Noname manuscript No.
(will be inserted by the editor)

Generalized risk parity portfolio optimization: An ADMM approach

Giorgio Costa · Roy H. Kwon

Received: date / Accepted: date

Abstract The risk parity solution to the asset allocation problem yields portfolios where the risk contribution from each asset is made equal. We consider a generalized approach to this problem. First, we set an objective that seeks to maximize the portfolio expected return while minimizing portfolio risk. Second, we relax the risk parity condition and instead bound the risk dispersion of the constituents within a predefined limit. This allows an investor to prescribe a desired risk dispersion range, yielding a portfolio with an optimal risk–return profile that is still well-diversified from a risk-based standpoint. We add robustness to our framework by introducing an ellipsoidal uncertainty structure around our estimated asset expected returns to mitigate estimation error. Our proposed framework does not impose any restrictions on short selling. A limitation of risk parity is that allowing of short sales leads to a non-convex problem. However, we propose an approach that relaxes our generalized risk parity model into a convex semidefinite program. We proceed to tighten this relaxation sequentially through the alternating direction method of multipliers. This procedure iterates between the convex optimization problem and the non-convex problem with a rank constraint. In addition, we can exploit this structure to solve the non-convex problem analytically and efficiently during every iteration. Numerical results show that this algorithm converges to a higher quality optimal solution when compared to the competing non-convex problem.

Keywords Non-convex optimization · Robust optimization · ADMM · Risk parity · Asset allocation

1 Introduction

Since the introduction of modern portfolio theory [24], the application of optimization theory in asset allocation and portfolio construction has been widely studied.

G. Costa · R. H. Kwon (✉)
Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, M5S 3G8, Canada
E-mail: {gcosta, rkwon}@mie.utoronto.ca

Moreover, its application by industry practitioners has been consistently growing ever since. In particular, the mean–variance optimization (MVO) mathematical framework serves to deliver an optimal compromise between financial risk and reward. Although MVO emphasizes a quantitative approach to asset management, it suffers from its susceptibility to estimation error. Having noisy input parameters may lead to what has been referred to as 'error maximization', yielding poor out-of-sample portfolio performance. Indeed, the sensitivity of portfolio optimization to estimation error has been studied in detail [2, 4, 25]. The two key input parameters in MVO are the asset variances and covariances to measure risk and the asset expected returns to measure reward. Errors in the estimated returns have an impact an order of magnitude larger than errors in the asset variances and covariances during optimization [6].

Traditionally, this shortfall is addressed by either of two methods. We can improve the quality of our estimated parameters by using factor models [10,21,26,31] or through Bayesian shrinkage [18, 20]. Alternatively, we can apply robust optimization techniques. A typical implementation of robust portfolio optimization mitigates the effect of uncertainty through the use of either box constraints or ellipsoidal constraints [22, 32]. A different approach incorporates the uncertainty of the choice of probability distribution governing the asset returns, formulating a distributionally robust optimization problem [7]. More generally, robust optimization can be used in tandem with factor models, where the uncertainty is derived from a factor model structure as described in [15].

A modern asset allocation framework that has recently gained traction in both academia and industry is risk parity [1, 23, 28]. The objective is to allocate wealth in such a way as to equalize the risk contribution per asset, meaning any resulting portfolio is fully diversified from a risk perspective. This framework excludes the use of expected returns, thereby negating the drawbacks that arise from their estimation error. Risk budgeting, a more general form of risk parity, yields portfolios where assets or collections of assets are given a predetermined risk weight as a percentage of the total portfolio risk, or 'risk budget' [5]. A robust risk budgeting framework is presented in [19]. Although rejecting the use of expected returns has the benefit of stability, it is counter-intuitive for most investors, who consider the use of portfolio return a necessity. Moreover, risk parity optimization has the fundamental limitation of only being convex for 'long-only' portfolios, i.e., the problem becomes non-convex if short selling is allowed[1].

