on this logic, it is natural to define our arbitrage portfolio selection problem as follows.

At time t, consider the training time interval $[t - h, t]$. Find the optimal portfolio $(a_1, a_2, ..., a_n)$ which minimizes the sample mean of τ's in this interval, given $\bar{X} = \theta$ and $S^2(X) < M$ for a constant M.

The reason we impose the upper bound M on the sample variance is because of the constraint of the initial wealth and to prevent a large leverage, which makes the portfolio unstable.

4.1.1 Empirical Mean Reversion Time

In this part, we introduce the conception of empirical mean reversion time (EMRT) based on the idea above. Inspired by Fink and Gandhi [65], we firstly define the important extremes of time series. Let s be the sample standard deviation of a time series X_t, where $t \in [0, T]$. In real market, we can only get the discrete data points. So we assume a time series $(X_1, \cdots, X_n)$. Let C be a positive constant. A point X_m is an important minimum of the time series if there are indices i and j, where $\mathrm{i} \leq m \leq \mathrm{j}$, such that

- X_m is the minimum among $X_\mathrm{i}, \cdots, X_\mathrm{j}$;
- $X_\mathrm{i} - X_m \geq C \cdot s$ and $X_\mathrm{j} - X_m \geq C \cdot s$.

Intuitively, X_m is the minimal value of some segment $X_\mathrm{i}, \cdots, X_\mathrm{j}$, and the endpoint values of this segment are much larger than X_m. Similarly, X_m is an important maximum if there are indices i and j, where $\mathrm{i} \leq m \leq \mathrm{j}$, such that

- X_m is the maximum among $X_\mathrm{i}, \cdots, X_\mathrm{j}$;
- $X_m - X_\mathrm{i} \geq C \cdot s$ and $X_m - X_\mathrm{j} \geq C \cdot s$.

Drawing on the conceptual framework introduced by Equation (4.1), our objective is to propose an empirical mean reversion time that quantifies the duration required for the spread to revert to its long-term mean, starting from the maximum deviation observed. It enables us to infer the optimal coefficients for securities by minimizing the spread's empirical reversion time.

We now proceed to construct a sequence of time moments $\{\tau_\mathrm{i}\}_{\mathrm{i}=0}^{N}$, derived recursively from the significant local extremes within the actual asset price process. More precisely, we define the initial time moment as

$$\tau_1 = \inf\{u \in [0, T] : X_u \text{ is a local extreme}\}.$$

Subsequently, τ_2 is identified as the first instance when the series crosses the sample mean $\hat{\theta}$, defined by

$$\tau_2 = \inf\{u \in [\tau_1, T] : X_u = \hat{\theta}\}.$$

Recursively, τ_3 is the first local extreme following τ_2, and τ_4 is the first crossing of the long-term mean after τ_3, and so on. Thus, all odd-numbered time moments $\{\tau_n\}_{n=1,3,5,\cdots}$ correspond to local extremes and are defined as

$$\tau_n = \inf\{u \in [\tau_{n-1}, T] : X_u \text{ is a local maximum}\}.$$

Conversely, all even-numbered time moments $\{\tau_n\}_{n=2,4,6,\cdots}$ are associated with the crossings of the long-term mean, specified by

$$\tau_n = \inf\{u \in [\tau_{n-1}, T] : X_u = \hat{\theta}\}.$$

The complete sequence $\{\tau_n\}_{n=1}^{N}$ is constructed in an inductive manner.

Once we get the time stamps of iterated time stamps $\{\tau_n\}$, the empirical reversion time r is defined as the average of the time interval from local extremes to crossing times. That is,

$$r = \frac{2}{N} \sum_{\substack{\mathrm{i}=2 \\ \mathrm{i\,even}}}^{N} (\tau_n - \tau_{n-1})$$

Next, we briefly introduce a *grid search algorithm* that can help us find the optimal coefficients based on the empirical mean reversion time. Assume the price processes of n similar assets are denoted by $S_1, S_2, \ldots, S_n$. Our aim is to find the optimal coefficients $(a_1, a_2, \ldots, a_n)$ such that the portfolio $X = \sum_{\mathrm{i}=1}^{n} a_{\mathrm{i}} S_{\mathrm{i}}$ exhibits the minimal empirical mean reversion time. Without loss of generality, we set the first coefficient to $a_1 = 1$. We then evaluate the empirical mean-reversion time of X for each coefficient a_{i}, where $a_{\mathrm{i}} \in [-3.00, -2.99, -0.98, \ldots, 2.99, 3.00]$ for $2 \leq \mathrm{i} \leq N$. The optimal coefficients are determined by selecting the set that minimizes the empirical mean reversion time of X.

4.2 Reinforcement Learning: Advanced Trading Strategies

The final stage of statistical arbitrage involves developing a trading strategy based on a mean-reverting spread. Traditional approaches assume parameters' stability from formation phase to trading phase, which market changes can challenge. Moreover, the choice of hyper-parameters, such as the deviation threshold, critically affects performance, yet a standard method for their optimal selection is lacking.

Our motivation is to leverage reinforcement learning algorithms to help us decide the optimal trading actions dynamically over time, other than design some preset rules manually. Reinforcement learning framework is a machine learning method concerned with how intelligent agents ought to take optimal actions in an environment in order to maximize the cumulative reward.

4.2.1 Preliminaries of Reinforcement Learning

In RL, the sequential decision-making problem is modeled as Markov decision process (MDP), which is an augmented structure of Markov process. In addition to a Markov process, one has the possibility of choosing an action from an available action space and get some reward that tells us how good our choices were at each step.

The environment is defined as the part of the system outside of the RL agents control. At each time step t, we observe the current state of the environment $S_t \in \mathcal{S}$ and then chooses an action $A_t \in \mathcal{A}$. The choice of action influences both the transition to the next state, as well as the reward received, R_t. Thus, we will get a sequence in MDP as:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, \cdots$$

Every MDP is uniquely determined by a multivariate conditional probability distribution $p(s', r \mid s, a)$, which is the joint probability of transitioning to state s' and receiving reward r, conditional on the previous state being s and taking action a.

