Finding Moving-Band Statistical Arbitrages via Convex-Concave Optimization

Kasper Johansson Thomas Schmelzer Stephen Boyd

December 13, 2023

Abstract

We propose a new method for finding statistical arbitrages that can contain more assets than just the traditional pair. We formulate the problem as seeking a portfolio with the highest volatility, subject to its price remaining in a band and a leverage limit. This optimization problem is not convex, but can be approximately solved using the convex-concave procedure, a specific sequential convex programming method. We show how the method generalizes to finding moving-band statistical arbitrages, where the price band midpoint varies over time.

Contents

1 Introduction

We consider the problem of finding a *statistical arbitrage* (stat-arb), *i.e.*, a portfolio with mean-reverting price. Roughly speaking, this means that the price of the portfolio stays in a band, and varies over it. Such a stat-arb is traded in the obvious way, buying when the price is in the low part of the band and selling when it is in the high part of the band.

Traditional stat-arbs focus on portfolios consisting of two or possibly three underlying assets. When the portfolio contains two assets, trading the stat-arb is called *pairs trading*. Pairs can be found by exhaustive search over all $n(n-1)/2$ pairs of assets in a universe of n assets. When the weights of the two assets in pairs trading are $+1$ and -1, the portfolio value is the *spread* (between the two prices). The assets in a pair are called *co-moving assets*.

In this paper we propose a new method for finding stat-arbs that can contain multiple (more than two) assets, with general weights. The problem is formulated as a nonconvex optimization problem in which we maximize the portfolio price variation subject to the price staying within a fixed band, along with a leverage limit, over some training period. Although this approach requires maximizing a convex function, we show how to approximately solve it using the convex-concave procedure [SDGB16, LB16].

Our second contribution is to introduce the concept of a *moving-band stat-arb*, in which the price of the portfolio varies in a band that changes over time, centered at the recent average portfolio price. (We refer to a traditional stat-arb as a *fixed-band stat-arb*.) We show that the same method we use to find fixed-band stat-arbs can be used to find moving-band stat-arbs, despite the apparent complexity of the average price also depending on the portfolio. Moving-band stat-arbs are traded in the same obvious way as fix-band stat-arbs, buying when the price is in the low part of the band and selling when it is in the high part of the band; but with moving-band stat-arbs, the center of the band changes over time. Moving-band stat-arb trading resembles trading using *Bollinger bands* [Bol02, Bol92], except that the bands are associated with the price of a carefully constructed portfolio, and not a single asset. Our empirical studies show that moving-band stat-arbs out-perform fixed-band stat-arbs in terms of profit, and remain profitable for longer out-of-sample periods.

1.1 Related work

Stat-arbs. Stat-arb trading strategies date back to the 1980s when a group of Morgan Stanley traders, led by Nunzio Tartaglia, developed pairs trading [Pol11, GGR06]. This strategy involves identifying pairs of assets whose price tends to move together, hence referred to as pairs trading. In pairs trading, the spread, *i.e.*, price difference between two assets, is tracked and positions entered when this difference deviates from its mean. This trading strategy has enjoyed widespread popularity, with its success substantiated by numerous empirical studies in various markets like equities [AL10], commodities [Nak19, VM17], and currencies [FKD19]; see, *e.g.*, [GGR06, AL10, Per09, HJTW04, KDH17, CM13, Huc19, DGLR10].

In the general setting, a stat-arb consists of multiple assets in a portfolio that exhibits a mean-reverting behavior [FP16, §10.5]. Stat-arb trading is a widely used strategy in quantitative finance. The literature on stat-arbs is extensive and generally splits into several categories: finding stat-arbs, modeling the (mean-reverting) portfolio price, and trading stat-arbs. We give a brief review of these here and refer the reader to [Kra17] for a comprehensive overview of the literature.

Finding stat-arbs. Probably the simplest approach to finding pairs of co-moving securities is the distance approach. The distance pairs trading strategy finds assets whose (normalized) prices have moved closely historically, in an exhaustive search through pairs of assets [GGR06]. Assets whose prices have a low sum of squared deviation from each other are considered for trading. The distance approach is simple and intuitive, although it does not necessarily find good pairs [Kra17]. The objective itself is to minimize the distance between two asset prices, which does not directly relate to the desired properties of a stat-arb, which crucially should also have a high variance. This paper addresses this issue by directly optimizing for large fluctuations around the mean.

The co-integration approach is another popular method for finding co-moving securities. Co-integration is an important concept in the econometrics literature [Joh00, AD05], and dates back to Engle and Granger's works in the 1980s (for which Granger was awarded the 2003 Nobel Memorial Prize in Economic Sciences) [Gra83, EG87]. The idea is that the absence of stationarity in a multivariate time series may be explained by common trends, which would make it possible to find linear combinations of assets that are stationary and hence mean-reverting. Thus, the co-integration approach is based on identifying linear combinations of assets that result in a stationary time series [Kra17]. In [Vid04], the most cited work on co-integration based pairs trading, potential asset pairs are found based on statistical measures, which are then tested for co-integration using an adapted version of the Engle-Granger test. Several co-integration based methods have been proposed to extend the pairs trading strategy to more than two assets. For example, in [ZP18, ZZWP18, ZZP19] the authors consider a (non-convex) optimization problem for finding high variance, mean-reverting portfolios. Their strategy is based on finding a portfolio of spreads, defined by a co-integration subspace, and implemented using sequential convex optimization. Their proposed optimization problem, *i.e.*, maximizing variance subject to a mean reversion criterion is similar to our problem formulation. However, our problem differs significantly in that we do not rely on any co-integration analysis or statistical testing. Rather, we directly optimize for a high variance portfolio that is mean-reverting.

Asset pairs can also be found using machine learning methods. In [SH20], the authors use unsupervised learning and propose a density-based clustering algorithms to cluster assets. Then, within asset clusters, pairs of assets are chosen for trading depending on co-integration, as well as mean-reversion tendency and frequency. Modern machine learning methods are also explored in [KDH17], where the authors propose the use of deep neural networks, gradient-boosted trees, and random forests for finding stat-arb portfolios. Another recent study of deep-learning stat-arb finding is [GOPZ21]. Earlier work on using machine learning for finding stat-arbs includes, *e.g.*, [DKB15, MZ14, TL13, Huc10, Huc09].

Modeling the stat-arb spread. When a co-moving set of assets has been identified, the next step is to model the portfolio price (or spread between the assets for a pair). Perhaps the most popular approach is to model the spread using stochastic control theory. It is common to consider investments in a mean-reverting spread and a risk-free asset and to model the spread as an Ornstein-Uhlenbeck process [MPW08, JY07]. Other methods include those borrowing tools from time series analysis [Kra17]. In [EVDHM05] the authors propose a mean-reverting Gaussian Markov chain model for modeling the spread between two assets. Copulas have also been proposed to model the joint distribution of the spread, both for pairs and for larger sets of assets [DMPZ16, SMK18, KS17]. In [DMPZ16] the authors suggest modeling the spread using linear state space models.

