Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Portfolio Re-balancing and Optimization using Directional Changes and Genetic Algorithms

Rui Jorge Almeida, Nalan Bastürk, Paulo Rodrigues

DP 12/2023-064

Portfolio Re-balancing and Optimization using Directional Changes and Genetic Algorithms

Rui Jorge Almeida
Dept. of Quantitative Economics
Dept. of Data Analytics and Digitization
School of Business and Economics
Maastricht University
Maastricht, The Netherlands
rj.almeida@maastrichtuniversity.nl

Nalan Baştürk
Dept. of Quantitative Economics
School of Business and Economics
Maastricht University
Maastricht, The Netherlands
n.basturk@maastrichtuniversity.nl

Paulo Rodrigues
Dept. of Finance
School of Business and Economics
Maastricht University
Maastricht, The Netherlands
p.rodrigues@maastrichtuniversity.nl

2023 IEEE Congress on Evolutionary Computation (CEC) | 979-8-3503-1458-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/CEC53210.2023.10254022

***Abstract*—Dynamic portfolio optimization is a crucial but complex task due to financial market dynamics and the difficulty of disentangling noise from substantial changes in stock prices. In most existing methods, portfolios are re-optimized, hence re-balanced, at pre-specified time periods, return properties of each asset are dynamically computed, and portfolio weights are optimized according to an objective function. We propose a novel algorithm for dynamic portfolio optimization with a two-step signaling mechanism for re-balancing the portfolio including the optimization of re-balancing points and portfolio weights. The first step signals portfolio re-balancing only if there is a substantial price change in one or more of the portfolio constituents. These substantial price changes are defined according to directional change (DC) methods. DC methods create an intrinsic time series for each asset according to whether or not the change in the asset price exceeds a threshold level, hence removing part of the noise in asset prices. The second signaling mechanism uses genetic algorithms (GA) to assess if re-balancing is indeed profitable at each point indicated by the first signaling mechanism. The genetic algorithm is set up such that it simultaneously optimizes the weights of the re-balanced portfolio. For GA, we input the asset price summaries retrieved from DC methods to ensure that the GA can learn from the relatively less noisy data compared to observed asset prices. We show that the GA fit function can be set up to include several conventional trading strategies. As a first step, we apply the proposed method to a portfolio of 30 assets including 29 Exchange Traded Funds (ETF) and one risk-free asset where daily prices are observed during the period between 2 January 2018 and 30 December 2021. Second, we apply the method to 100 individual stocks for the same time period. We compare the obtained portfolio results with benchmarks, such as the simple buy and hold strategy of the S&P 500 index, the naive $1/N$ portfolio, and a minimum variance portfolio in terms of standard portfolio evaluation methods including the Sharpe ratio.**

***Index Terms*—Directional changes, financial portfolio optimization, genetic algorithms.**

I. Introduction

Portfolio optimization is one of the major tasks for individuals or institutions aiming to maintain good financial performance. There are several modeling techniques proposed in the literature for portfolio optimization, the most conventional ones being based on Markowitz mean-variance optimization [1]. Optimizing a portfolio of a typically large number of assets is a difficult process. Obtained portfolio results, such as realized returns, are often unstable due to reasons including numerical problems in estimating dynamic variance-covariance matrices, or more general dynamics of financial markets. We propose dynamic portfolio optimization using Directional Changes (DC) [2] as an alternative to conventional portfolio optimization based on returns obtained in fixed time intervals such as trading days. In addition, we base the dynamic portfolio optimization on novel genetic algorithms (GA) [3] for directional changes.

DC methods are proposed to summarize substantial price changes in an asset. A time series price curve is transformed into an intrinsic time curve that records upward or downward price changes that exceed a threshold level [2], [4], [5]. Price movements are then summarized as upward and downward trends where these trends usually continue for a period, defined by an overshoot (OS) event. In this setting, a change in the direction of an asset price gives a natural indication to buy or sell the asset in anticipation of future price trends. The use of a threshold to determine substantial price changes indicates that the noise in prices is partially removed, hence the obtained trading signals are more reliable compared to those obtained from observed returns [6]. Several studies analyzed this aspect and used confirmed price change points indicated by DCs to make investment decisions [6], [7]. The literature so far has focused on investments in a single asset. We extend this literature by considering a portfolio of assets.

Despite their advantages, DC methods even for a single asset lead to complex information, and DC signals are not necessarily accurate in forecasting price movements such as the OS event. GAs have been applied to portfolio optimization of (observed) time series [8]–[14] and extended to obtain buy and sell signals based on the intrinsic time series obtained from DCs [3], [6], [15], [16]. The use of GAs for DCs have been shown to be profitable for several assets. However, the literature on the use of GAs for DCs focuses on obtaining buy or sell signals on a single asset at a time, potentially due to DCs so far being applied to individual assets. We

N. Baştürk is partially supported by an NWO grant number 195.187.
P. Rodrigues is partially supported by a grant from the Network for Studies on Pensions, Aging, and Retirement (Netspar).

extend this literature to a more realistic case of portfolio optimization using DCs and GAs, where buy and sell decisions of a large number of assets have to be determined together with the weights allocated to each asset. To our knowledge, we provide the first study to consider the use of DCs for multiple assets in a portfolio optimization setting. Portfolio optimization is a multi-objective problem aiming to maximize returns and minimize risk. We simplify this multi-objective problem, derive it as an unconstrained optimization problem and solve it using genetic algorithms.