A policy $\boldsymbol{\pi}$ is a mapping from states to probability distributions over the action space. If the RL agent is following policy $\boldsymbol{\pi}$, then in state s it will choose action a with probability

$\pi(a \mid s)$. To find the optimal policy, one must specify a goal function. A wide-used discounted goal function is defined as

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots \\ &= R_{t+1} + \gamma G_{t+1}, \end{aligned} \tag{4.2}$$

where R_t is the instant reward at time t and $\gamma \in (0,1)$ is a discount factor expressing that rewards further in the future are worth less than rewards which are closer in time. Our goal is to search for the optimal policy that maximizes the expectation of the goal function, namely

$$\max_{\pi} \mathbb{E}[G_t]$$

Next we introduce related concepts of Q-learning. The action-value function for policy π is the expectation of goal function, assuming we start in state s, take action a and then follow the policy π from then on

$$q_{\pi}(s,a) := E[G_t \mid S_t = s, A_t = a].$$

The optimal action-value function is then defined as

$$q_*(s,a) = \max_{\pi} q_{\pi}(s,a).$$

If we knew the optimal action-value function, we would know the optimal policy itself easily, that is, choose $a \in \mathcal{A}$ to maximize $q_*(s,a)$. Hence we can reduce the problem to finding q_*, which is solved iteratively based on the Bellman equations. It is straightforward to establish the Bellman equation of action-value function:

$$q_*(s,a) = \sum_{s',r} p(s',r \mid s,a)[r + \gamma \max_{a'} q_*(s',a')]. \tag{4.3}$$

The core of the Q-learning is to leverage Bellman equation as a simple value iteration update, using the weighted average of the old value and the new information.

Before learning begins, the approximate action-value function Q is initialized to a possibly arbitrary value. Then, the corresponding sample update for q-function of S_t, A_t, given a sample next state and instant reward, S_{t+1} and R_{t+1} (from the model), is the Q-learning update:

$$Q^{new}(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha \cdot \left(R_{t+1} + \gamma \cdot \max_a Q(S_{t+1}, a) - Q(S_t, A_t)\right), \tag{4.4}$$

where α is the learning rate, γ is the discount factor, R_{t+1} is the reward received after taking action A_t in state S_t, and S_{t+1} is the next state. Note that $Q^{new}(S_t, A_t)$ is the sum of three factors:

1. $(1-\alpha)Q(S_t, A_t)$: the current value weighted by the learning rate.
2. $\alpha \cdot R_{t+1}$: the weighted instant reward to obtain if action A_t is taken when in state S_t.
3. $\alpha \cdot \gamma \cdot \max_a Q(S_{t+1}, a)$: the maximum cumulative reward that can be obtained from next state S_{t+1} (weighted by learning rate and discount factor)

Generally, $R_{t+1} + \gamma \cdot \max_a Q(S_{t+1}, a)$ is referred as target Y_t. Thus, the iteration (4.4) updates the current value $Q(S_t, A_t)$ towards a target value Y_t.

An epsilon-greedy strategy is employed for action selection. In the training phase, actions are chosen at random with a probability of ϵ, whereas the action with the highest Q-value is selected with a probability of $1 - \epsilon$. This approach facilitates a balance between exploration of new actions and exploitation of known values. In the testing phase, when the trained agent is assessed using new incoming data, ϵ is adjusted to 0. This modification ensures that action selection is solely based on the highest Q-value, thereby focusing entirely on exploitation based on the acquired knowledge. Thus, the model incorporates both exploration and exploitation during training, while adopting a strategy of pure exploitation during testing.

4.2.2 RL Model for Mean Reversion Trading

Now we can introduce our reinforcement learning model for an optimal mean reversion trading strategy.

The state space is constructed based on the trajectory of price movements over a recent sequence of time points. At any particular moment t, the state, denoted S_t, is encapsulated by the vector

$$S_t = [d_{t-l+1}, d_{t-l+2}, \ldots, d_t],$$

where each d_i characterizes the direction and magnitude of price changes at time i. A positive value of d_i signifies a price increase relative to time $\mathrm{i}-1$, and conversely, a negative value indicates a decline. The magnitude of d_i quantifies the extent of this change. Formally, for each i within the interval $t-l+1 \leq \mathrm{i} \leq t$, let $\pi_\mathrm{i} = \left(\frac{P_\mathrm{i}-P_{\mathrm{i}-1}}{P_{\mathrm{i}-1}}\right) \times 100$ represent the percentage price change from $\mathrm{i}-1$ to i. The definition of states is given by

$$d_\mathrm{i} = \begin{cases} +2 & \text{if } \pi_\mathrm{i} > k, \\ +1 & \text{if } 0 < \pi_\mathrm{i} < k, \\ -1 & \text{if } -k < \pi_\mathrm{i} < 0, \\ -2 & \text{if } \pi_\mathrm{i} < -k. \end{cases}$$

Here, '+2' indicates a significant increase, '+1' a moderate increase, '-2' a significant decrease, and '-1' a moderate decrease. The choice of threshold k, such as 3%, is adjustable to accommodate different sensitivity levels. This approach results in a state space comprising 4^l unique states, effectively capturing a wide spectrum of recent price movement scenarios.

The action space is composed of three possible actions: selling one share is represented by -1, taking no action is denoted by 0, and buying one share is indicated by $+1$. The set of available actions at any given time is contingent upon the agent's current position. Specifically, when the agent does not hold any position, the permissible actions include buying $(+1)$ or holding (0). Conversely, if the agent is currently in a long position, the options are limited to selling (-1) or holding (0). Note that our model does not account for initially entering a short position.

The immediate reward, R_{t+1}, earned by the agent for taking action A_t under prevailing environmental conditions, is mathematically defined as:

$$R_{t+1} = A_t \cdot (\theta - X_t) - c \cdot \mid A_t \mid, \tag{4.5}$$

where X_t denotes the current price of the spread, and θ represents the true global mean of X_t. The formulation is designed such that a buy action ($A_t = +1$) is rewarded positively when X_t is below its long-term mean, θ, encouraging purchases at lower prices. Conversely, a sell action ($A_t = -1$) incurs a negative reward under the same conditions. If X_t exceeds the long-term mean, resulting in a negative value for $\theta - X_t$, the rewards for buy and sell actions are adjusted accordingly to discourage buying at high prices and encourage selling. The term c represents the transaction cost per trade.