Trading stat-arbs. We mention here a number of stat-arb trading methods, ranging from simple ones based on the intuitive idea of buying when the price is low and selling when it is high, to more complex ones based on learning the price statistics and using stochastic control. One simple method is based on *hysteresis*, as is used in a conventional thermostat. In this approach we buy (enter into a long position) when the price drops below a threshold, and sell (switch to a short position) when the price goes above another threshold. The thresholds are typically based on price bands, as discussed below. A variation on this method sets the thresholds based on the standard deviation of the price, as proposed in [GGR06]. Another simple method is *linear trading*, where we take a position proportional to the difference between the band midpoint price and the current price. This method can also be modified to use volatility-based bands instead of fixed bands. (Such a trading policy is a simple Markowitz policy, with the mean return given by the difference between the midpoint price and the current price.)

Other methods for trading stat-arbs follow directly from the spread models described above. When the spread is modeled as an Ornstein-Uhlenbeck process, the optimal trading strategy is found by solving a stochastic control problem [MPW08, JY07, Ber10]. In [YP12, PY18, YP18] the authors model the spread as an autoregressive process, a discretization of the Ornstein-Uhlenbeck process, and show how to trade a portfolio of spreads under proportional transaction costs and gross exposure constraints using model predictive control. With the copula approaches of [DMPZ16, SMK18, KS17], the trading strategy is based on deviations from confidence intervals. The machine learning methods of [SH20, KDH17, GOPZ21, DKB15, MZ14, TL13, Huc10, Huc09] also include trading strategies.

Exiting a stat-arb. A stat-arb will not keep its mean-reverting behavior forever. Hence, a strategy for exiting a stat-arb, *i.e.*, closing the position, is needed. One approach is to exit when the spread reaches a certain threshold in terms of its standard deviation or in terms of a price-band, as described below. Another simple method is to exit after a fixed time-period, as was done in [GGR06].

A simple variation on any of these exit methods does not exit the position immediately when the exit condition is first satisfied. Instead it reduces the position to zero slowly (*e.g.*, linearly) over some fixed number of periods.

Price-bands. Price-bands are popular in technical analysis, and are used to identify trading opportunities based on the price of an asset relative to its recent price history [LLP21]. The most popular is the Bollinger band [Bol92]. It is constructed by computing the M-day moving average (with common choice $M = 21$) of the asset price, denoted μ_t at time t, and the corresponding price standard deviation σ_t. The Bollinger band is then defined as the interval $[\mu_t - k\sigma_t, \mu_t + k\sigma_t]$, where $k > 0$ is a parameter, typically taken to be $k = 2$. A price signal is extracted based on if the price is near the top or bottom of the band. For a detailed description of the Bollinger band and other trading-bands, we refer the reader to [Bol02, Bol92].

1.2 This paper

This paper proposes a new method for finding stat-arbs, with two main contributions. The first contribution of our method is to formulate the search for stat-arbs as an optimization problem that intuitively and directly relates to the desired properties of a stat-arb: its price should remain in a band (*i.e.*, be mean-reverting) and also should have a high variance. Since our method is based on convex optimization, it readily scales to large universes of assets. It can and does find stat-arbs with ten or more assets, well beyond our ability to carry out an exhaustive search.

Our second contribution is to introduce the concept of a moving-band stat-arb. While the idea of a price band is widely known and used in technical analysis trading, the main difference is that we apply it to carefully constructed portfolios (*i.e.*, stat-arbs), instead of single assets.

Our focus is on finding both fixed-band and moving-band stat-arbs, and not on trading them. Our numerical experiments use a simple linear trading policy, and a simple time-based exit condition. (We have also verified that similar results are obtained using hysteris-based trading policies.) We also do not address the question of how one might trade a portfolio of stat-arbs, the focus of a forth-coming paper.

1.3 Outline

The rest of the paper is organized as follows. In §2 we propose a new method for finding traditional stat-arbs, *i.e.*, with a fixed band. We extend this method to moving bands in §3. In §4 we present experiments on real data, and in §5 we conclude the paper.

2 Finding fixed-band stat-arbs

We consider a vector time series of prices of a universe of n assets, denoted $P_t \in \mathbf{R}^n$, $t = 1, 2, \ldots, T$, denoted in USD per share. (We presume these are adjusted for dividends and splits.) We consider a portfolio of these assets given by $s \in \mathbf{R}^n$, denoted in shares, with $s_i < 0$ denoting short positions. The price or net value of the portfolio is the scalar time series $p_t = s^T P_t$, $t = 1, \ldots, T$. The asset prices P_t are positive, but the portfolio price p_t

need not be, since the entries of s can be negative. For future use, we define the average price of the n assets as the vector $\bar{P} = (1/T)\sum_{t=1}^{T} P_t$.

We seek a portfolio for which p_t consistently varies over a band (interval of prices) with two goals. It should stay in the price band, and also, vary over the band consistently; that is, it should frequently vary between the high end of the band and the low end of the band. We refer to such a portfolio s as a stat-arb, so we use the term to refer to the general concept as well as a specific portfolio. Our method differs from the traditional statistical framework, where one would seek a portfolio s for which p_t is co-integrated.

2.1 Formulation as convex-concave problem

We formulate the search for a stat-arb s as an optimization problem. The condition that p_t varies over a band is formulated as $-1 \leq p_t - \mu \leq 1$, where μ is the midpoint of the band. Here we fix the width of the price band as 2; s and μ can always be scaled so this holds. We express the desire that p_t vary frequently over the band by maximizing its volatility. Since s, p, and μ can all be multiplied by -1 without any effect on the constraints or objective, we can assume that $\mu \geq 0$ without loss of generality.

We arrive at the problem

$$\begin{array}{ll} \text{maximize} & \sum_{t=2}^{T}(p_t - p_{t-1})^2 \\ \text{subject to} & -1 \leq p_t - \mu \leq 1, \quad p_t = s^T P_t, \quad t = 1, \ldots, T \\ & |s|^T \bar{P} \leq L, \quad \mu \geq 0, \end{array} \tag{1}$$

with variables $s \in \mathbf{R}^n$, $p \in \mathbf{R}^T$, and $\mu \in \mathbf{R}$, where the absolute value in the last constraint is elementwise. The problem data are the vector price time series P_t, $t = 1, \ldots, T$, and the positive parameter L.

Note that $|s|^T \bar{P}$ is the average total position of the portfolio, sometimes called its leverage, so the constraint $|s|^T \bar{P} \leq L$ is a leverage constraint; it limits the total position of the portfolio. The leverage is a weighted ℓ_1 norm of s, and so tends to lead to sparse s, *i.e.*, a portfolio that concentrates in a few assets, a typical desired quality of a stat-arb. The problem (1) is a nonconvex optimization problem, since the objective is a convex function, and we wish to maximize it; we will explain below how we can approximately solve it.