DC-based portfolio optimization has two main challenges, namely the synchronization of asset-specific DCs and the definition of a portfolio re-balancing method. The DC-based portfolio optimization requires the synchronization of the intrinsic time series since DC creates intrinsic time series for individual assets. There are potential return benefits from re-balancing a portfolio as soon as the expected return properties of any one of its component assets change. We, therefore, propose to obtain this joint intrinsic time series as the union of individual intrinsic time series. The latter challenge, portfolio re-balancing, occurs since DCs give information on the potential direction of the asset price, but not a clear indication of the amount of increase or decrease in prices. DC summaries can be used to obtain such information [17], but the performance of the methods has only been analyzed for individual assets. Furthermore, forecasts or portfolio decisions based solely on DCs can be sub-optimal since each DC is not necessarily followed by an OS event [16].

Our proposal is a two-step mechanism for portfolio re-balancing. In the first step, the proposed portfolio DC method gives an indication to re-balance the portfolio when at least one asset price has changed direction. In the second step, the GA signals whether the DC signal should be followed given the properties of the DC. Once a time point is chosen as a re-balancing point, the portfolio weights are obtained using Markowitz portfolio optimization without short-selling. We apply the proposed method to a portfolio consisting of prices of 30 assets - 29 Exchange Traded Funds (ETF) and one risk-free asset - and a set of 100 stocks included in the S&P index for the period between 2 January 2018 and 30 December 2021. We compare the obtained portfolio results with benchmarks, such as a simple buy and hold strategy of the S&P 500 index, the naïve $1/N$ portfolio, and minimum variance portfolios in terms of the standard portfolio evaluation methods including the Sharpe ratio. We show that the proposed method compares well with the baselines, especially in some periods. The performance of the proposed strategy, however, depends on the chosen DC threshold as well as the number of assets.

II. Directional Change Representation of Multiple Time Series

A. Directional Changes and Intrinsic Time Series Representation

Directional Change (DC) models are proposed to obtain an intrinsic time series that summarize market price movements of a financial instrument as upturn or downturn changes [2]. This intrinsic time series is identified by a change in the price given a pre-defined threshold value θ. If the absolute value of the price change exceeds the threshold at a given time period, this time period is labeled as a DC time point in the intrinsic time series. DC events can be either a downturn or an upturn event. A downturn DC event is defined as an event where the absolute price change between the current price p_t and the last high price p_h is lower than a fixed threshold θ :

$$p_t \leq p_h(1-\theta), \tag{1}$$

and an upturn DC event is defined as an event where the absolute price change between the current market price p_t and the last low price p_l is higher than a fixed threshold θ:

$$p_t \geq p_l(1+\theta). \tag{2}$$

A DC event is followed by an overshoot (OS) until an opposite, upward, or downward DC event occurs. A downward trend is defined as a downturn event followed by a downward overshoot, and an upward trend is defined as an upturn event followed by an upward overshoot. Thus the observed time series is converted to an intrinsic time series composed of DC and OS events, where a threshold θ determines this transformation [4], [6].

The summary of the obtained intrinsic time series together with the DC and OS events can be based on several metrics [2]. In this paper, we focus on the following metrics that will be used for portfolio construction. Let $t = 1, \ldots, T$ denote the time intervals during which prices are observed, $p_{n,t}$ denote the price level of asset $n = 1, \ldots, N$ at time $t = 1, \ldots, T$. Furthermore, let $r_{n,t} = 100 \times \ln(p_{n,t} - p_{n,t-1})$ denote the percentage returns of asset n at time t. Starting from an initial directional change point $\mathrm{DC}_{n,k-1}$, directional changes of each asset are defined with respect to the earlier directional change point as the reference point. For $k = 1, \ldots, K_n$ directional change points with $K_n < T$, each directional change and the sign of the directional change can be calculated iteratively. Specifically, if $DC_{n,k-1}$ is a downward directional change (DDC), the next directional change is an upward directional change (UDC) with the following timing:

$$\mathrm{DC}_{n,k} = \underset{t \geq \mathrm{DC}_{n,k-1}}{\arg\min} \left(p_{n,t} \geq p_{n,\mathrm{DC}_{n,k-1}}(1+\theta) \right), \tag{3}$$

$$p_{n,h} = p_{\mathrm{DC}_{n,k}}, \tag{4}$$

$$\mathrm{EDC}_{n,k} = \mathrm{UDC}, \tag{5}$$

where $p_{n,h}$ denotes the latest 'high price' in the market, and 'event' (EDC) indicates the sign of the directional change at every time period. If $\mathrm{DC}_{n,k-1}$ is an upward DC, the next directional change is a DDC with the following timing:

$$\mathrm{DC}_{n,k} = \underset{t \geq \mathrm{DC}_{n,k-1}}{\arg\min} \left(p_{n,t} \leq p_{n,\mathrm{DC}_{n,k-1}}(1-\theta) \right), \tag{6}$$

$$p_{n,l} = p_{n,\mathrm{DC}_{n,k}}, \tag{7}$$

$$\mathrm{EDC}_{n,k} = \mathrm{DDC}, \tag{8}$$

where $p_{n,l}$ denotes the latest 'low price' in the market and $\text{DC}_{n,0}$ can be initialized as an upward DC point at the beginning of the sample.