The cumulative return from time t to the terminal time T is expressed as:

$$G_t = \sum_{s=t+1}^{T} \mathrm{e}^{-r\cdot(s-t)} \cdot R_s + I_T \cdot X_T, \tag{4.6}$$

where r denotes the interest rate, reflecting the time value of money, and I_T signifies the position held at the terminal time. This formulation accounts for the exponential decay of rewards over time due to the discounting effect of the interest rate, emphasizing the importance of immediate gains and the impact of holding a position until the end of the considered period.

To accurately fit the optimal Q-table, ample training data is essential. However, the real market offers only a limited observation path of spreads, presenting a significant challenge for effective training. Additionally, our reward function, as defined in Equation (4.5), incorporates the true long-term mean of the spread value that remains elusive in actual market scenarios.

To overcome these obstacles, our strategy involves initially simulating a multitude of mean reversion spreads with different parameters to train the reinforcement learning (RL) agent. This step allows for extensive exposure to various market conditions, enhancing the agent's learning and decision-making capabilities. Subsequently, the model, now adept from the simulation training, is applied to execute trades in the real market. In the language of RL, this method entails utilizing a simulated environment for the agent's training phase.

4.3 Experiments

In this chapter, the performance of the proposed method is thoroughly evaluated. Initially, tests are conducted on simulated data to assess the efficacy of the proposed empirical mean reversion time, and the reinforcement learning (RL) model. Subsequently, mean reversion trading experiments are carried out on the S&P 500 using the proposed model-free framework, with outcomes compared against those achieved with other classical statistical arbitrage methods.

4.3.1 Empirical Mean Reversion Time

In this section, we investigate the empirical mean reversion time by conducting simulations of the Ornstein-Uhlenbeck (OU) process. We fix the parameters $\theta = 0$ and $\sigma = 1$, and vary μ from 2 to 20 in increments of 2. For each parameter combination, we simulate 100 paths of the OU process, each with a terminal time of $T = 1.0$ and $n = 1000$ data points. Subsequently, we calculate the average empirical mean reversion time using a threshold of $C = 2$ for these paths. The primary objective is to compute the average empirical mean reversion time for these paths using a threshold value of $C = 2$, thereby examining the impact of the mean reversion parameter μ on the empirical mean reversion time.

Figure 4.1 presents the identified local extremes on a simulated Ornstein-Uhlenbeck (OU) path characterized by parameters $\mu = 10$, $\theta = 0$, and $\sigma = 1$. The analysis reveals that the defined criteria for extreme points are highly effective in pinpointing nearly every instance where the time series reaches a local maximum or minimum. This capability underscores the precision of our approach in capturing significant turning points within the simulated path, providing a robust method for analyzing mean reversion characteristics. The accurate identification of these extremes is critical for the development of mean reversion time.

The results of the impact of the mean reversion parameter μ on the empirical mean reversion time are succinctly presented in Table 4.1, which supports our initial hypothesis

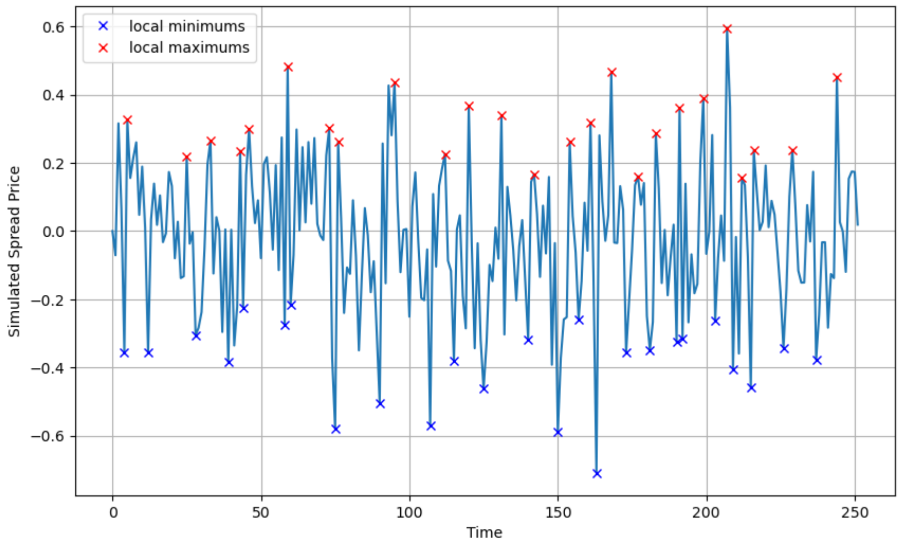


Figure 4.1. Local extremes calculated on a simulated OU spread with $\theta = 10$, $\theta = 0$ and $\sigma = 1$.

Table 4.1. Variation of average empirical mean reversion time (EMRT) with parameter μ in the Ornstein-Uhlenbeck Process.

Parameter μ	Average EMRT	Parameter μ	Average EMRT
2.0	98.79	12.0	49.22
4.0	83.45	14.0	45.10
6.0	78.09	16.0	38.04
8.0	59.22	18.0	35.63
10.0	58.51	20.0	31.15

by demonstrating a clear inverse relationship between the mean reversion speed, μ, and the empirical mean reversion time. As μ increases, the mean reversion time decreases, indicating a faster adjustment of the process back to its mean level. This finding confirms our hypothesis that the empirical mean reversion time reflects the mean-reverting speed of financial time series.

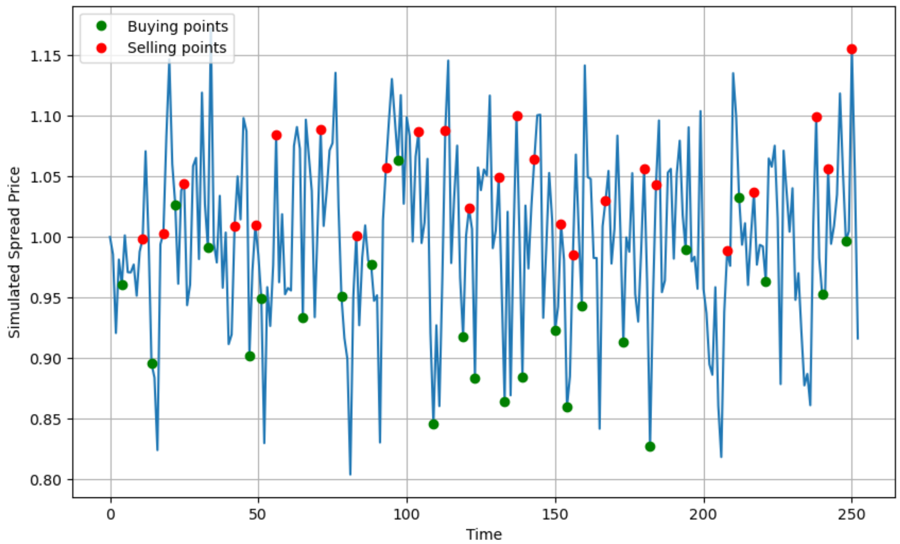


Figure 4.2. Simulated trading on OU process based on reinforcement learning framework.