2.2 Interpretation via a simple trading policy

While our formulation of the stat-arb optimization problem (1) makes sense on its own, we can further motivate it by looking at the profit obtained using a simple trading policy. Suppose we hold quantity $q_t \in \mathbf{R}$ of the portfolio, *i.e.*, we hold the portfolio $q_t s \in \mathbf{R}^n$ (in shares). We assume that $q_0 = q_T = 0$, *i.e.*, we start and end with no holdings. In period t we buy $q_t - q_{t-1}$ and pay $p_t(q_t - q_{t-1})$. The total profit is then

$$\sum_{t=1}^{T-1} q_t(p_{t+1} - p_t). \tag{2}$$

We will relate this to our objective in (1) above, with the simple trading policy

$$q_t = \mu - p_t, \quad t = 1, \ldots, T-1, \tag{3}$$

which we refer to as a linear trading policy since the holdings are a linear function of the difference between the band midpoint and the current price. This trading policy holds nothing when $p_t = \mu$, *i.e.*, the price is in the middle of the band. When the price is low, $p_t = \mu - 1$, we hold $q_t = +1$, and when it is high, $p_t = \mu + 1$ we hold $q_t = -1$.

With the simple linear policy (3) and the boundary conditions $q_0 = q_T = 0$, the profit (2) is, after some algebra,

$$\frac{1}{2}\sum_{t=2}^{T}(p_t - p_{t-1})^2 + \frac{(p_1 - \mu)^2 - (p_T - \mu)^2}{2}.$$

The first term is one half our objective. The second term is between $-1/2$ and $1/2$, since $(p_T - \mu)^2$ and $(p_1 - \mu)^2$ are both between 0 and 1. Thus the profit is at least

$$\frac{1}{2}\left(\sum_{t=2}^{T}(p_t - p_{t-1})^2 - 1\right). \tag{4}$$

This shows that our objective is the profit of the simple linear policy (3), scaled by one-half, plus a constant. In particular, if the objective of the problem (1) exceeds one, the simple linear policy makes a profit.

2.3 Solution method

Convex-concave procedure. We solve the problem (1) approximately using sequential convex programming, specifically the convex-concave procedure [SDGB16, LB16]. Let k denote the iteration, with s^k the portfolio and p_t^k the portfolio price in the kth iteration. In each iteration of the convex-concave procedure, we linearize the objective, replacing the quadratic function $f(p) = \sum_{t=2}^{T}(p_t - p_{t-1})^2$ with the affine approximation

$$\hat{f}(p; p^k) = f(p^k) + \nabla f(p^k)^T(p - p^k) = \nabla f(p^k)^T p + c,$$

where c is a constant (*i.e.*, does not depend on p). This linearization is a lower bound on the true objective, *i.e.*, we have $f(p) \geq \hat{f}(p; p^k)$ for all p. For completeness we note that

$$(\nabla f(p))_t = \begin{cases} 2(p_1 - p_2) & t = 1 \\ 2(2p_t - p_{t-1} - p_{t+1}) & t = 2, \ldots, T-1 \\ 2(p_T - p_{T-1}) & t = T. \end{cases}$$

We now solve the linearized problem

$$\begin{array}{ll} \text{maximize} & \hat{f}(p; p^k) \\ \text{subject to} & -1 \leq p_t - \mu \leq 1, \quad p_t = s^T P_t, \quad t = 1, \ldots, T \\ & |s|^T \bar{P} \leq L, \quad \mu \geq 0, \end{array} \tag{5}$$

with variables p_t, s and μ. This is a convex problem, in fact a linear program (LP), and readily solved [BV04]. We take the solution of this problem as the next iterate p^{k+1}, s^{k+1}, μ^{k+1}. This simple algorithm converges to a local solution of (1), typically in at most a few tens of iterations.

Cleanup phase. The leverage constraint $|s|^T \bar{P} \leq L$ encourages sparse solutions, but in some cases the convex-concave procedure converges to a portfolio with a few small holdings. To achieve even sparser portfolios, we can carry out a clean-up step once the convex-concave procedure has converged. We first determine the subset of assets for which s_i is zero or small, as measured by its relative weight in the portfolio, *i.e.*, i for which

$$|s_i|\bar{P}_i \leq \eta |s|^T \bar{P},$$

where η is a small positive constant such as 0.05. We then solve the problem again, this time with the constraint that all such s_i are zero. This takes just a few convex-concave iterations, and can be repeated, which results in sparse portfolios in which every asset has weight at least η.

Implementation. To make the optimization problem better conditioned, we scale the prices P_t so that (after scaling) $\bar{P} = \mathbf{1}$, the vector with all entries one. Thus after scaling, the leverage $|s|^T \bar{P}$ becomes the ℓ_1 norm of s. We also scale the gradient (or objective) to be on the order of magnitude one. These scalings do not affect the solution, but make the method less vulnerable to floating point rounding errors.

Initialization. The final portfolio found by the convex-concave procedure, plus the cleanup phase, depends on the initial portfolio s^1. It can and does converge to different final portfolios for different starting portfolios. With a random initialization we can find multiple stat-arbs for the same universe. (We also get some duplicates when the method converges to the same final portfolio from different initial portfolios.) We have found that uniform initialization of the entries of s^1 in the interval $[0, 1]$ works well in practice. Thus from one universe and data set, we can obtain multiple stat-arbs.

3 Finding moving-band stat-arbs

3.1 Moving-band stat-arbs

In the fixed-band stat-arb problem (1) μ is constant, so the midpoint of the trading band does not vary with time. In this section we describe a simple but powerful extension in which the stat-arb band midpoint changes over time. One simple (and traditional) choice is to define μ_t as the mean of the trailing prices p_t, for example the mean over the last M periods,

$$\mu_t = \frac{1}{M} \sum_{\tau=t-M+1}^{t} p_\tau.$$

(This requires knowledge of the prices $P_0, P_{-1}, \ldots, P_{-M+1}$.) In this formulation, μ_t is also a function of s, the portfolio, but it is a known linear function of it. Any other linear expression for the average recent price could be used, *e.g.*, exponentially weighted moving average (EWMA). A moving-band stat-arb is a portfolio s in which the price p_t stays in a moving band with width two and midpoint μ_t, and also has high variance.

Trading policy. The simple linear trading policy (3) can be modified in the obvious way, as

$$q_t = \mu_t - p_t, \quad t = 1, \ldots, T-1. \tag{6}$$

3.2 Finding moving-band stat-arbs

We arrive at the optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{t=2}^{T} (p_t - p_{t-1})^2 \\ \text{subject to} & -1 \leq p_t - \mu_t \leq 1, \quad p_t = s^T P_t, \quad t = 1, \ldots, T \\ & |s|^T \bar{P} \leq L, \quad \mu_t = (1/M) \sum_{\tau=t-M+1}^{t} p_\tau, \quad t = 1, \ldots, T, \end{array} \tag{7}$$

with variables $s \in \mathbf{R}^n$, $p_1, \ldots, p_T$, and $\mu_1, \ldots, \mu_T$. (The latter two sets of variables are simple linear functions of s.) In this problem we have an additional parameter M, the memory for the band midpoint.

Note that in the fixed-band stat-arb problem (1), μ is a scalar variable that we freely choose; in the moving-band stat-arb problem (7), μ_t varies over time, and is itself a function of s. Despite this complication, the moving-band stat-arb problem (7) can be (approximately) solved using exactly the same convex-concave method as the fixed-band stat-arb problem (1); the only difference is in the convex constraints.