The relevant summary metrics reported for each asset n are the average time of a DC_n (TDC_n), the average time of an OS_n (TOS_n) after a confirmed directional change, and the average ratio of OS event length over the average ratio of DC event length ($\text{TOS}_n/\text{TDC}_n$), number of directional changes ($\#\text{DC}_n$, K_n), number of upward directional changes ($\#\text{UDC}_n$), number of downward directional changes ($\#\text{DDC}_n$) [17]:

$$\text{TDC}_n = \frac{1}{K_n}\sum_{k=1}^{K_n}\left(\text{DC}_{n,k} - \text{DC}_{n,k-1}\right), \tag{9}$$

$$\text{TOS}_n = \frac{1}{K_n}\sum_{k=1}^{K_n}\left(\text{OS}_{n,k} - \text{DC}_{n,k}\right), \tag{10}$$

$$\text{TOS}_n/\text{TDC}_n = \frac{1}{K_n}\sum_{k=1}^{K_n}\frac{\text{OS}_{n,k} - \text{DC}_{n,k}}{\text{DC}_{n,k} - \text{DC}_{n,k-1}}, \tag{11}$$

$$\#\text{DC}_n = K_n, \tag{12}$$

$$\#\text{UDC}_n = \frac{1}{K_n}\sum_{k=1}^{K_n} I[\text{EDC}_{k_n} = \text{UDC}_n], \tag{13}$$

$$\#\text{DDC}_n = \#\text{DC}_n - \#\text{UDC}_n, \tag{14}$$

where $I[.]$ is an indicator function which takes the value of 1 if its argument is true, and the value of 0 otherwise.

The total number of DC events for a given θ, $\#\text{DC}_n$, measures the variation of DC events for asset n. Based on the same θ, a time period with a lower #DC value will have less volatility than another time period of the same length. We note that the threshold parameter θ can be defined differently for each stock as an extension of the methodology.

B. Joint DC Representations of Multiple Assets

The defined DC approach in Section II-A creates an intrinsic time series for a single asset while a joint intrinsic time series for multiple assets requires a synchronization of the intrinsic time series. In the multivariate DC setting, a change in overshoot returns for one or more of the assets can indicate a substantial change in overall portfolio returns. We, therefore, define a portfolio DC point, PDC_k, as a point that at least one asset has a confirmed directional change:

$$\text{PDC}_k = \underset{t \geq \text{PDC}_{k-1}}{\arg\min}\ \exists n \text{ s.t. } \text{DC}_{n,k} = t\,, \tag{15}$$

where $k = 1, \ldots, K$ and PDC_0 is initialized at time 0, similar to the DC approach for individual assets. By this definition, the number of observations K of the joint intrinsic time series has the following properties:

$$K \geq \max_n K_n,\ \ K \leq \min\left(T, \sum_{n=1}^{N} K_n\right). \tag{16}$$

From the directional changes we obtain the OS events for each stock at each directional change point k:

$$\text{OS}_{n,k} = \begin{cases} \text{OS}_{n,k} & \text{if } \text{PDC}_k \in \text{DC}_n, \\ \text{OS}_{n,k-1} & \text{otherwise,} \end{cases} \tag{17}$$

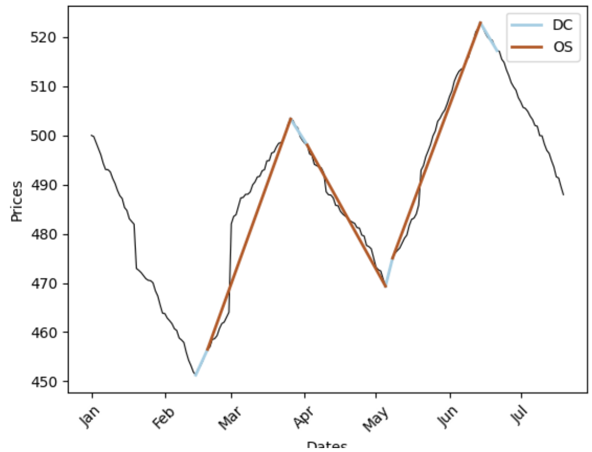

Fig. 1. DC representations of two artificially generated asset prices.

where $\text{DC}_n = \{\text{DC}_{n,1}, \ldots, \text{DC}_{n,K_n}\}$ is the set of DC points for asset n. I.e. the estimated overshoot event is only applied to assets that had an individual confirmed directional change at the time of the joint directional change. The remaining assets which do not have a directional change at time k are assumed to follow the direction they had at time $k-1$.

An example of DC representations of a single asset is given in Figure 1, where we illustrate four confirmed DCs and three OS events. Our proposal of portfolio rebalancing is based on such DC representations of each asset where a portfolio rebalancing signal is given at every point that one or more of the assets has a confirmed directional change in this figure.

III. Portfolio Optimization with Directional Changes and Genetic Algorithms

In this section, we introduce two portfolio optimization methods using DCs and GAs. We first present our proposal to join intrinsic time series of multiple assets. Second, we present the proposed GA for portfolio optimization.

A. Portfolio Optimization using joint DCs

The literature considers several measures to summarize the intrinsic time series created by DCs, as summarized in Section II. For portfolio optimization, we need additional measures that represent the relative profitability and return of each asset. The standard measures for these are the expected return and the variance-covariance matrix estimates. Since the DC method creates a new intrinsic time series, these measures cannot be applied directly to the DC time series. We propose to use the following metrics for the returns and volatilities from the DC-generated intrinsic time series:

$$\hat{r}_{n,\mathrm{UDC}} = \frac{\left(p_{\mathrm{OS}_{n,k}} - p_{\mathrm{OS}_{n,k-1}}\right) I(\mathrm{EDC}_{n,k} = \mathrm{UDC})}{\sum_{k=2}^{K} I(\mathrm{EDC}_{n,k} = \mathrm{UDC})}, \quad (18)$$

$$\hat{r}_{n,\mathrm{DDC}} = \frac{\left(p_{\mathrm{OS}_{n,k}} - p_{\mathrm{OS}_{n,k-1}}\right) I(\mathrm{EDC}_{n,k} = \mathrm{DDC})}{\sum_{k=2}^{K} I(\mathrm{EDC}_{n,k} = \mathrm{DDC})}, \quad (19)$$

$$\hat{s}_n = \frac{1}{K-1} \sum_{k=2}^{K} \left(p_{\mathrm{OS}_{n,k}} - p_{\mathrm{OS}_{n,k-1}}\right)^2, \quad (20)$$

where $p_{\mathrm{OS}_{n,k}}$ is defined as the price at the end of the OS period following the kth directional change for asset n. In other words, the expected return and risk properties are based on the historical OS events following each confirmed DC point.