To address these limitations, we introduce a generalized risk parity framework. This framework emphasizes a portfolio risk–return profile as the investor's objective, while also allowing short selling. Moreover, we relax the risk parity condition, allowing us to control the degree of risk-based diversification between the portfolio constituents. In essence, we outline a framework where an investor with a given risk–return profile can impose a predefined limit on the portfolio risk dispersion, maintaining a desired level of risk-based diversification while giving the optimization problem flexibility to attain a more desirable objective value. We extend this framework by considering a robust formulation of the asset expected returns in an effort to mitigate the impact of estimation error. However, as we will

[1] Short selling pertains to taking negative positions on our investments. 'Long-only' refers to portfolios where we can only take non-negative positions.

discuss later on, allowing for 'long–short' portfolios while constraining individual risk contributions means this is a non-convex optimization problem.

Managing the non-convexity of both risk parity and risk budgeting is the subject of ongoing research. Feng and Palomar [11] propose an algorithm that sequentially solves a series of convex problems by using a first-order approximation to the original non-convex risk parity problem. Their research is centred around the nominal risk parity problem, and their first-order formulation does not allow for the type of non-convex quadratic constraint we will need to bound the asset risk dispersion. Moreover, their formulation is able to guarantee global convergence to a stationary point, but cannot make any claims about the quality of this local optimum relative to all extant optimal solutions. Another more general algorithm is proposed by Haugh et al. [17] for a generalized risk budgeting framework that seeks to maximize the risk–return utility of an investor while having binding risk contribution constraints. However, an investor must predetermine the exact risk budgets assigned to each asset (or basket of assets) before optimization takes place. To overcome the issue of non-convexity, Haugh et al. [17] introduce an algorithm that combines the augmented Lagrangian and Markov chain Monte Carlo methods. This heuristic approach relies on sampling different starting points from the feasible set, and then proceeding to solve the non-convex problem repeatedly using traditional non-linear optimization algorithms. The best solution is then chosen as the quasi-global optimum.

We propose a different approach, which relies on sequentially tightening a semidefinite program (SDP) relaxation of the original non-convex problem. By definition, SDPs are convex. However, the solution to a SDP relaxation can only provide a lower bound to the global optimum. In the case of risk parity, the solution of the SDP relaxation generally violates any constraints pertaining to the asset risk contributions. As we will see in Section 4, a feasible solution can be attained if we impose a non-convex rank-1 constraint on the SDP. Imposing a non-convex constraint on our SDP might seem counterproductive, but proceeding in this fashion allows us to decompose this into two sub-problems: a convex SDP, and a non-convex rank-1 approximation. The first step involves solving a modified version of our SDP relaxation, while the second step requires us to find the closest rank-1 solution to the first step. Although non-convex, the second step has a closed-form expression that allows us to efficiently find an exact solution. We proceed to implement the alternating direction method of multipliers (ADMM) algorithm [3, 12, 14]. We favour this approach because, through the tightening of our relaxation, we approximate the global optimum from its lower bound. Relaxing a non-convex into a convex SDP and using ADMM to sequentially impose a rank-1 constraint is not a new approach. Indeed, this has been successfully applied in optimal power flow problems [33]. This paper proposes a similar method based on an ADMM algorithm to solve the non-convex generalized risk parity problem. This algorithm sequentially tightens the relaxed SDP, converging to a rank-1 constrained SDP solution, which is equivalent to our non-convex formulation. The sequential tightening of the relaxed feasible set ensures we approach our optimal solution from a lower bound, converging to a higher quality optimal solution.

The outline of this paper is the following. Section 2 discusses the nominal risk parity framework in detail, emphasizing the non-convexity that arises when short selling is allowed. Section 3 introduces the generalized risk parity framework. The corresponding SDP relaxation is presented in Section 4, followed by the introduction of the ADMM algorithm. The numerical experiments in Section 5 corroborate this by comparing the performance of our proposed algorithm against the non-convex formulation, even after we warm-start the non-convex problem. Finally, Section 6 summarizes the findings and contribution of this paper.