4 Numerical experiments

We illustrate our method with an empirical study on historical asset prices. Everything needed to reproduce the results is available online at

`https://github.com/cvxgrp/cvxstatarb`.

4.1 Experimental setup

Data set. We use daily data of the CRSP US Stock Databases from the Wharton Research Data Services (WRDS) portal [WRD23]. The data set consists of adjusted prices of 15405 stocks from January 4th, 2010, to December 30th, 2023, for a total of 3282 trading days.

Monthly search for stat-arbs. Starting December 23, 2011, and every 21 trading days thereafter, until July 6th, 2022, we use the convex-concave method with 10 different random initial portfolios. From these 10 stat-arbs we add the unique ones, defined by the set of assets in the stat-arb, to our current set of stat-arbs. All together we solve fixed-band and moving-band problems 1270 times.

Parameters. We set $L = \$50$ for the fixed-band stat-arb and $L = \$100$ for the moving-band stat-arb. For the moving-band stat-arb, we take $M = 21$ days, *i.e.*, we use the trailing month average price as the midpoint, a value commonly used for Bollinger bands [Bol92, Bol02]. We note that our results are not very sensitive to the choices of L and M, and that choosing a larger L for the moving-band than for the fixed-band stat-arb is reasonable since we expect the price of a portfolio to vary less around its short-term moving midpoint than around a fixed midpoint.

Trading policy. We use the simple linear trading policies defined in (3) and (6) for the fixed-band and moving-band stat-arbs, respectively. We use a simple time-based exit condition, where we trade a stat-arb for $T^{\max}$ trading days, and then exit the position uniformly (linearly) over the next T^{exit} trading days. This means we take

$$q_t = (1 - \alpha_t)(\mu - p_t), \quad t = T^{\max}, \ldots, T^{\max} + T^{\text{exit}} - 1,$$

where $\alpha_t = (t + 1 - T^{\max})/T^{\text{exit}}$. We use parameter values $T^{\max} = 63$ for the fixed-band stat-arb and $T^{\max} = 125$ for the moving-band stat-arb, and $T^{\text{exit}} = 21$ for both. We also exit a stat-arb if the value of the stat-arb plus a cash account drops below a given level, as described below.

4.2 Simulation and metrics

We simulate a stat-arb as follows.

Cash account. Each stat-arb is initialized with a cash account value

$$C_0 = \nu |s|^T P_0,$$

where ν is a positive parameter and P_0 is the price vector the day before we start trading the stat-arb. We take $\nu = 0.5$ in our experiments. The cash account is updated as

$$C_{t+1} = C_t - (q_{t+1} - q_t)p_{t+1} - \phi_t, \quad t = 1, \ldots, T,$$

where ϕ_t is the transaction and holding cost, consisting of the trading cost at time $t + 1$ and the holding cost over period t, described below. The cash account plus the long position is meant to be the collateral for the short positions.

Portfolio net asset value. The net asset value (NAV) of the portfolio, including the cash account, at time t is then

$$V_t = C_t + q_t p_t.$$

(Note that $V_0 = C_0$.) The profit at time t is

$$V_t - V_{t-1} = q_t p_t + C_t - q_{t-1} p_{t-1} - C_{t-1} = q_t(p_t - p_{t-1}) - \phi_t,$$

which agrees with the profit formula (2), after accounting for transaction and holding costs.

NAV based termination. If the NAV goes below 50% of C_0, we liquidate the stat-arb portfolio. We do this for two reasons. First, we do not want the portfolio to have negative value, *i.e.*, to go bust. Second, this constraint ensures that the short positions are always fully collateralized by the long positions plus the cash account, with a margin of at least half of our initial investment. All of our metrics include early termination stat-arbs.

Trading and shorting costs. Our numerical experiments take into account transaction costs, *i.e.*, we buy assets at the ask price, which is the (midpoint) price plus one-half the bid-ask spread, and we sell assets at the bid price, which is the price minus one-half the bid-ask spread. Note that while we do not take into account transaction cost in our simple trading policy, we do in simulation and accounting. We use 3 basis points per day as a proxy for shorting costs. This corresponds to 7.5% annualized.

Metrics. The profit of a stat-arb is

$$\sum_{t=1}^{T} (V_t - V_{t-1}),$$

where $T = T^{\max} + T^{\text{exit}}$ is the number of trading days in the evaluation period. The return at time t is

$$r_t = \frac{V_t - V_{t-1}}{V_{t-1}}, \quad t = 1, \ldots, T.$$

We report several standard metrics based on the daily returns r_t. The average return is

$$\overline{r} = \frac{1}{T} \sum_{t=1}^{T} r_t,$$

which we multiply by 250 to annualize. The risk (return volatility) is

$$\left(\frac{1}{T} \sum_{t=1}^{T} (r_t - \overline{r})^2 \right)^{1/2},$$

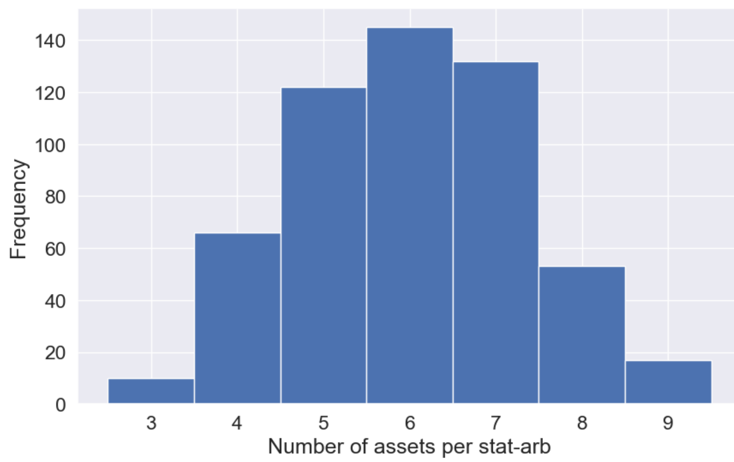

Figure 1: Distribution of the number of assets per fixed-band stat-arb.

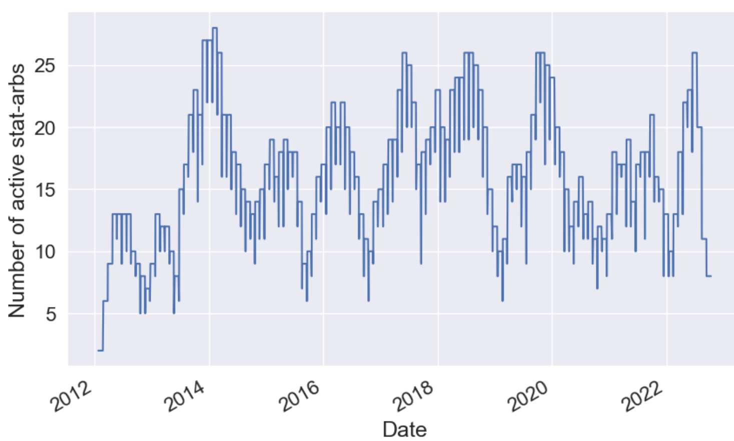

Figure 2: Number of active fixed-band stat-arbs over time.

which we multiply by $\sqrt{250}$ to annualize. The annualized Sharpe ratio is the ratio of the annualized average return to the annualized risk. Finally, the maximum drawdown is

$$\max_{1 \le t_1 < t_2 \le T} \left(\frac{V_{t_1}}{V_{t_2}} - 1 \right),$$

the maximum drop in value form a previous high.