Equations (18) and (19) represent the return properties of upward and downward DCs separately. These separate definitions allow a more flexible portfolio optimization than assuming the same average return under the two cases. Since the DC is designed to separate two directions, it is intuitive to also obtain return estimates for the two directions separately.

B. Genetic Algorithms for DC-based Portfolio Optimization

Genetic algorithms [18], [19] are robust iterative optimization methods that are based on a population of structures. GAs consider many structures as potential candidate solutions to the optimization problem, cover a large search space through sampling, and thus increase the probability of convergence to a global optimum compared to alternative and more local optimization methods. GAs have been applied to a wide range of problems including problems with uncertainty and problems that are not easily reduced to a precise mathematical formulation [20], and have been applied to different portfolio optimization problems [21]–[24].

The search within a defined GA is carried out through a set of potential solutions until the most superior ones come to dominate. GAs have been applied for obtaining trading signals and in general for portfolio optimization successfully [8]–[14]. More recently, GAs have been applied to DC models for obtaining trading signals [16], but to our knowledge, their application to directional-change implied portfolio optimization does not exist.

The design of a GA has three categories, namely the chromosome representation, fit function, and genetic operations of selection, crossover, and mutation. Chromosome representations can be binary representations that take the values of 0 or 1 to indicate whether the asset is included in the final portfolio [25]. The disadvantage of this setting is that the weights of each asset are not optimized directly with the defined GA. Alternatively, chromosomes are defined to represent the weight of each asset in the portfolio [26] and GAs have been extended to obtain solutions to this constrained optimization problem where the portfolio weights sum up to 1 [27], [28].

We define the real-valued chromosomes of the GA as vectors of size N where the portfolio weights are obtained as a transformation of variables from GA chromosomes:

$$\omega_n = \mathrm{softmax}(\delta_n),\ \delta_n \text{ for } n = 1, \ldots, N, \quad (21)$$

where δ_n is unconstrained.

The chromosome definition in (21) corresponds to portfolio optimization without short-selling since $\omega_n \in [0, 1]$ and the variable transformation mitigates the problem of constrained optimization since $\omega_n \in (0, 1)$ and $\sum_{n=1}^{N} \omega_n = 1$. This setting does not allow for exact values of 0 or 1 in portfolio optimization, but the weights can get arbitrarily close to 0, indicating that an asset is almost excluded from the portfolio.

The outlined GA algorithm constitutes the second step of the proposed two-step portfolio re-balancing and optimization method. In the first step, the proposed portfolio DC method gives an indication to re-balance the portfolio when at least one asset price has changed direction and the fitness function value is positive. In the second step, the GA signals the appropriate portfolio weights for each asset.

The fitness function of the GA is based on a return-risk metric we define for the whole portfolio:

$$PR(\omega_1, \ldots, \omega_N) = \frac{\sum_{n=1}^{N} \omega_n \hat{r}_n}{\sum_{n=1}^{N} \omega_n^2 \hat{s}_n}, \quad (22)$$

where $\hat{r}_n$ and $\hat{s}_n$ are intrinsic time series quasi-indicators of risk and return defined in (18)-(20). The intuition for this metric is akin to a Sharpe ratio where the metric increases with expected returns and decreases with volatility indicators of the assets, accounting for their weights in the portfolio. The GA then optimizes new weights according to the improvement in $PR(\omega_1, \ldots, \omega_N)$ compared to the last optimized weights.

We apply the proposed GA with rank selection, 80 generations, a population size of 40, 4 solutions selected as parents, single-point crossover with 80% probability, random mutation with 30% probability, and lower and upper bounds set as 0 and 5, respectively. Chromosome sizes are real-valued, with size equal to the number of assets in the application (30 or 100). GAs are performed with 20 runs for ETF returns and 10 runs for stock returns, and the best one is selected. The GAs were implemented using PyGAD version 2.18.1 [29].

IV. Application to Daily ETF Data

In this section, we apply the proposed method to daily ETF data and present the results of the obtained portfolio compared to the benchmarks. We use the mutual fund database of the Center for Research in Security Prices (CRSP). The return data is the daily total returns for the ETFs for the time period of January 2005 - December 2021. CRSP separates the ETFs into several style codes. For our analysis, we chose the largest equity ETFs in terms of average net asset values for the time period under consideration within each style category provided by CRSP. The final list of selected ETFs is available upon request. Finally, the risk-free rate is taken from the Fama-French factor database of Keneth French, accessed through Wharton Research Data Services.

A. Joint DC Representation of CRSP Data

We first summarize the DC representations of each asset and the proposed joint DC representation of all assets. We apply the DC algorithm with, $\theta = 0.04$ (GA 004) and $\theta = 0.05$ (GA

005). We choose two similar threshold values to study the sensitivity of the proposed methodology. Since we are using daily data, very small thresholds (e.g. $\theta = 0.01$) lead to almost one DC per time period (and no associated overshoot), while larger thresholds (e.g. $\theta = 0.1$) lead to large time-windows without any DC. We select a moving window size equal to the time between the first observation and the first time all intrinsic times series had at least one confirmed DC.