4.3.2 RL Trading on Simulated Data

In this part, we introduce a preliminary simulated experiment into the effectiveness of the proposed reinforcement learning strategy tailored for mean reversion trading. With the parameters fixed at $\mu = 1$, $\theta = 1$, and $\sigma = 0.1$, we simulate 10,000 paths of the Ornstein-Uhlenbeck process. Each path is designed to reach a terminal time of $T = 252$ and includes $n = 252$ data points, effectively simulating one year of data for use as training samples in our study. This process was chosen due to its relevance in modeling mean-reverting financial instruments, thereby providing a realistic and challenging environment for training our reinforcement learning model.

Our reinforcement learning (RL) model configuration employs a lookback window size of $l = 4$, generating 16 distinct states. Hyper-parameters are set with a learning rate of 0.1, a discount factor of 0.99, an epsilon of 0.1 for the epsilon-greedy strategy and 10 training episodes.

Following the training phase, we apply the trained model to a new, distinct OU sample path to evaluate its decision-making power. This simulated trading are visually presented in Figure 4.2, where buy and sell actions executed by the RL agent are denoted by green and red points, respectively. The outcome demonstrates that the trained agent is capable of executing a series of strategic buy and sell decisions.

Subsequently, we evaluate the trained reinforcement learning model on 100 new samples, calculating the average accumulated profits across these samples. An initial investment of 100 dollars is allocated to each new OU sample. At each purchase point recommended by the RL agent, the entire available cash is used to take a long position on the spread. These positions are then closed at the selling points suggested by the agent. The simulation results in an average profit exceeding 600% across these 100 paths, underscoring the model's adeptness at identifying and leveraging trading opportunities following its training period.

4.3.3 Real World Experiments

In this section, we conduct real-world experiments on S&P500 to evaluate the performance of our proposed strategy, comparing it against established benchmarks such as the classic distance method (DM) [1] and the Ornstein-Uhlenbeck (OU) mean reversion trading strategy [2], [4].

Benchmarks

We begin by introducing the details of the implementation of these benchmark strategies. For the distance method [1], the initial step involves calculating the sum of squared deviations among all potential pairs' normalized price series. This is followed by the identification and selection of pairs of securities that yield the minimum sum of squared deviations. Subsequently, a mean-reverting spread is constructed, denoted as $X = S_1 - S_2$, where S_1 and S_2 represent two analogous stocks. In this context, a long position is assumed in one security of the pair, while a short position is taken in the other. Additionally, estimates of the long-term mean and standard deviations are determined during the formation period.

Transitioning to the trading phase of distance method, the strategy prescribes initiating a long position when the spread's price deviation falls below multiples of estimated standard deviations from the long-term mean. The trade is then exited upon the subsequent reversion of prices. To be more clear, we clarify the trading criterion that we use in our experiment:

- buy to open if $X_t - \bar{x} < -k \cdot s$
- close long position if $X_t - \bar{x} > k \cdot s$

where $\bar{x}$ and s are the sample mean and standard deviance estimated from the formation period. The threshold parameter k is set to 1 in our experiment, Note that we solely focuses on the initial engagement in a long position within the portfolio.

For the OU mean reversion trading strategy [2], [4], we also select the pairs of securities that yield the minimum sum of squared deviations. The next step consists of constructing mean-reverting spreads for the pairs, denoted as $X = S_1 - B \cdot S_2$, where S_1 and S_2 represent two analogous stocks and B is determined by maximizing the likelihood score of fitting the spread to an OU process. Furthermore, the parameters of the spreads are estimated as an OU process, including the mean reversion speed $\hat{\mu}$, the long-term mean $\hat{\theta}$, and the volatility $\hat{\sigma}$.

In the OU trading phase, the equilibrium variance is calculated as $\hat{\sigma}_{\text{eq}} = \frac{\hat{\sigma}}{\sqrt{2\hat{\mu}}}$, according to Avellaneda and Lee (2010). The basic trading signals are based on the estimations of the OU parameters:

- buy to open if $X_t - \hat{\theta} < -k \cdot \hat{\sigma}_{\text{eq}}$
- close long position if $X_t - \hat{\theta} > k \cdot \hat{\sigma}_{\text{eq}}$

where k represents the cutoff value and we set it to 0.5 in our experiment. The trading remains exclusively on initially entering a long position in the portfolio.

Data

Our experimental framework is designed to encompass a one-year formation period, subsequently followed by a trading period spanning the subsequent year. We utilize the daily adjusted closing prices of representative stocks from different sectors within the U.S. market to construct mean reversion spreads. Our selection includes pairs such as MSFT-GOOGL from Technology, CVS-JNJ from Healthcare, CL-KMB from Consumer Goods, V-MA from Financials, GE-BA from Industrials, OXY-XOM from Energy, WELL-VTR from Real Estate, PPG-SHW from Materials, VZ-TMUS from Telecommunication, and CSX-NSC from

Transportation. Data on the daily closing prices for these stocks was collected over the period from January 1, 2022, to December 31, 2023. The data was sourced from the Yahoo! Finance API.

Following the completion of the annual trading, we will collate and analyze the data to calculate the trading performance across various sectors. This process is designed to rigorously evaluate the efficacy of our strategy across different market segments, thereby demonstrating its consistency and adaptability in the face of financial market uncertainties.