4.3 Results for fixed-band stat-arbs

Stat-arb statistics. After solving the fixed-band stat-arb problem 1270 times, we found 545 unique stat-arbs. These stat-arbs contained between 3 and 9 assets, with a median value of 6, as shown in figure 1. Over time the number of active stat-arbs ranges up to 28, with a median value of 17, as shown in figure 2. Out of the 545 stat-arbs, 26 (around 5%) were

Profitability	
Fraction of profitable stat-arbs	63%
Annualized return	
Average	10%
Median	18%
75th percentile	33%
25th percentile	-2%
Annualized risk	
Average	32%
Median	21%
75th percentile	36%
25th percentile	12%
Annualized Sharpe ratio	
Average	0.81
Median	1.05
75th percentile	1.80
25th percentile	-0.06
Maximum drawdown	
Average	15%
Median	10%
75th percentile	19%
25th percentile	5%

Table 1: Metric summary for 545 fixed-band stat-arbs.

terminated before the end of the evaluation period, due to the NAV falling below 50% of the initial investment.

Metrics. Table 1 summarizes metrics related to the profitability of the fixed-band stat-arbs. Of the 545 stat-arbs, 63% were profitable. The average annualized return is 10%, with an average annualized risk of 32%, and an average annualized Sharpe ratio of 0.81. The maximum drawdown was on average 15% over the four-month trading period for each stat-arb.

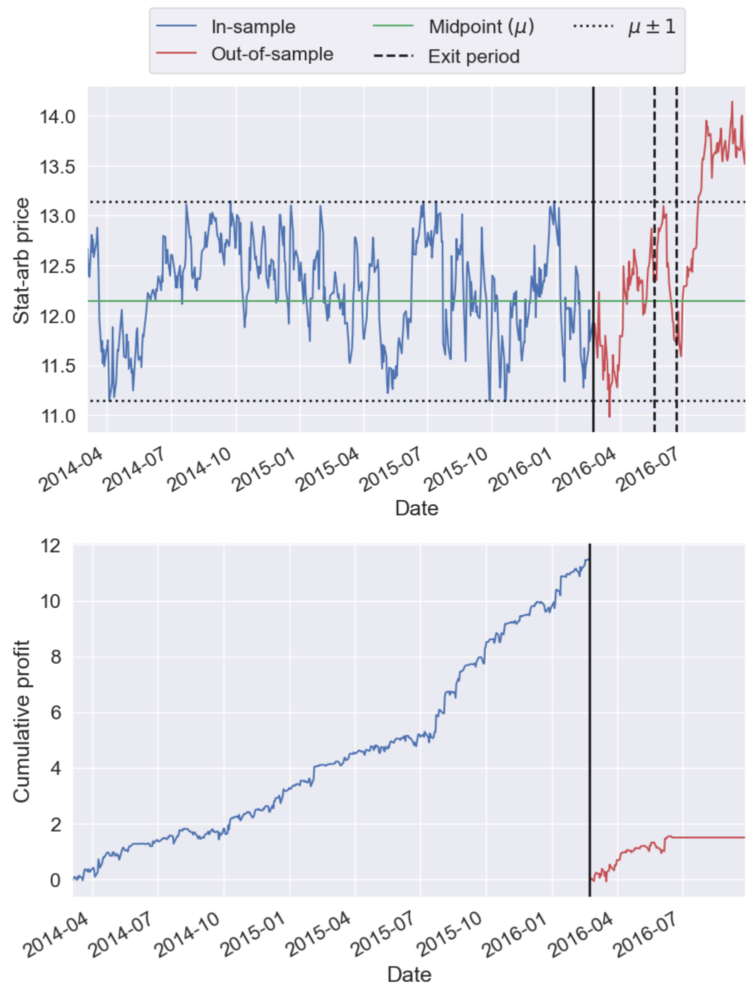

Figure 3: A fixed-band stat-arb strategy that made money. *Top.* Price. *Bottom.* Cumulative profit.

Example stat-arbs. We show the detailed evolution of two stat-arbs, one that made money and one that lost money, in figures 3 and 4, respectively. These were chosen to have average returns around the 75th and 25th percentiles of the return distribution across our 545 stat-arbs. As expected both stat-arbs are very profitable in-sample. The first one continues to be profitable out-of-sample.

The first one, which made money, contained the assets

Facebook
The Walt Disney Company
Eli Lilly and Company
Biogen
Occidental Petroleum Corporation
Alexion Pharmaceuticals.

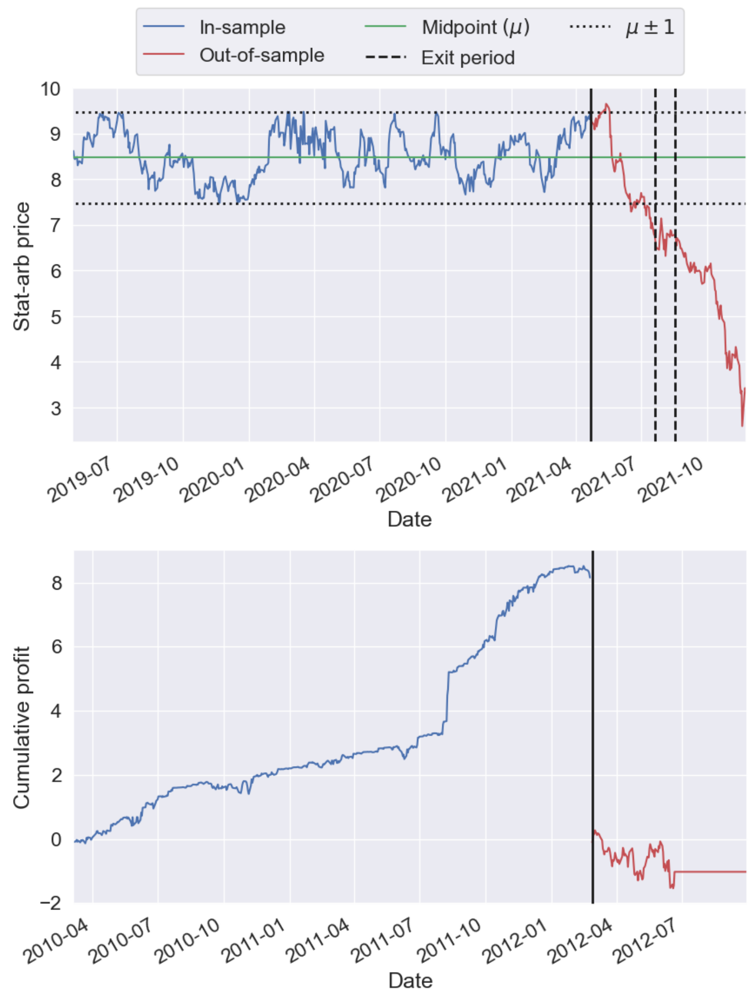

Figure 4: A fixed-band stat-arb strategy that lost money. *Top.* Price. *Bottom.* Cumulative profit.