2 Risk parity

This section presents the original risk parity problem and examines the issue of non-convexity that arises when we relax the 'long-only' condition. First, we discuss the input parameters required to measure reward and risk. For a portfolio with n assets, the typical choices are the asset expected returns, $\boldsymbol{\mu} \in \mathbb{R}^n$, and the asset covariance matrix, $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$. These measure reward and risk, respectively. Thus, our portfolio expected return, $\mu_p \in \mathbb{R}$ and variance, $\sigma_p^2 \in \mathbb{R}$, are

$$\begin{aligned} \mu_p &= \boldsymbol{\mu}^T \boldsymbol{x}, \\ \sigma_p^2 &= \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x}, \end{aligned}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is our decision variable and measures proportion of wealth invested in each asset, i.e., it serves as the vector of asset weights.

The risk parity framework solves the asset allocation problem by assigning wealth in such a way that the individual asset risk contributions are equalized. To achieve this, we must first determine the risk contribution per asset. As shown in [23], these can be derived through an Euler decomposition of a portfolio risk measure. The only condition is that the risk measure is a homogeneous function of degree one. The portfolio standard deviation, which is the square root of the portfolio variance, complies with this condition. As such, we can decompose it as follows

$$\sigma_p = \sqrt{\boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x}} = \sum_{i=1}^{n} x_i \frac{\partial \sigma_p}{\partial x_i} = \sum_{i=1}^{n} x_i \frac{(\boldsymbol{\Sigma} \boldsymbol{x})_i}{\sqrt{\boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x}}}, \tag{1}$$

where $\partial \sigma_p / \partial x_i$ is the marginal risk contribution of asset i. We note the denominator within the sum of the partitioned standard deviation is consistent for all parts, and this denominator is equal to σ_p. If we consider portfolio variance as our risk measure, Equation (1) allows us to derive the following useful property

$$\sigma_p^2 = \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} = \sum_{i=1}^{n} x_i (\boldsymbol{\Sigma} \boldsymbol{x})_i = \sum_{i=1}^{n} R_i,$$

where $R_i = x_i (\boldsymbol{\Sigma} \boldsymbol{x})_i$ is the individual risk contribution of asset i to the overall portfolio variance. The risk parity framework seeks to compute a portfolio where $R_i = R_j \ \forall \ i, j$. This can be achieved through a least squares approach

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} \left(x_i (\boldsymbol{\Sigma} \boldsymbol{x})_i - x_j (\boldsymbol{\Sigma} \boldsymbol{x})_j \right)^2 \\ \text{s.t.} \quad & \mathbf{1}^T \boldsymbol{x} = 1, \\ & \boldsymbol{x} \geq 0, \end{aligned} \tag{2}$$

where $\mathbf{1} \in \mathbb{R}^n$ denotes a vector where all elements are equal to one. The objective of Model (2) is to minimize the difference between the asset risk contributions. In addition, the first constraint ensures all available wealth is invested while the second constraint prohibits short selling. Both of these constraints are affine. Moreover, the asset covariance matrix, $\boldsymbol{\Sigma}$, is positive semidefinite[2] (PSD). However, partitioning the portfolio variance into the individual asset risk contributions leads to a non-convex expression. The issue of non-convexity is prevalent in any risk parity formulation where the individual risk contributions are employed.

2.1 Non-convexity of risk parity

The non-convexity of Model (2) becomes apparent if we inspect the individual risk contributions, R_i. We begin by recasting R_i in standard quadratic notation

$$R_i = x_i(\boldsymbol{\Sigma x})_i = \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x},$$

where $\boldsymbol{A}_i \in \mathbb{R}^{n\times n}$ captures the individual risk contribution of asset i. The symmetric matrices $\boldsymbol{A}_i$ are composed of the superposition of row i and column i from the original covariance matrix $\boldsymbol{\Sigma}$ multiplied by one half, with all other elements in the matrix equal to zero, i.e.,

$$\boldsymbol{A}_i = \frac{1}{2}(\boldsymbol{e}_i \boldsymbol{e}_i^T \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \boldsymbol{e}_i \boldsymbol{e}_i^T),$$

where $\boldsymbol{e}_i \in \mathbb{R}^n$ denotes the i^{th} column of the identity matrix. Inspecting the sparse matrices $\boldsymbol{A}_i$ reveals these are indefinite, each having a single positive eigenvalue, a single negative eigenvalue, and all other eigenvalues being equal to zero. Thus, any optimization problem that employs the risk contributions R_i will be non-convex.