Table I provides summaries of DC representations of a selection of assets constituting the portfolio for the two DC threshold values. The number of observations in the intrinsic time series through DCs is much smaller than the number of observations in the full sample. This indicates that the first step of our method limits substantially the number of portfolio re-balancing points compared to, e.g. re-balancing the portfolio at each time period, and only takes place when there is a change of θ in at least one of the asset prices. In addition, the intrinsic time series across different stocks have different properties such as the average time to OS. We finally note that the joint intrinsic time series for the 30 assets has in total 1881 confirmed DCs for $\theta = 0.04$ and 1109 confirmed DCs for $\theta = 0.05$ and these numbers correspond to the number of times a portfolio would be re-balanced using the proposed method.

TABLE I
SUMMARY OF DC REPRESENTATIONS FOR A SELECTION OF ASSETS

	EFRP	EFRL	EFRE	EFRX	...	VIPERs Share Class
$\theta = 0.04$						
#DC	256	414	229	274	...	176
TOS	11.4	6.8	12.6	10.6	...	17.3
$\theta = 0.05$						
#DC	188	336	178	218	...	124
TOS	15.3	8.4	16.0	13.0	...	25.5
TOS/TDC	3.4	2.7	3.2	2.9	...	5.1

B. *Portfolio Investment Results*

The optimization of the GA for each DC converged without reaching the maximum number of generations. An analysis of the obtained weights shows that the optimization has a tendency to set non-zero weights to a small number of assets, usually including the one responsible for the confirmed DC that triggered portfolio re-balancing. For $\theta = 0.05$, 1023 out of 1109 portfolio weights are above 0.9 and for $\theta = 0.04$, 1666 out of 1881 optimized portfolio weights are above 0.9.

We compare our GA portfolio strategy to three other strategies. The first comparison is to a buy-and-hold portfolio that is invested in the S&P 500 index. The reason for this comparison is that this strategy is the default recommendation for most retail investors, so it is a natural benchmark. The second strategy consists of a naive 1/N portfolio, i.e., the weights are equal for all assets in the portfolio. It has been shown that this naive diversification strategy is hard to beat by a mean-variance optimizing strategy [30] and potential reasons for the good performance of the naive diversification strategy are discussed [31]. In this estimation strategy, we re-balance our portfolio to 1/N at the end of each month in our analysis. Finally, we compare the GA strategy to a Markowitz-type strategy [1]. However, instead of a mean-variance optimization, we use a minimum variance strategy in our comparison since expected return estimations for the mean-variance optimization using simple historical means are unreliable [32] and the common practice is a minimum variance approach [33]. As an estimator of the covariance matrix, we use the procedure of [34]. As re-balancing points, we use our GA re-balancing days.

TABLE II
PORTFOLIO RESULTS THROUGH ROLLING WINDOWS

	Exc. Ret.	Std. Dev.	Sharpe Ratio	$\hat{z}_{\mathrm{JK}}$
S&P 500	0.081	0.209	0.387	-15.970
1/N	0.113	0.147	0.768	2.760
Min. Var.	0.109	0.111	0.983	0.856
GA 004	0.027	0.217	0.124	-
GA 005	0.108	0.172	0.627	-

The investment strategies we consider take the PDC points defined as Section II as portfolio re-balancing points. For each one of these portfolio re-balancing points, the portfolio strategies require the calculation of quasi-indicators of risk and return, as defined in (18)-(20), based on historical data. Based on the portfolio optimization at each PDC point $k = 1, \dots, K-1$, the portfolio is held until the next PDC point $k+1$ for $k = 1, \dots, K-2$. The last portfolio is held until the end of the time period T. The returns of the portfolio holding periods are then stored as realized returns.

Table II presents the standard portfolio evaluation measures, the mean, variance and Sharpe Ratio of these realized returns. As we can see, three strategies, namely 1/n, Min. Var., and GA 005, which are based on an active re-balancing of the portfolio improve the performance compared to the buy and hold strategy in the S&P 500 index. However, the GA 004 strategy underperforms and achieves the lowest Sharpe ratio. The larger excess returns, i.e., the average return above the risk-free rate, for the 1/N, Min. Var., and GA 005 strategies are accompanied by a lower standard deviation. This results in larger Sharpe ratios for all three strategies. However, although the GA 005 strategy beats the passive holding of the S&P 500, it does not beat the 1/N or Min. Var. strategy. A test for statistical significance in performance when comparing the Sharpe Ratios of the benchmark strategies to the GA005 strategy can be found in the last column of Table II. To test if the Sharpe-Ratios of portfolio i and n are significantly different, [35], and the correction given in [36], show that

$$\hat{z}_{\mathrm{JK}} = \frac{\hat{\sigma}_n \hat{\mu}_i - \hat{\sigma}_i \hat{\mu}_n}{\sqrt{\hat{\vartheta}}}, \tag{23}$$

is asymptotically standard normal where

$$\hat{\vartheta} = \frac{1}{T}\left(2(\hat{\sigma}_i^2 \hat{\sigma}_n^2 - \hat{\sigma}_i \hat{\sigma}_n \hat{\sigma}_{i,n}) + \frac{\hat{\mu}_i^2 \hat{\sigma}_n^2 + \hat{\mu}_n^2 \hat{\sigma}_i^2}{2} - \frac{\hat{\mu}_i \hat{\mu}_n \hat{\sigma}_{i,n}^2}{\hat{\sigma}_i \hat{\sigma}_n}\right). \tag{24}$$

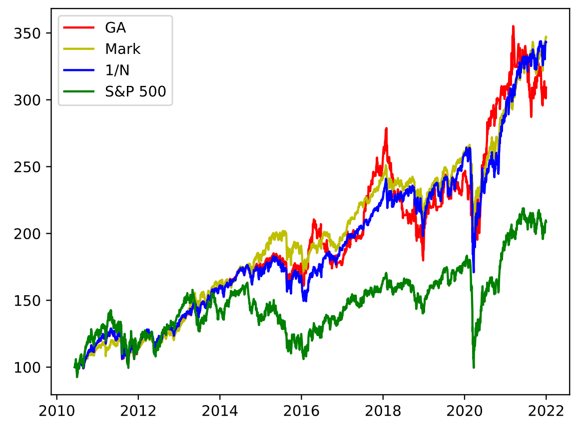

Fig. 2. Portfolio results through rolling windows with GA 005.