The second one, which lost money, contained the assets

Berkshire Hathaway
Mastercard
ArcelorMittal
El Paso Corporation
Kinder Morgan Energy Partners.

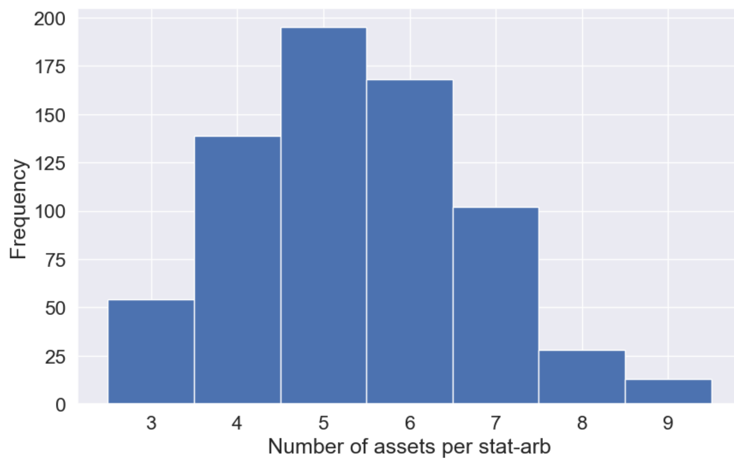

Figure 5: Distribution of the number of assets per moving-band stat-arb.

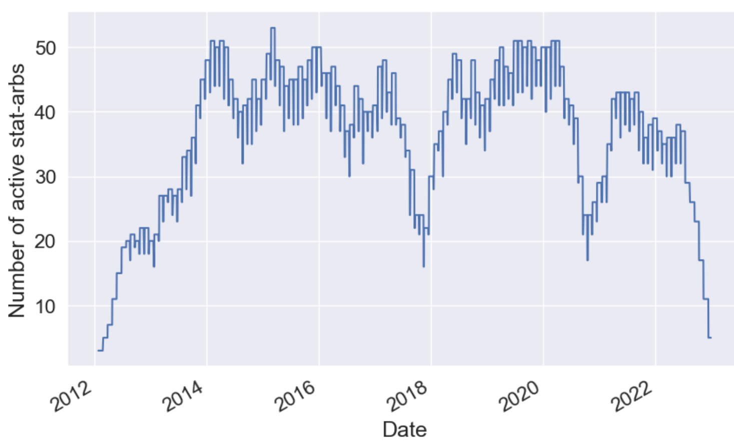

Figure 6: Number of active moving-band stat-arbs over time.

4.4 Results for moving-band stat-arbs

Stat-arb statistics. We found 712 unique moving-band stat-arbs (compared with 545 fixed-band stat-arbs). These stat-arbs contained between 1 and 10 assets, with a median value of 5. The full distribution is shown in figure 5. The number of active moving-band stat-arbs over time is shown in figure 6. The median number of active stat-arbs is 40. This is considerably larger than the number of active fixed-band stat-arbs since we find more of them, and they are active (by our choice) almost twice as long. Only three (around 0.4%) out of the 712 moving-band stat-arbs were terminated before the end of the evaluation period, due to the NAV falling below 50% of the initial investment.

Metrics. Table 2 summarizes the profitability of the moving-band stat-arbs. A large majority (70%) of the stat-arbs are profitable. The average annualized return was 15%, with an average annualized risk of 20%, and an average annualized Sharpe ratio of 0.84. The

Profitability	
Fraction of profitable stat-arbs	70%
Annualized return	
Average	15%
Median	12%
75th percentile	24%
25th percentile	3%
Annualized risk	
Average	20%
Median	15%
75th percentile	25%
25th percentile	9%
Annualized Sharpe	
Average	0.84
Median	0.88
75th percentile	1.52
25th percentile	0.21
Maximum drawdown	
Average	12%
Median	9%
75th percentile	15%
25th percentile	5%

Table 2: Metric summary for 712 moving-band stat-arbs.

maximum drawdown was on average 12% over the seven-month trading period for each stat-arb.

Comparison with fixed-band stat-arbs. Our first observation is that far fewer of the moving-band stat-arbs were terminated early due to low NAV than the fixed-band stat-arbs, despite their running for a period almost twice as long. Comparing tables 1 and 2 we see that the metrics for fixed-band stat-arbs are more variable, with a larger range in each of the metrics. The moving-band stat-arbs are more profitable than the fixed-band stat-arbs, but the difference is not large.

Example stat-arbs. Two stat-arbs, picked to represent roughly the 70th and 15th percentiles of the return distribution across the 712 stat-arbs, are illustrated in figures 7 and 8, respectively. Again, both stat-arbs are profitable in-sample, and the first one continues to be profitable out-of-sample.