Experimental Results

In the forthcoming part, we delve into a detailed analysis of a one-year study, with 2022 designated as the formation period and 2023 as the trading period. During the formation phase, we construct mean reversion portfolios denoted as $X = S_1 - B \cdot S_2$, where S_1 and S_2 symbolize the first and second stocks, respectively, as listed above. Here, B represents a positive coefficient tailored to each trading strategy. Specifically, for the Distance Method, this coefficient is uniformly set to 1 across all pairs. In contrast, for OU pairs trading, B is determined by optimizing the likelihood score of the pair's fit to an Ornstein-Uhlenbeck (OU) process. Within our proposed methodology, B is calibrated by aiming to minimize the empirical mean reversion time of the spread. Table 4.2 compiles the pairs trading coefficient B for each selected pair's mean reversion portfolio, comparing the benchmarks with our novel method.

Figure 4.3 presents the evolution of total wealth throughout the year 2023, offering an initial comparison of trading performance between the proposed method and established baselines. This visualization provides a preliminary insight into the efficacy of our strategy relative to conventional benchmarks.

The trading performance of our proposed method in comparison to established benchmarks is detailed in Tables 4.3 and 4.4. We present a comprehensive set of performance metrics including daily returns (DailyRet), daily standard deviation (DailyStd), daily Sharpe Ratio (DailySR), annual maximum drawdown (MaxDD), and the annual cumulative profit and loss (CumulPnL).

Table 4.2. Comparison of pairs coefficients derived from various methods.

	Pairs Trading Coefficient B		
Pairs Index	DM	OU	EMRT
MSFT-GOOGL	1.0	0.99	0.89
CVS-JNJ	1.0	0.43	-0.24
CL-KMB	1.0	0.39	0.46
V-MA	1.0	0.53	0.33
GE-BA	1.0	0.20	0.34
OXY-XOM	1.0	0.77	0.22
WELL-VTR	1.0	0.99	0.98
PPG-SHW	1.0	0.33	0.12
VZ-TMUS	1.0	0.10	0.01
CSX-NSC	1.0	0.12	0.14

Our experimental results demonstrate that the proposed reinforcement learning approach significantly outperforms traditional benchmarks in terms of daily Sharpe Ratio and cumulative returns, thereby evidencing its effectiveness and robustness in executing mean reversion trading strategies across diverse market sectors. This distinct out-performance underscores the potential benefits of incorporating reinforcement learning techniques into mean reversion trading frameworks. A pivotal factor in achieving such success is the careful design of the reinforcement learning framework, tailored to align with the specific nuances and challenges of the financial applications.

4.4 Conclusions

This study has presented a novel approach to statistical arbitrage by integrating a model-free framework with reinforcement learning techniques. By establishing an empirical mean reversion time metric and optimizing asset coefficients to minimize this duration, our work has substantially refined the process of constructing mean reversion spreads. Furthermore, we have formulated a reinforcement learning framework for the trading phase, carefully designing the state space to encapsulate the recent trends in price movements and the reward functions to align with the distinct attributes of mean reversion trading. The empirical analysis

Table 4.3. Performance summary for trading mean reversion portfolios by baselines and the proposed RL method.

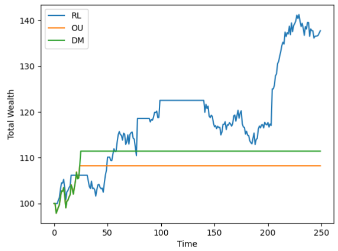
Index	MSFT-GOOGL	CVS-JNJ	CL-KMB	V-MA	GE-BA
			DM Method		
DailyRet (%)	0.0446	0.0440	0.0659	-0.0244	0.2771
DailyStd (%)	0.4670	2.0692	1.2314	1.1796	3.9356
DailySR	0.0955	0.0213	0.0535	-0.0207	0.0704
MaxDD (%)	-2.1344	-24.6778	-13.6791	-10.7238	-18.1823
CumulPnL (%)	11.4443	5.7581	15.6385	-7.4888	64.2387
			OU Method		
DailyRet (%)	0.0327	-0.0073	0.0198	0.0342	0.0392
DailyStd (%)	0.4285	1.4950	0.7589	0.3475	0.3890
DailySR	0.0764	-0.0049	0.0261	0.0985	0.1007
MaxDD (%)	-2.1427	-25.6665	-5.7253	-1.0185	0.000
CumulPnL (%)	8.2443	-4.5179	4.298	8.7348	10.046
			RL Method		
DailyRet (%)	0.1344	0.0585	0.0826	0.0330	0.1679
DailyStd (%)	1.0754	0.7506	0.6000	0.3144	1.2803
DailySR	0.1250	0.0780	0.1377	0.1049	0.1312
MaxDD (%)	0.0000	0.0000	-1.9476	-0.6211	0.0000
CumulPnL (%)	37.7555	14.8895	22.2879	8.4248	48.8196

conducted over the several sectors within the US market has underscored the proposed method's effectiveness and consistency.

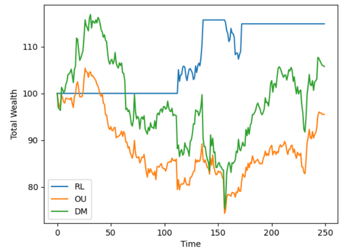
For future work, we aim to explore more sophisticated reinforcement learning algorithms to further optimize trading strategies. This will include the application of deep reinforcement learning and exploration of various reward structures to enhance strategy performance.

Table 4.4. Performance summary for trading mean reversion portfolios by baselines and the proposed RL method.