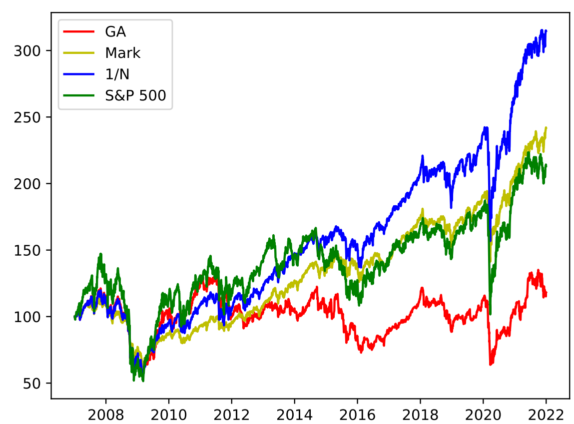

Fig. 3. Portfolio results through rolling windows with GA 004.

In Table II we present test statistics of the three benchmark strategies against the best-performing GA strategy, in this case, GA 005. We see that the GA 005 strategy significantly outperforms the S&P 500 strategy, underperforms the 1/N strategy, and is not significantly different from the minimum variance strategy.

Figure 2 presents cumulative realized returns for the GA 005 strategy and the benchmarks. For each strategy, we start an initial investment of 100 units and plot the evolution of the portfolio value over time. We see that the 1/N, the Min. Var., and GA 005 portfolio values start to diverge from the S&P 500 around 2015. The GA strategy shows periods of better performance than the 1/N, e.g., around 2016 and 2018. However, it is more prone to large corrections than the 1/N or the Min. Var strategy. This property results in a lower average return, but especially in a larger standard deviation, which explains the lower Sharpe ratio when compared to 1/N and Min. Var. We note that despite similar performances in some periods, our portfolio strategy leads to very sparse weights compared to 1/N: For $\theta = 0.05$, 1023 out of 1109 portfolio weights are above 0.9, and for $\theta = 0.04$, 1666 out of 1881 optimized portfolio weights are above 0.9. In the majority of cases, portfolios allocate a high weight to a single asset while a 1/N portfolio always diversifies across assets.

Figure 3 shows the same results and comparisons for the GA 004 strategy. Note that this lower threshold implies that there is a higher number of confirmed DC points in each asset, leading to more frequent portfolio rebalancing. The results of the GA 004 strategy are worse than those of GA 005 especially in cumulative. We see that the underperformance of the GA 004 portfolio starts around 2012 and generates a realized portfolio value that stays consistently below the benchmark values.

The main takeaway from this analysis is that it is possible to construct a portfolio strategy that, although underperforming 1/N, can rival it in terms of outcomes. Note that our algorithm is trained on daily data and as such could be missing important intraday information. Moreover, it could be the case that a portfolio containing a collection of ETFs does not provide the method with as many advantageous trading signals as a rather larger collection of single stocks. Compared to single stocks, ETF in itself is already a diversified investment vehicle that averages information across several assets. Finally, we note that the use of the same value for θ for all assets is a restriction that can impact the results since our results show large variations in performance for close values.

V. Application to Daily Stock Return Data

We next investigate the properties of the proposed methods in comparison to alternatives in a relatively large collection of single stocks. We use the stock return database of CRSP. The return data is the daily total returns for the time period of January 2005 - December 2021. Our stock selection proceeds as follows. We identify all stocks contained in the S&P 500 index for the totality of the time period of our analysis using the information provided by CRSP. After that, we compute the average market value of equity for the time period for each stock by multiplying the shares outstanding by the stock price. Finally, we select the 100 stocks with the largest average value. A list of all companies selected is available upon request.

TABLE III
Portfolio results through rolling windows

	Exc. Ret.	Std. Dev.	Sharpe Ratio	$\hat{z}_{\text{JK}}$
S&P 500	0.073	0.261	0.279	-1.535
1/N	0.138	0.209	0.659	-1.140
Min. Var.	0.128	0.157	0.817	-0.407
GA 003	0.164	0.328	0.499	-
GA 004	0.291	0.324	0.899	-
GA 005	0.247	0.304	0.815	-
GA 006	0.234	0.290	0.809	-
GA 008	0.155	0.255	0.607	-

A. Portfolio Investment Results

We follow the same procedure described in Section IV-B for the evaluation of the portfolio strategies when the asset

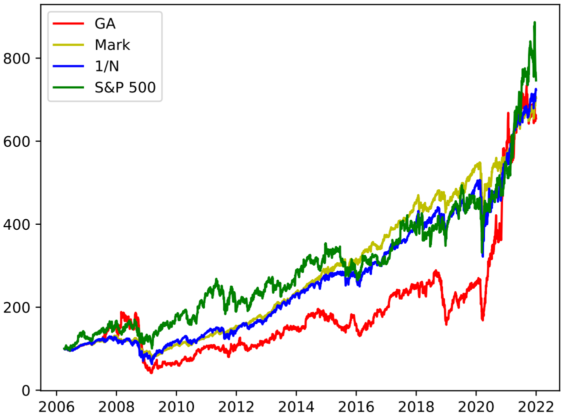

Fig. 4. Portfolio results through rolling windows with GA 003.