Nevertheless, Model (2) is numerically efficient and is able to consistently attain a unique global solution within the feasible set of long-only portfolios [23]. Imposing a long-only constraint is not financially motivated, but rather it is a necessity arising from the non-convexity of the optimization model itself. A unique global solution, where the asset risk contributions are equalized, is only guaranteed if we shrink the feasible set to only have long positions.

To elaborate on this subject, we refer to the convex formulation of risk parity proposed by Bai et al. [1]

$$\begin{aligned} \min_{x} \quad & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q x} - b\sum_{i=1}^{n} \ln \beta_i x_i \\ \text{s.t.} \quad & \beta_i x_i > 0, \quad \text{for } i = 1, ..., n, \end{aligned} \tag{3}$$

where $\beta_i \in \{-1, 1\}\ \forall\ i$ and $b \geq 0$ is an arbitrary positive constant. At optimality, we find that

$$(\boldsymbol{Q x})_i = \frac{b}{x_i} \ \ \forall\ i \quad \Rightarrow \quad x_i(\boldsymbol{Q x})_i = b \ \ \forall\ i,$$

retrieving a solution where all individual risk contributions are equal to a constant, and, therefore, to each other. The lack of a budget constraint in Model (3) provides

[2] An estimated covariance matrix of financial assets has the useful property of being symmetric and positive semidefinite, provided sufficient data is used for estimation.

flexibility to the decision variable $\boldsymbol{x}$ to match the risk contributions to b. However, there is no guarantee that the sum of weights will be equal to one. Instead, the optimal solution can be recovered as $x_i^* = x_i / \sum_{i=1}^n x_i$. Since $\boldsymbol{Q}$ is PSD and the sum of logarithms is strictly concave, the objective function in Model (3) is strictly convex. This, in turn, means any optimal solution will be a unique global solution. This is particularly relevant to us because, by design, we should be able to find up to 2^{n-1} risk parity solutions corresponding to every possible combination of $\beta_i x_i > 0$ for each $\beta_i \in \{-1, 1\}$ [1], with each solution corresponding to every possible long–short combination between the assets.

From an optimization perspective, all of these solutions are global optima. However, these solutions may lead to very different portfolios, with varying expected rates of return and risk. This highlights two fundamental deficiencies of the nominal risk parity framework. First, risk parity seeks only to equalize individual risk contributions, disregarding any impact this may have on the overall portfolio risk. Other things equal, an investor will always prefer to hold a portfolio with the lowest possible risk. The second deficiency is that any short positions we wish to hold must be predetermined before optimization takes place. Indeed, should we wish to find the most desirable risk parity solution, we will first need to find the subset of 2^{n-1} solutions before selecting the most desirable based on a secondary level of criteria (e.g., lowest portfolio variance, highest rate of return).

The introduction of the logarithmic barrier term in Model (3) presents a convex counterpart to the non-convex least squares Model (2). However, the sum of logarithms not only lacks financial relevance, it also restricts the feasible set to positive elements, $\beta_i x_i \geq 0$. Thus, any flexible framework that relaxes the conditions on long–short portfolios must exploit the individual asset risk contributions, R_i. Moreover, a generalized framework should consider overall portfolio risk and return, emphasizing the risk–return profile of an investor. We proceed to address these conditions in the next section.

3 Generalized risk parity

This section presents a generalized risk parity framework. This framework is designed so investors can attain an optimal risk–return profile while maintaining a certain degree of risk-based diversification. Moreover, it provides the flexibility to construct long–short portfolios. This falls in line with other generalized risk budgeting frameworks [17], but we note two important distinctions proposed in this paper. First, we allow the investor to preset a desired range of risk-based diversification instead of having to preset fixed risk budgets for its assets. This provides a degree of control over the dispersion of risk within the assets while offering greater flexibility to attain a more desirable risk–return profile. Second, we introduce a robust structure around the asset expected returns to reduce the impact of estimation error, which we discuss in Section 3.1.