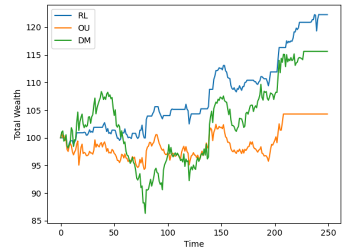
Index	OXY-XOM	WELL-VTR	PPG-SHW	VZ-TMUS	CSX-NSC
			DM Method		
DailyRet (%)	0.0373	0.0694	-0.0772	-0.112	0.0000
DailyStd (%)	2.0950	0.6555	1.0113	3.7311	0.0000
DailySR	0.0178	0.1058	-0.0764	-0.0300	0.0000
MaxDD (%)	-19.1535	-1.4004	-19.1196	-37.4779	0.0000
CumulPnL (%)	3.9194	18.2114	-18.5547	-36.1756	0.0000
			OU Method		
DailyRet (%)	0.0238	0.0539	0.0000	-0.0123	0.0199
DailyStd (%)	1.6012	0.5480	0.0000	1.3241	0.2879
DailySR	0.0149	0.0983	0.0000	-0.0093	0.0693
MaxDD (%)	-14.5001	-1.3798	0.0000	-18.652	0.0000
CumulPnL (%)	2.7812	13.9245	0.0000	-5.0869	4.9825
			RL Method		
DailyRet (%)	0.0609	0.0745	0.1124	0.0412	0.0496
DailyStd (%)	0.9446	0.6794	1.0600	0.8037	0.7101
DailySR	0.0861	0.1097	0.1061	0.0513	0.0698
MaxDD (%)	-2.0008	-3.2895	0.0000	-3.7306	0.0000
CumulPnL (%)	15.0791	19.6910	30.4559	9.9163	12.4263



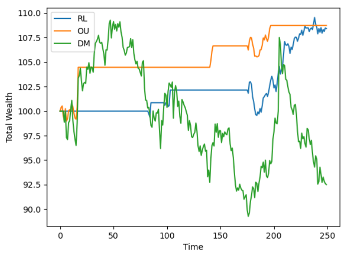
(a) MSFT-GOOGL



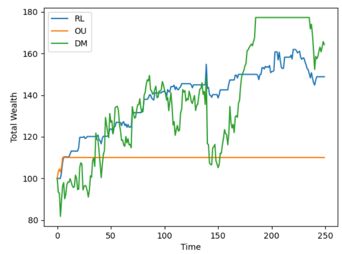
(b) CVS-JNJ



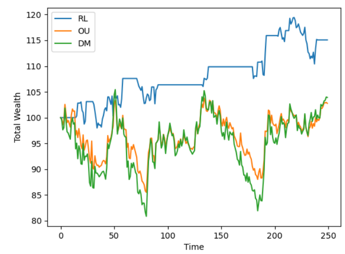
(c) CL-KMB



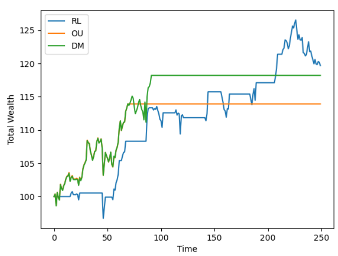
(d) V-MA



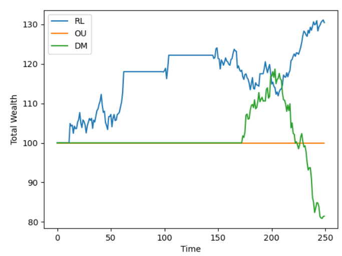
(e) GE-BA



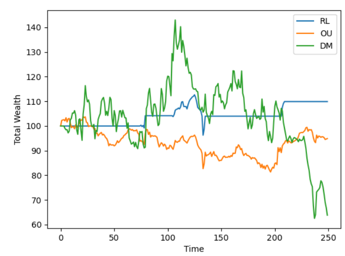
(f) OXY-XOM



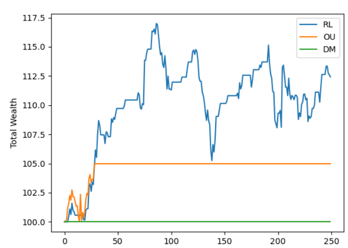
(g) WELL-VTR



(h) PPG-SHW



(i) VZ-TMUS



(j) CSX-NSC

Figure 4.3. 2023 total wealth growth for trading mean-reverting portfolios based on benchmarks and the proposed RL method with initial investment \$100.

5. SUMMARY

This dissertation has explored the intricate domain of statistical arbitrage, a financial strategy leveraging the mean-reverting property of asset price spreads. Through rigorous quantitative analysis and innovative methodological advancements, this work has significantly contributed to the statistical arbitrage literature, particularly in the development of diversified trading strategies, the application of signature-based optimal entry and exit points, and the integration of reinforcement learning techniques.

5.1 Overview of Methods

Chapter 2 introduced a novel framework for constructing diversified portfolios through multiple pairs trading strategies. By dynamically allocating capital among different pairs based on the statistical characteristics of the historical spreads, we demonstrated the framework's ability to enhance trading opportunities and manage risks more effectively. Monte Carlo simulations underscored the framework's potential, showing significant benefits from diversification and dynamic re-balancing based on mean reversion characteristics.

Chapter 3 delved into optimizing the timing for entering and exiting trades in price spreads displaying mean-reverting behavior. Utilizing a sequential optimal stopping framework, we applied a refined signature optimal stopping method, identifying precise entry and exit points to maximize gains. Numerical experiments, including simulations and real-market data analysis, validated the superior performance of our signature optimal mean reversion trading strategy over traditional trading rules.

In Chapter 4, we presented a refined reinforcement learning-based framework for statistical arbitrage. This approach deviated from traditional strategies by constructing a state space that encapsulates recent price trends and a tailored reward function, reflecting the unique dynamics of mean reversion trading. Experimental results from simulated and real-world data demonstrated the efficacy of our reinforcement learning model in outperforming existing benchmarks.

5.2 Contributions and Future Directions

This dissertation contributes to the statistical arbitrage field by presenting three novel approaches: a diversification framework for trading multiple pairs, an optimal entry and exit strategy based on signature theory, and a reinforcement learning framework for advanced trading strategies. These methodologies not only offer theoretical advancements but also practical tools for traders and financial analysts.

Future research could extend this work in several directions. First, exploring the applicability of these strategies across different asset classes and market conditions could provide deeper insights into their robustness and adaptability. Additionally, integrating macroeconomic factors and market sentiment analysis could enhance the predictive accuracy of the trading strategies. Lastly, advancements in computational finance and machine learning offer promising avenues for further refining these models, potentially unlocking new strategies and improving trade execution.

In conclusion, this dissertation advances the field of statistical arbitrage by developing quantitative methods that leverage the mean-reverting nature of financial markets. The frameworks and strategies introduced herein pave the way for more sophisticated and effective trading systems, contributing to the ongoing evolution of quantitative finance.

REFERENCES

[1] E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst, “Pairs trading: Performance of a relative-value arbitrage rule,” *Review of Financial Studies*, vol. 19, no. 3, pp. 797–827, 2006.