The first one, which made money, contained the assets

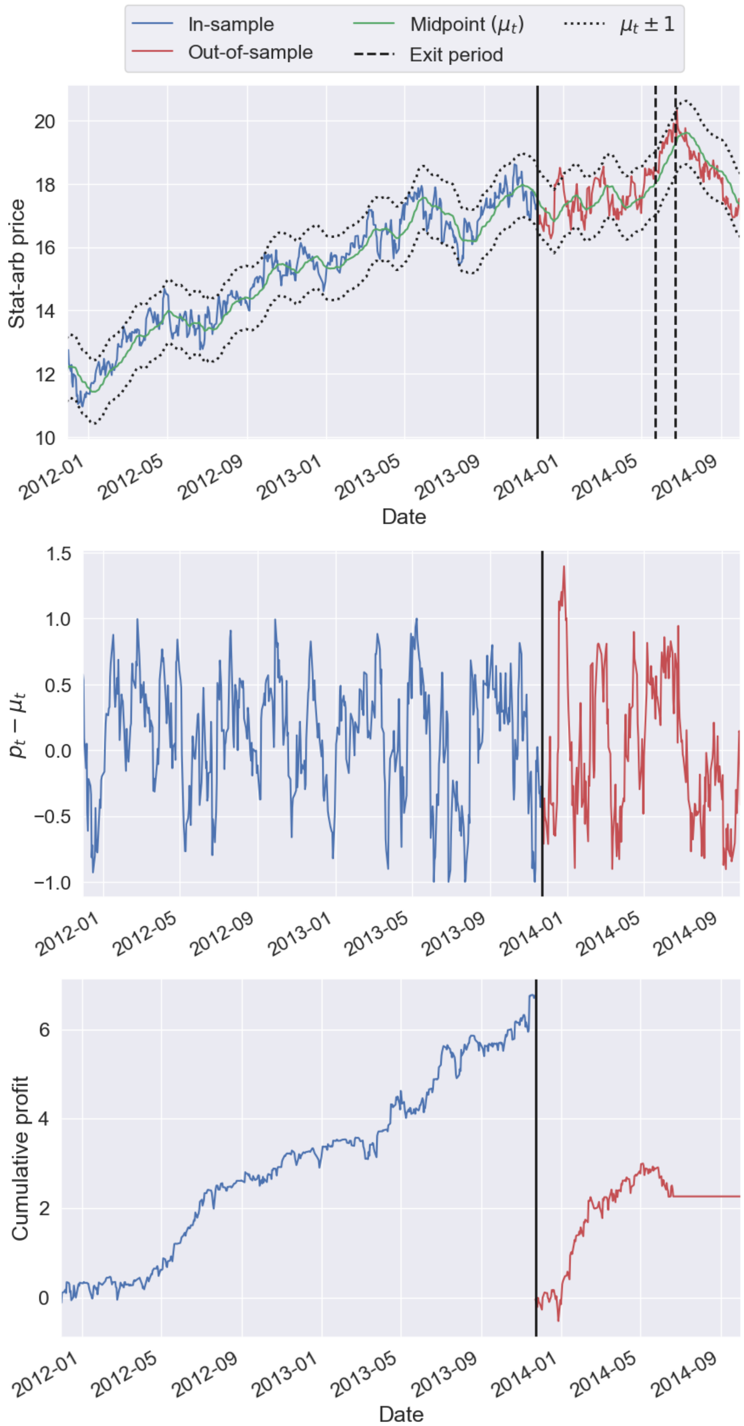

Figure 7: A moving-band stat-arb strategy that made money. *Top.* Price. *Middle.* Price relative to the trailing mean. *Bottom.* Cumulative profit.

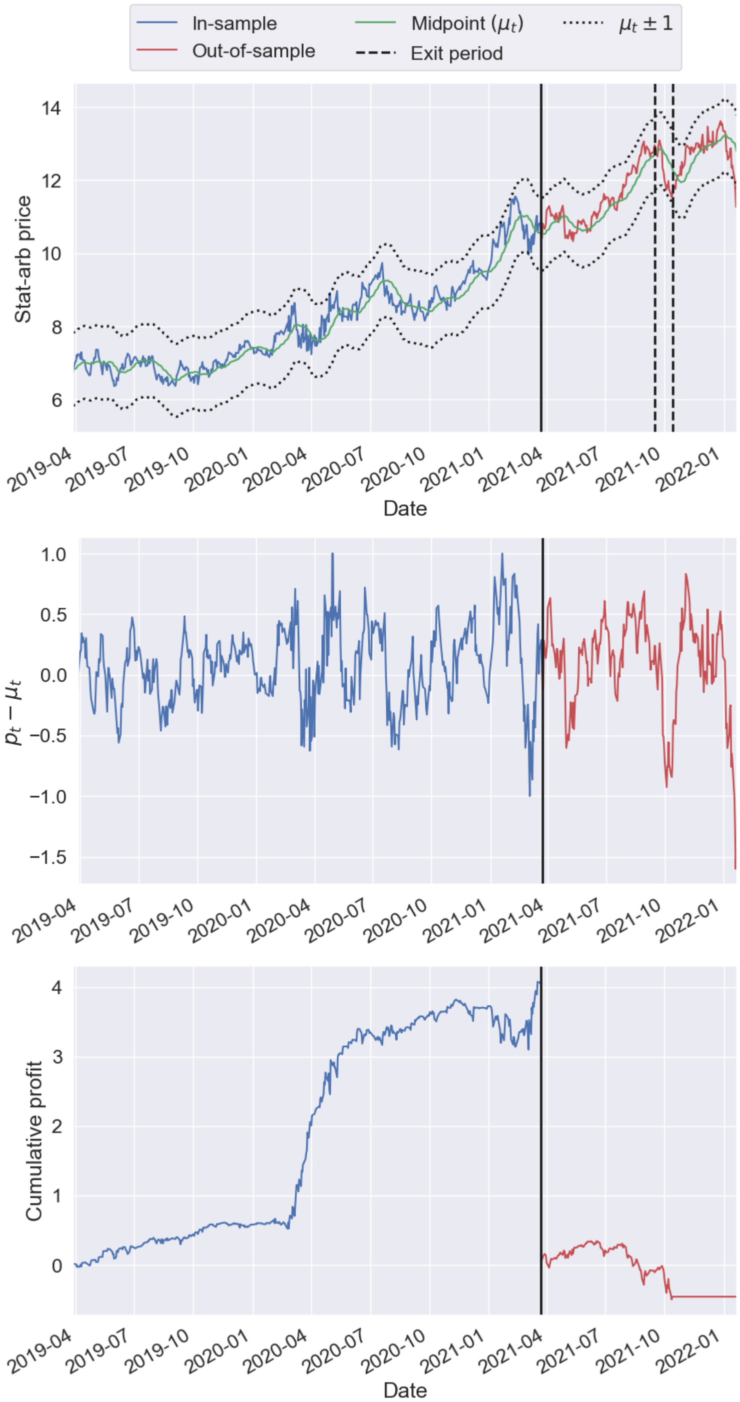

Figure 8: A moving-band stat-arb strategy that lost money. *Top.* Price. *Middle.* Price relative to the trailing mean. *Bottom.* Cumulative profit.

Morgan Stanley
Monsanto
Walgreen
Accenture
Pioneer Natural Resources.

The second one, which lost money, contained the assets

Lockheed Martin
ServiceNow
Gilead Sciences
NXP Semiconductors.

5 Conclusions and comments

We have formulated the problem of finding stat-arbs as a nonconvex optimization problem which can be approximately solved using the convex-concave procedure. We have introduced moving-band stat-arbs, which combine ideas from statistical arbitrage and price band trading.

Our empirical study on historical data shows that moving-band stat-arbs perform better than fixed-band stat-arbs, and remain profitable for longer out-of-sample periods. Our empirical study uses very simple trading and exit policies; we imagine that with more sophisticated ones such as those cited above, the results would be even better. Our focus in this paper is on finding stat-arbs, and not on trading them.

Variations and extensions. We mention here several ideas that we tried out, but were surprised to find did not improve the empirical results.

- *Asset screening.* We construct stat-arbs using assets only within an industry or sector.
- *Validation.* We split past asset prices into a training and a test set. We find candidate stat-arbs using the training data and then test them on the test data. We then only trade those with good test performance.
- *Incorporating transaction costs in the trading policy.* We modify the linear policy to take into account transaction costs. (Our simulations take trading cost into account, but our simple linear trading policy does not.)
- *Hysteresis-based trading.* We use a hysteresis-based trading policy, which can help reduce transaction costs compared to the linear policy.

Trading a portfolio of stat-arbs. We have focussed on finding individual stat-arbs. A next obvious topic is how to trade a portfolio of stat-arbs. This will be addressed in an upcoming paper by the authors. We simply note here that the results presented in this paper are all for single stat-arbs, and so fall somewhere in between individual assets and a full portfolio.

Acknowledgments

The idea of using the convex-concave procedure to find fixed-band stat-arbs is from initial unpublished work by Stephen Boyd and Jonathan Tuck. We thank Ron Kahn and Mark Mueller for detailed reviews of the paper and many helpful comments and suggestions. We gratefully acknowledge support from the Office of Naval Research. This work was partially supported by ACCESS – AI Chip Center for Emerging Smart Systems. Kasper Johansson was partially funded by the Sweden America Foundation.