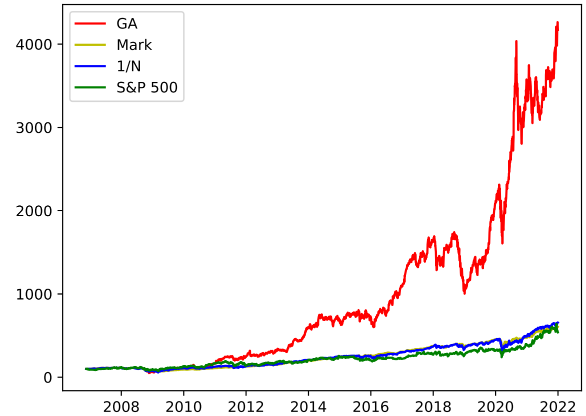

Fig. 5. Portfolio results through rolling windows with GA 004.

menu consists of 100 single stocks. The results of this exercise are given in Table III. The values of the test statistic are reported against the best performing GA model in terms of Sharpe Ratio, in this case, GA004. The Figures 4-6 depicts the cumulative returns of the GA strategies against the benchmark portfolios for varying values of the θ parameter. We show the results for GA003 and GA008 in Figures 4 and 6, respectively. Figure 5 shows the results for the best-performing model. The two other strategies show comparable returns as the one for GA004. Three things immediately stand out from the results: First, the best-performing GA strategy is no longer outperformed by any of the benchmarks. Second, the GA strategies are able to achieve an enormous return, however, at the cost of an equally large portfolio variance. Third, we see that the selection of the θ parameter has a crucial effect on the portfolio performance. A value that is too low, GA 003, has an equally deteriorating effect on the Sharpe Ratio as a value that is too high, GA 008.

This analysis shows that the performance of the portfolio strategy improves considerably when applied to an asset menu that consists of single stocks instead of ETFs. The reason for this is that the ETF is already a diversified portfolio, which eliminates the idiosyncratic risk component in the stock. By using stocks, the algorithm is able to identify DC events that are not averaged away in the diversified return series of the ETF. We note two points of caution. The data we analyze, stock return series from CRSP, are informative about the realized returns an investor can achieve, but this information does not take into account the market microstructure issues when actually trading on capital markets. Furthermore, trading in real markets incurs costs that are not accounted for in the analysis. Both of these issues present a direction for future research to refine the algorithm.

VI. Conclusion and Future Work

We propose a novel two-step algorithm for dynamic portfolio optimization using directional changes and genetic algorithms. The first step of the algorithm signals portfolio rebalancing only if there is a substantial price change in one or more of the portfolio constituents using DCs. The second step uses a GA to optimize portfolio weights at each point indicated by the first signaling mechanism. We apply the proposed method to a portfolio of 29 Exchange Traded Funds and the risk-free asset, and to a collection of 100 stocks contained in the S&P 500 index for the period between 2 January 2018 and 30 December 2021. We present the results of this application in comparison to several benchmarks and we report the effects of increasing the number of assets and the choice of the threshold parameter in the DC method. We find that the choice of the DC threshold parameter has a crucial effect on the portfolio performance. Furthermore, our method performs considerably better when applying it to an asset universe that is more diverse and has a higher number of assets.

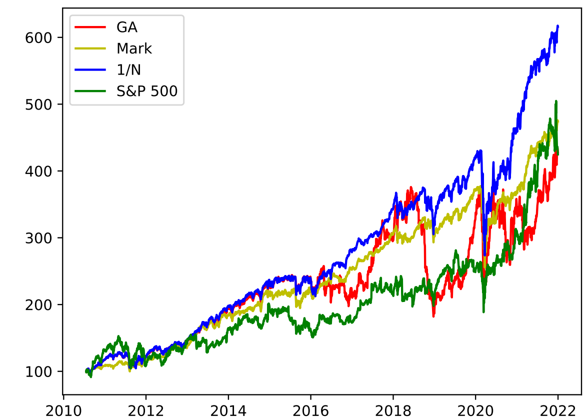

Fig. 6. Portfolio results through rolling windows with GA 008.

The proposed method can be extended in two directions. First, we consider a single threshold for every asset calculating

DCs. Recent methods to accommodate for multiple thresholds for a single asset [16] which can be extended to the multiple asset framework. Furthermore, the GA design can affect trading signals substantially [8]. We propose a relatively simple GA for portfolio optimization. Other GA designs, such as constrained optimization of parameters [28], can be used to directly estimate portfolio weights instead of the transformed parameters we propose. Finally, the algorithm can be refined to account for trading frictions.

References

[1] H. M. Markowitz, *Portfolio selection: efficient diversification of investments*. Yale University Press, 1968, vol. 16.

[2] E. P. Tsang, R. Tao, A. Serguieva, and S. Ma, "Profiling high–frequency equity price movements in directional changes," *Quantitative Finance*, vol. 17, no. 2, pp. 217–225, 2017.

[3] J. Gypteau, F. E. Otero, and M. Kampouridis, "Generating directional change based trading strategies with genetic programming," in *European Conference on the Applications of Evolutionary Computation*. Springer, 2015, pp. 267–278.

[4] J. Glattfelder, A. Dupuis, and R. Olsen, "Patterns in high-frequency FX data: discovery of 12 empirical scaling laws," *Quantitative Finance*, vol. 11, no. 4, pp. 599–614, 2011.

[5] R. J. Almeida, N. Baştürk, and R. Golan, "Intraday value-at-risk estimation for directional change events and investment strategies," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–8.

[6] M. Kampouridis and F. E. Otero, "Evolving trading strategies using directional changes," *Expert Systems with Applications*, vol. 73, pp. 145–160, 2017.

[7] A. Bakhach, E. P. K. Tsang, and H. Jalalian, "Forecasting directional changes in the FX markets," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2016, pp. 1–8.