The objective of our generalized risk parity framework is to minimize the portfolio variance, σ_p^2, while simultaneously maximizing portfolio return, μ_p. An investor's risk–return profile is defined by the preset trade-off coefficient $\lambda \in \mathbb{R}_+$. The investor can control his or her level of risk-based diversification by limiting the dispersion of the individual risk contributions, R_i. A simple yet highly tractable method to achieve this is to impose a limit on the difference between the highest and the lowest values of R_i. This is better described by the constraints given in the complete generalized risk parity framework, which is shown below

$$\begin{aligned} \min_{x,\theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} - \lambda \boldsymbol{\mu}^T \boldsymbol{x} \\ \text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 0, \quad \text{for } i = 1, ..., n \\ & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} - (1-c)\theta \geq 0, \quad \text{for } i = 1, ..., n \\ & \mathbf{1}^T \boldsymbol{x} = 1, \end{aligned} \tag{4}$$

where $\theta \in \mathbb{R}$ is an auxiliary variable that serves as a placeholder for the midpoint between the highest and the lowest asset risk contributions, and $c \in \mathbb{R}_+$ is a preset risk diversification coefficient that serves to bound the dispersion of the asset risk contributions. We note that setting $c = 0$ enforces perfect risk parity, while $c = +\infty$ collapses the problem to the nominal MVO model. Finally, the last constraint ensures all available wealth is invested. To the best of our knowledge, this is the first model to propose a relaxed risk-based diversification with tight bounds on the dispersion of the risk contribution per asset.

We note that the formulation in Model (4) is non-convex due to the constraints on the asset risk contributions. Thus, the model may be prone to converging to a local minimum due to the non-convexity of the feasible set. This is exemplified later in Section 5, where numerical experiments show the solution to this model tends to be far from a global optimum. The non-convexity of this model is addressed in Section 4 where we propose an algorithm to overcome this problem. For now, the remainder of this section presents two extensions of our generalized risk parity framework.

3.1 Robust generalized risk parity

As discussed in Section 1, the uncertainty arising from the estimation of parameters can lead to poor out-of-sample portfolio performance [2]. This, in turn, served as motivation for risk parity, a modern asset allocation framework that does not necessitate the expected returns as an input, avoiding the pitfalls arising from their uncertainty. Nevertheless, our proposed generalized risk parity framework allows an investor to consider both portfolio risk and return. To mitigate the impact of uncertainty, we impose a robust structure on the estimated expected returns. We start by introducing an ellipsoidal uncertainty set around the true (but unknown) expected returns as

$$\mathcal{U}_\mu = \left\{ \boldsymbol{\mu} \in \mathbb{R}^n : \ (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) \leq \omega^2 \right\},$$

where $\widehat{\boldsymbol{\mu}} \in \mathbb{R}^n$ is the vector of estimated asset expected returns, $\boldsymbol{\Omega} \in \mathbb{R}^{n \times n}$ is the covariance matrix of estimation errors for the vector of expected returns $\boldsymbol{\mu}$, and $\omega \in \mathbb{R}_+$ is a measure of distance that scales the uncertainty set in proportion to some probabilistic guarantee. Simply put, this uncertainty set states that the sum of squared deviations between the point estimate $\widehat{\boldsymbol{\mu}}$ and any other point in the set cannot be greater than ω^2. Instead of having individual confidence intervals for each estimate $\widehat{\mu}_i$, an ellipsoidal uncertainty set describes a joint confidence region [9].

The parameters $\boldsymbol{\Omega}$ and ω serve to quantify the error from estimation and the size of the uncertainty set, respectively. Different methods are available to calibrate these parameters. In this paper we present an example typically seen in robust portfolio optimization. The matrix Ω can be set as the diagonal matrix of squared standard errors, i.e.,

$$\boldsymbol{\Omega} = \frac{1}{T}\text{diag}(\boldsymbol{\Sigma}),$$

where $T \in \mathbb{R}_+$ is the number of observations used during estimation, and the operator diag$(\cdot)$ creates a diagonal matrix by setting every off-diagonal element to zero. If we approximate the joint confidence region as a sum of the squares of independent standard normal variables, we can use a χ^2-distribution with n degrees of freedom to determine ω^2, i.e.,

$$\omega^2 = \chi^2_n(\delta).$$

where δ is some confidence interval measuring the probability that $\boldsymbol{\mu}$ is within the set $\mathcal{U}_\mu$. For example, a popular choice in finance is to assign $\delta = 0.95$ to have a 95% confidence interval. However, in practice, using a χ^2-distribution may be prohibitively conservative. Additional detail on this calibration method, as well as other alternative methods, are shown in [9].