References

[AD05] C. Alexander and A. Dimitriu. Indexing and statistical arbitrage. *The Journal of Portfolio Management*, 31(2):50–63, 2005.

[AL10] M. Avellaneda and J. Lee. Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7):761–782, 2010.

[Ber10] W. Bertram. Analytic solutions for optimal statistical arbitrage trading. *Physica A: Statistical mechanics and its applications*, 389(11):2234–2243, 2010.

[Bol92] J. Bollinger. Using bollinger bands. *Stocks & Commodities*, 10(2):47–51, 1992.

[Bol02] J. Bollinger. *Bollinger on Bollinger bands*. McGraw-Hill New York, 2002.

[BV04] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[CM13] J. Caldeira and G. Moura. Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Available at SSRN 2196391*, 2013.

[DGLR10] C. Dunis, G. Giorgioni, J. Laws, and J. Rudy. Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities. *Liverpool Business School*, 2010.

[DKB15] M. Dixon, D. Klabjan, and J. Bang. Implementing deep neural networks for financial market prediction on the Intel Xeon Phi. In *Proceedings of the 8th Workshop on High Performance Computational Finance*, New York, NY, USA, 2015. Association for Computing Machinery.

[DMPZ16] C. Eduardo De Moura, A. Pizzinga, and J. Zubelli. A pairs trading strategy based on linear state space models and the Kalman filter. *Quantitative Finance*, 16(10):1559–1573, 2016.

[EG87] R. Engle and C. Granger. Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987.

[EVDHM05] R. Elliott, J. Van Der Hoek, and W. Malcolm. Pairs trading. *Quantitative Finance*, 5(3):271–276, 2005.

[FKD19] T. Fischer, C. Krauss, and A. Deinert. Statistical arbitrage in cryptocurrency markets. *Journal of Risk and Financial Management*, 12(1):31, 2019.

[FP16] Y. Feng and D. Palomar. A signal processing perspective on financial engineering. *Foundations and Trends® in Signal Processing*, 9(1–2):1–231, 2016.

[GGR06] E. Gatev, W. Goetzmann, and G. Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3):797–827, 2006.

[GOPZ21] J. Guijarro-Ordonez, M. Pelger, and G. Zanotti. Deep learning statistical arbitrage. *arXiv preprint arXiv:2106.04028*, 2021.

[Gra83] C. Granger. *Co-integrated variables and error-correcting models*. PhD thesis, Discussion Paper 83-13. Department of Economics, University of California, 1983.

[HJTW04] S. Hogan, R. Jarrow, M. Teo, and M. Warachka. Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial economics*, 73(3):525–565, 2004.

[Huc09] N. Huck. Pairs selection and outranking: An application to the S&P 100 index. *European Journal of Operational Research*, 196(2):819–825, 2009.

[Huc10] N. Huck. Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, 207(3):1702–1716, 2010.

[Huc19] N. Huck. Large data sets and machine learning: Applications to statistical arbitrage. *European Journal of Operational Research*, 278(1):330–342, 2019.

[Joh00] S. Johansen. Modelling of co-integration in the vector autoregressive model. *Economic Modelling*, 17(3):359–373, 2000.

[JY07] J. Jurek and H. Yang. Dynamic portfolio selection in arbitrage. In *EFA 2006 Meetings Paper*, 2007.

[KDH17] C. Krauss, X. Do, and N. Huck. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2):689–702, 2017.

[Kra17] C. Krauss. Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, 31(2):513–545, 2017.

[KS17] C. Krauss and J. Stübinger. Non-linear dependence modelling with bivariate copulas: Statistical arbitrage pairs trading on the S&P 100. *Applied Economics*, 49(52):5352–5369, 2017.

[LB16] T. Lipp and S. Boyd. Variations and extension of the convex–concave procedure. *Optimization and Engineering*, 17:263–287, 2016.

[LLP21] J. Lee, J. Lee, and A. Prékopa. Price-bands: A technical tool for stock trading. *Fintech with Artificial Intelligence, Big Data, and Blockchain*, pages 221–246, 2021.

[MPW08] S. Mudchanatongsuk, J. Primbs, and W. Wong. Optimal pairs trading: A stochastic control approach. In *2008 American control conference*, pages 1035–1039. IEEE, 2008.

[MZ14] B. Moritz and T. Zimmermann. Deep conditional portfolio sorts: The relation between past and future stock returns. *LMU Munich and Harvard University Working paper*, 2014.

[Nak19] T. Nakajima. Expectations for statistical arbitrage in energy futures markets. *Journal of Risk and Financial Management*, 12(1):14, 2019.

[Per09] M. Perlin. Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives & Hedge Funds*, 15:122–136, 2009.

[Pol11] A. Pole. *Statistical arbitrage: algorithmic trading insights and techniques*. John Wiley & Sons, 2011.

[PY18] J. Primbs and Y. Yamada. Pairs trading under transaction costs using model predictive control. *Quantitative Finance*, 18(6):885–895, 2018.

[SDGB16] X. Shen, S. Diamond, Y. Gu, and S. Boyd. Disciplined convex-concave programming. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 1009–1014. IEEE, 2016.

[SH20] S. Sarmento and N. Horta. *A Machine Learning Based Pairs Trading Investment Strategy*. Springer, 2020.

[SMK18] J. Stübinger, B. Mangold, and C. Krauss. Statistical arbitrage with vine copulas. *Quantitative Finance*, 18(11):1831–1849, 2018.

[TL13] L. Takeuchi and Y. Lee. Applying deep learning to enhance momentum trading strategies in stocks. In *Technical Report*. Stanford University Stanford, CA, USA, 2013.

[Vid04] G. Vidyamurthy. *Pairs Trading: quantitative methods and analysis*, volume 217. John Wiley & Sons, 2004.

[VM17] M. Vaitonis and S. Masteika. Statistical arbitrage trading strategy in commodity futures market with the use of nanoseconds historical data. In *Information and Software Technologies*, pages 303–313, Cham, 2017. Springer International Publishing.

[WRD23] Wharton research data services. `https://wrds-web.wharton.upenn.edu/wrds/`, 2023.

[YP12] Y. Yamada and J. Primbs. Model predictive control for optimal portfolios with co-integrated pairs of stocks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 5705–5710. IEEE, 2012.

[YP18] Y. Yamada and J. Primbs. Model predictive control for optimal pairs trading portfolio with gross exposure and transaction cost constraints. *Asia-Pacific Financial Markets*, 25:1–21, 2018.

[ZP18] Z. Zhao and D. Palomar. Mean-reverting portfolio with budget constraint. *IEEE Transactions on Signal Processing*, 66(9):2342–2357, 2018.

[ZZP19] Z. Zhao, R. Zhou, and D. Palomar. Optimal mean-reverting portfolio with leverage constraint for statistical arbitrage in finance. *IEEE Transactions on Signal Processing*, 67(7):1681–1695, 2019.

[ZZWP18] Z. Zhao, R. Zhou, Z. Wang, and D. Palomar. Optimal portfolio design for statistical arbitrage in finance. In *2018 IEEE Statistical Signal Processing Workshop (SSP)*, pages 801–805. IEEE, 2018.