[2] M. Avellaneda and J.-H. Lee, “Statistical arbitrage in the us equities market,” *Quantitative Finance*, vol. 10, no. 7, pp. 761–782, 2010.

[3] G. Montana and K. Triantafyllopoulos, “Dynamic modeling of mean reverting spreads for statistical arbitrage,” *Computational Management Science*, vol. 8, pp. 23–49, 2011.

[4] T. Leung and X. Li, *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications* (Modern Trends in Financial Engineering). World Scientific Publishing Company, 2016.

[5] M. J. Brennan and E. S. Schwartz, “Arbitrage in stock index futures,” *Journal of Business*, vol. 63, no. 1, S7–S31, 1990.

[6] M. Dai, Y. Zhong, and Y. K. Kwok, “Optimal arbitrage strategies on stock index futures under position limits,” *Journal of Futures Markets*, vol. 31, no. 4, pp. 394–406, 2011.

[7] T. Kanamura, S. Rachev, and F. Fabozzi, “A profit model for spread trading with an application to energy futures,” *The Journal of Trading*, vol. 5, no. 1, pp. 48–62, 2010.

[8] T. Leung and H. Nguyen, “Constructing cointegrated cryptocurrency portfolios for statistical arbitrage,” *Studies in Economics and Finance*, 2019.

[9] T. G. Fischer, C. Krauss, and A. Deinert, “Statistical arbitrage in cryptocurrency markets,” *Journal of Risk and Financial Management*, vol. 12, no. 1, p. 31, 2019.

[10] B. Do and R. Faff, “Are pairs trading profits robust to trading costs?” *Journal of Financial Research*, vol. 35, no. 2, pp. 261–287, 2012.

[11] B. Do and R. Faff, “Does simple pairs trading still work?” *Financial Analysts Journal*, vol. 66, no. 4, pp. 83–95, 2010.

[12] H. Chen, S. Chen, Z. Chen, and F. Li, "Empirical investigation of an equity pairs trading strategy," *Management Science*, vol. 65, no. 1, pp. 370–389, 2019.

[13] M. S. Perlin, "Evaluation of pairs-trading strategy at the brazilian financial market," *Journal of Derivatives & Hedge Funds*, vol. 15, pp. 122–136, 2009.

[14] S. Andrade, V. Di Pietro, and M. Seasholes, "Understanding the profitability of pairs trading," *Unpublished working paper, UC Berkeley, Northwestern University*, 2005.

[15] J. Engelberg, P. Gao, and R. Jagannathan, "An anatomy of pairs trading: The role of idiosyncratic news, common information and liquidity," in *Third Singapore International Conference on Finance*, 2009.

[16] H. Jacobs and M. Weber, "On the determinants of pairs trading profitability," *Journal of financial markets*, vol. 23, pp. 75–97, 2015.

[17] N. Huck, "The high sensitivity of pairs trading returns," *Applied Economics Letters*, vol. 20, no. 14, pp. 1301–1304, 2013.

[18] N. Huck and K. Afawubo, "Pairs trading and selection methods: Is cointegration superior?" *Applied Economics*, vol. 47, no. 6, pp. 599–613, 2015.

[19] D. Bowen, M. C. Hutchinson, and N. OSullivan, *High frequency equity pairs trading: transaction costs, speed of execution and patterns in returns*. SSRN, 2019.

[20] G. Vidyamurthy, *Pairs Trading: quantitative methods and analysis*. John Wiley & Sons, 2004, vol. 217.

[21] R. F. Engle and C. W. Granger, "Co-integration and error correction: Representation, estimation, and testing," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987.

[22] A. Galenko, E. Popova, and I. Popova, "Trading in the presence of cointegration," *The Journal of Alternative Investments*, vol. 15, no. 1, pp. 85–97, 2012.

[23] H. Puspaningrum, Y.-X. Lin, and C. M. Gulati, "Finding the optimal pre-set boundaries for pairs trading strategy based on cointegration technique," *Journal of Statistical Theory and Practice*, vol. 4, pp. 391–419, 2010.

[24] C. Dunis, J. Laws, and B. Evans, "Trading futures spreads: An application of correlation and threshold filters," *Applied Financial Economics*, vol. 16, no. 12, pp. 903–914, 2006.

[25] S. Broumandi and T. Reuber, "Statistical arbitrage and fx exposure with south american adrs listed on the nyse," *Financial Assets and Investing*, vol. 3, no. 2, pp. 5–18, 2012.

[26] J. Caldeira and G. V. Moura, "Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy," *Available at SSRN 2196391*, 2013.

[27] M. L. Li, C. M. Chui, and C. Q. Li, "Is pairs trading profitable on china ah-share markets?" *Applied Economics Letters*, vol. 21, no. 16, pp. 1116–1121, 2014.

[28] R. Elliott, J. Van Der Hoek, and W. Malcolm, "Pairs trading," *Quantitative Finance*, vol. 5, no. 3, pp. 271–276, 2005.

[29] B. Do, R. Faff, and K. Hamza, "A new approach to modeling and estimation for pairs trading," in *Proceedings of 2006 financial management association European conference*, Citeseer, vol. 1, 2006, pp. 87–99.

[30] K. Triantafyllopoulos and G. Montana, "Dynamic modeling of mean-reverting spreads for statistical arbitrage," *Computational Management Science*, vol. 8, pp. 23–49, 2011.

[31] J. Berkowitz and L. Kilian, "Recent developments in bootstrapping time series," *Econometric Reviews*, vol. 19, no. 1, pp. 1–48, 2000.

[32] M. Cummins and A. Bucca, "Quantitative spread trading on crude oil and refined products markets," *Quantitative Finance*, vol. 12, no. 12, pp. 1857–1875, 2012.

[33] Z. Zeng and C.-G. Lee, "Pairs trading: Optimal thresholds and profitability," *Quantitative Finance*, vol. 14, no. 11, pp. 1881–1893, 2014.

[34] M. Bock and R. Mestel, "A regime-switching relative value arbitrage rule," in *Operations Research Proceedings 2008: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR) University of Augsburg, September 3-5, 2008*, Springer, 2009, pp. 9–14.