We proceed to use the uncertainty set to formulate a robust counterpart of the generalized risk parity framework

$$\begin{aligned}
\min_{x,\theta} \quad & \boldsymbol{x}^T\boldsymbol{\Sigma}\boldsymbol{x} - \lambda\left(\widehat{\boldsymbol{\mu}}^T\boldsymbol{x} - \omega||\boldsymbol{\Omega}^{1/2}\boldsymbol{x}||_2\right) \\
\text{s.t.} \quad & (1+c)\theta - \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} \geq 0, \quad \text{for } i = 1, ..., n \\
& \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} - (1-c)\theta \geq 0, \quad \text{for } i = 1, ..., n \\
& \mathbf{1}^T\boldsymbol{x} = 1,
\end{aligned} \tag{5}$$

where $||\cdot||_2$ denotes the Euclidean norm operator. The norm term in the objective function of Model (5) denotes the ellipsoidal uncertainty around the estimated portfolio return, $\widehat{\boldsymbol{\mu}}^T\boldsymbol{x}$, giving us the robust expression of the portfolio return.

3.2 Lowest variance risk parity

Thus far, our proposed generalized risk parity framework has emphasized two benefits over its nominal counterpart. First, it allows an investor to optimize with respect to a predetermined risk–return profile, attempting to minimize risk while maximizing return. Second, it allows for the construction of long–short portfolios.

However, in certain scenarios, an investor who lacks confidence or is unable to produce reliable estimates of the asset expected returns might wish to ignore portfolio returns and focus solely on risk-based diversification. In essence, such an investor will revert back to the nominal risk parity framework. As discussed in Section 2.1, there exist up to 2^{n-1} risk parity solutions, except some of these solutions may be more desirable than others if they have a lower portfolio variance. Indeed, the concept of the lowest variance risk parity portfolio is not new and is considered by Bai et al. [1], where a heuristic algorithm is proposed that sequentially moves from the minimum variance portfolio to the nearest risk parity portfolio.

We propose to use our generalized risk parity framework to solve this problem. We can do this by referencing Model (4), setting the trade-off coefficient and the risk diversification coefficient to zero, i.e., $\lambda = c = 0$. Thus, the lowest variance risk parity framework is

$$\begin{aligned} \min_{x,\theta} \quad & \boldsymbol{x}^T \boldsymbol{\Sigma} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} = \theta, \quad \text{for } i = 1, ..., n \\ & \mathbf{1}^T \boldsymbol{x} = 1. \end{aligned} \tag{6}$$

After setting $c = 0$, we can see that the $2n$ inequality constraints that bound the individual risk contributions collapse to n equality constraints, enforcing perfect risk parity. As with Models (4) and (5), this is also a non-convex optimization problem. Nevertheless, the issue of non-convexity can be overcome through an ADMM algorithm, which we present in the next section.

4 Sequential tightening via ADMM

This section presents a heuristic algorithm that sequentially approximates a global optimal solution to the non-convex generalized risk parity framework. We begin by relaxing the generalized risk parity framework from Model (4) as a SDP. To do so, we introduce a new variable $\boldsymbol{X} \in \mathbb{R}^{n \times n}$ as a convex relaxation of the square of our decision variable $\boldsymbol{x}$ such that $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T$. We then introduce a new decision variable $\boldsymbol{Y} \in \mathbb{R}^{(n+1)\times(n+1)}$ as an expression of the Schur complement between $\boldsymbol{x}$ and $\boldsymbol{X}$. Finally, we reformulate the input parameters to align with the dimensions of $\boldsymbol{Y}$,

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} \succeq 0, \quad \boldsymbol{Q} = \begin{bmatrix} \boldsymbol{\Sigma} & -\frac{\lambda}{2}\widehat{\boldsymbol{\mu}} \\ -\frac{\lambda}{2}\widehat{\boldsymbol{\mu}}^T & 0 \end{bmatrix}, \quad \boldsymbol{B}_i = \begin{bmatrix} \boldsymbol{A}_i & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \text{ for } i = 1, ..., n,$$

where $\mathbf{0} \in \mathbb{R}^n$ denotes a vector where all elements are equal to zero.