[8] R. J. Bauer and G. E. Liepins, "Genetic algorithms and computerized trading strategies," *Expert Systems in Finance*, vol. 5, pp. 89–100, 1992.

[9] V. Karunamurthy, "A genetic programming approach to the dynamic portfolio rebalancing problem," *Genetic Algorithms and Genetic Programming at Stanford*, vol. 2003, pp. 100–108, 2003.

[10] A. Hirabayashi, C. Aranha, and H. Iba, "Optimization of the trading rule in foreign exchange using genetic algorithm," in *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, 2009, pp. 1529–1536.

[11] E. Alfaro-Cid, J. S. Baixauli-Soler, and M. O. Fernandez-Blanco, "Minimising value-at-risk in a portfolio optimisation problem using a multiobjective genetic algorithm," *International Journal of Risk Assessment and Management*, vol. 15, no. 5-6, pp. 453–477, 2011.

[12] J. Samuel Baixauli-Soler, E. Alfaro-Cid, and M. O. Fernandez-Blanco, "A naïve approach to speed up portfolio optimization problem using a multiobjective genetic algorithm," *Investigaciones Europeas de Dirección y Economía de la Empresa*, vol. 18, no. 2, pp. 126–131, 2012.

[13] S.-H. Chen, *Genetic algorithms and genetic programming in computational finance*. Springer Science & Business Media, 2012.

[14] R. Aguilar-Rivera, M. Valenzuela-Rendón, and J. Rodríguez-Ortiz, "Genetic algorithms and Darwinian approaches in financial applications: A survey," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7684–7697, 2015.

[15] A. Adegboye, M. Kampouridis, and C. G. Johnson, "Regression genetic programming for estimating trend end in foreign exchange market," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–8.

[16] A. Adegboye, M. Kampouridis, and F. Otero, "Algorithmic trading with directional changes," *Artificial Intelligence Review*, pp. 1–26, 2022.

[17] R. J. Almeida and N. Baştürk, "Forecasting directional change uncertainty using probabilistic fuzzy systems," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2018, pp. 1–8.

[18] J. H. Holland, "Adaptive algorithms for discovering and using general patterns in growing knowledge bases," *International Journal of Policy Analysis and Information Systems*, vol. 4, no. 3, pp. 245–268, 1980.

[19] ——, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[20] M. A. H. Dempster and C. M. Jones, "A real-time adaptive trading system using genetic programming," *Quantitative Finance*, vol. 1, no. 4, p. 397, 2001.

[21] K. J. Oh, T. Y. Kim, and S. Min, "Using genetic algorithm to support portfolio optimization for index fund management," *Expert Systems with applications*, vol. 28, no. 2, pp. 371–379, 2005.

[22] T.-J. Chang, S.-C. Yang, and K.-J. Chang, "Portfolio optimization problems in different risk measures using genetic algorithm," *Expert Systems with applications*, vol. 36, no. 7, pp. 10 529–10 537, 2009.

[23] C.-M. Lin and M. Gen, "An effective decision-based genetic algorithm approach to multiobjective portfolio optimization problem," *Applied mathematical sciences*, vol. 1, no. 5, pp. 201–210, 2007.

[24] K. Metaxiotis and K. Liagkouras, "Multiobjective evolutionary algorithms for portfolio management: A comprehensive literature review," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11 685–11 698, 2012.

[25] R. Hochreiter, "An evolutionary computation approach to scenario-based risk-return portfolio optimization for general risk measures," in *Applications of Evolutionary Computing: EvoWorkshops 2007: EvoCoMnet, EvoFIN, EvoIASP, EvoINTERACTION, EvoMUSART, EvoSTOC and EvoTransLog. Proceedings*. Springer, 2007, pp. 199–207.

[26] C. Aranha and H. Iba, "Modelling cost into a genetic algorithm-based portfolio optimization system by seeding and objective sharing," in *2007 IEEE Congress on Evolutionary Computation*. IEEE, 2007, pp. 196–203.

[27] M. Kaucic, M. Moradi, and M. Mirzazadeh, "Portfolio optimization by improved NSGA-II and SPEA 2 based on different risk measures," *Financial Innovation*, vol. 5, pp. 1–28, 2019.

[28] K. Tagawa and Y. Orito, "A mapping-based constraint-handling technique for evolutionary algorithms with its applications to portfolio optimization problems," *SICE Journal of Control, Measurement, and System Integration*, vol. 15, no. 1, pp. 62–76, 2022.

[29] A. F. Gad, "Pygad: An intuitive genetic algorithm python library," 2021.

[30] V. DeMiguel, L. Garlappi, and R. Uppal, "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *The Review of Financial Studies*, vol. 22, no. 5, pp. 1915–1953, 12 2007.

[31] Y. Plyakha, R. Uppal, and G. Vilkov, "Why does an equal-weighted portfolio outperform value- and price-weighted portfolios?" SSRN id2724535, Working Paper, 2012.

[32] R. C. Merton, "On estimating the expected return on the market: An exploratory investigation," *Journal of Financial Economics*, vol. 8, no. 4, pp. 323–361, 1980.

[33] V. DeMiguel, Y. Plyakha, R. Uppal, and G. Vilkov, "Improving portfolio selection using option-implied volatility and skewness," *Journal of Financial and Quantitative Analysis*, vol. 48, no. 6, p. 1813–1845, 2013.

[34] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[35] J. D. Jobson and B. M. Korkie, "Performance hypothesis testing with the sharpe and treynor measures," *The Journal of Finance*, vol. 36, no. 4, pp. 889–908, 1981.

[36] C. Memmel, "Performance hypothesis testing with the sharpe ratio," *Finance Letters*, no. 1, pp. 21–23, 2003.