[35] T. Kanamura, S. T. Rachev, and F. J. Fabozzi, “A profit model for spread trading with an application to energy futures,” KIT Working Paper Series in Economics, Tech. Rep., 2011.

[36] M. Boguslavsky and E. Boguslavskaya, “Arbitrage under power,” *Risk*, vol. 17, no. 6, pp. 69–73, 2004.

[37] C. W. Chen, M. Chen, and S.-Y. Chen, “Pairs trading via three-regime threshold autoregressive garch models,” in *Modeling Dependence in Econometrics: Selected Papers of the Seventh International Conference of the Thailand Econometric Society, Faculty of Economics, Chiang Mai University, Thailand, January 8-10, 2014*, Springer, 2014, pp. 127–140.

[38] S. Mudchanatongsuk, J. A. Primbs, and W. Wong, “Optimal pairs trading: A stochastic control approach,” in *2008 American control conference*, IEEE, 2008, pp. 1035–1039.

[39] E. Ekström, C. Lindberg, and J. Tysk, “Optimal liquidation of a pairs trade,” *Advanced mathematical methods for finance*, pp. 247–255, 2011.

[40] “An optimal pairs-trading rule,” *Automatica*, vol. 49, no. 10, pp. 3007–3014, 2013, ISSN: 0005-1098.

[41] C. Lindberg, “Pairs trading with opportunity cost,” *Journal of Applied Probability*, vol. 51, no. 1, pp. 282–286, 2014. DOI: 10.1239/jap/1395771429.

[42] J. Liu and A. Timmermann, “Optimal Convergence Trade Strategies,” *The Review of Financial Studies*, vol. 26, no. 4, pp. 1048–1086, Jan. 2013.

[43] J. Stübinger, B. Mangold, and C. Krauss, “Statistical arbitrage with vine copulas,” *Quantitative Finance*, vol. 18, no. 11, pp. 1831–1849, 2018.

[44] R. Q. Liew and Y. Wu, “Pairs trading: A copula approach,” *Journal of Derivatives & Hedge Funds*, vol. 19, no. 1, pp. 12–30, 2013.

[45] W. Xie, R. Q. Liew, Y. Wu, and X. Zou, “Pairs trading with copulas,” *The Journal of Trading*, vol. 11, no. 3, pp. 41–52, 2016.

[46] C. Krauss and J. Stübinger, “Non-linear dependence modelling with bivariate copulas: Statistical arbitrage pairs trading on the s&p 100,” *Applied Economics*, vol. 49, no. 52, pp. 5352–5369, 2017.

[47] Y. Stander, D. Marais, and I. Botha, “Trading strategies with copulas,” *Journal of Economic and Financial Sciences*, vol. 6, no. 1, pp. 83–107, 2013.

[48] O. Bartkoviak, V. Shpyrko, O. Chernyak, and Y. Chernyak, “Statistical arbitrage using cointegration and principal component analysis approach,” in *Business Development and Economic Governance in Southeastern Europe: 13th International Conference on the Economies of the Balkan and Eastern European Countries (EBEEC), Pafos, Cyprus, 2021*, Springer, 2022, pp. 167–182.

[49] J. Guijarro-Ordonez, M. Pelger, and G. Zanotti, “Deep learning statistical arbitrage,” *Available at SSRN 3862004*, 2021.

[50] G. Vergara and W. Kristjanpoller, “Deep reinforcement learning applied to statistical arbitrage investment strategy on cryptomarket,” *Applied Soft Computing*, vol. 153, p. 111 255, 2024.

[51] Y. Zhao, S. Xu, and J. Ossowski, “Deep learning meets statistical arbitrage: An application of long short-term memory networks to algorithmic trading.,” *Journal of Financial Data Science*, vol. 4, no. 4, 2022.

[52] N. Huck, “Large data sets and machine learning: Applications to statistical arbitrage,” *European Journal of Operational Research*, vol. 278, no. 1, pp. 330–342, 2019.

[53] B. Zhan, S. Zhang, H. S. Du, and X. Yang, “Exploring statistical arbitrage opportunities using machine learning strategy,” *Computational Economics*, vol. 60, no. 3, pp. 861–882, 2022.

[54] C. Krauss, X. A. Do, and N. Huck, “Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the s&p 500,” *European Journal of Operational Research*, vol. 259, no. 2, pp. 689–702, 2017.

[55] A. d’Aspremont, “Identifying small mean-reverting portfolios,” *Quantitative Finance*, vol. 11, no. 3, pp. 351–364, 2011.

[56] T. Leung, J. Zhang, and A. Aravkin, “Sparse mean-reverting portfolios via penalized likelihood optimization,” *Automatica*, vol. 111, p. 108 651, 2020.

[57] T. N. Li and A. Papanicolaou, “Statistical arbitrage for multiple co-integrated stocks,” *Applied Mathematics & Optimization*, vol. 86, no. 1, p. 12, 2022.

[58] Q. Song, G. Yin, and Q. Zhang, “Stochastic optimization methods for buying-low-and-selling-high strategies,” *Stochastic Analysis and Applications*, vol. 27, no. 3, pp. 523–542, 2009.

[59] T. Leung and X. Li, “Optimal mean reversion trading with transaction costs and stop-loss exit,” *International Journal of Theoretical and Applied Finance*, vol. 18, no. 03, p. 1 550 020, 2015.

[60] D. Lee and T. Leung, “On the efficacy of optimized exit rule for mean reversion trading,” *International Journal of Financial Engineering*, vol. 7, no. 03, p. 2 050 024, 2020.

[61] T. Leung and Y. Zhou, “Dynamic optimal futures portfolio in a regime-switching market framework,” *International Journal of Financial Engineering*, vol. 6, no. 4, p. 1 950 034, 2019.

[62] C. Bayer, P. Hager, S. Riedel, and J. Schoenmakers, “Optimal stopping with signatures,” *arXiv preprint arXiv:2105.00778*, 2021.

[63] B. Hambly and T. Lyons, “Uniqueness for the signature of a path of bounded variation and the reduced path group,” *Annals of Mathematics*, pp. 109–167, 2010.

[64] J. Yoon, D. Jarrett, and M. Van der Schaar, “Time-series generative adversarial networks,” *Advances in neural information processing systems*, vol. 32, 2019.

[65] E. Fink and H. S. Gandhi, “Important extrema of time series,” 2007.