Expressing our variables and input parameters in this fashion allows us to relax Model (4) into a SDP,

$$\begin{aligned} \min_{Y,\theta} \quad & \mathrm{Tr}(\boldsymbol{Q}\boldsymbol{Y}) \\ \text{s.t.} \quad & (1+c)\theta - \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) \geq 0 \quad \text{for } i = 1, ..., n, \\ & \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) - (1-c)\theta \geq 0 \quad \text{for } i = 1, ..., n, \\ & \sum_{i=1}^{n} Y_{i,n+1} = 1, \\ & Y_{n+1,n+1} = 1, \\ & \boldsymbol{Y} \succeq 0, \end{aligned} \tag{7}$$

where $\mathrm{Tr}(\cdot)$ is the trace operator. For reference, we will refer to the convex set formed by the constraints of Model (7) as $\mathcal{S}_1$. By definition, this model is a convex relaxation, but it is not guaranteed to respect our constraints on the dispersion of the risk contribution per asset. We note that imposing a rank-1 constraint on Model (7) where $\text{rank}(\boldsymbol{Y}) = 1$ is equivalent to $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$, recovering our original non-convex model. The SDP relaxation removes the rank-1 condition, thereby relaxing the non-convex feasible set from Model (4) into a convex one.

Our proposed approach attempts to attain a rank-1 solution to Model (4) by sequentially solving a series of convex problems, each tightening an approximation to a rank-1 solution. This method is motivated by the ADMM algorithm proposed by You and Peng [33], where they applied a similar method in the context of optimal power flow. As described by Boyd et al. [3], the ADMM algorithm takes the form of a decomposition–coordination procedure, where the solution to smaller sub-problems are coordinated to find the solution to a larger problem. ADMM combines the benefits of the dual decomposition and augmented Lagrangian methods, where the problem is partitioned through dual decomposition, and reconciled through the augmented Lagrangian method. Depending on the application, the smaller sub-problems can reduce the complexity of the original problem, or at least reduce the computational burden of having to directly solve the larger problem. Iteratively coordinating and reconciling the sub-problems eventually leads to convergence, yielding a solution to the original problem.

To implement a suitable ADMM algorithm for the generalized risk parity framework, we begin by transferring the rank-1 requirement to an auxiliary variable $\boldsymbol{Z} \in \mathbb{R}^{(n+1)\times(n+1)}$, followed by the imposition two additional constraints: $\boldsymbol{Y} = \boldsymbol{Z}$, and $\text{rank}(\boldsymbol{Z}) = 1$. Imposing these constraints recovers a non-convex formulation equivalent to Model (4). However, the inclusion of the auxiliary variable $\boldsymbol{Z}$ enables us to decompose the problem into a convex sub-problem and a non-convex sub-problem. The convex sub-problem is the same as Model (7), except we use an augmented Lagrangian as our objective function

$$\begin{aligned} L_1(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}) &= \text{Tr}(\boldsymbol{Q}\boldsymbol{Y}) + \text{Tr}\big(\boldsymbol{\Lambda}^T(\boldsymbol{Y} - \boldsymbol{Z})\big) + \frac{\rho}{2}||\boldsymbol{Y} - \boldsymbol{Z}||_F^2 \\ &= \text{Tr}(\boldsymbol{Q}\boldsymbol{Y}) + \frac{\rho}{2}\big|\big|\boldsymbol{Y} - \big(\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda}\big)\big|\big|_F^2, \end{aligned}$$

where $||\cdot||_F$ is the Frobenius norm operator. The augmented Lagrangian adds both a Lagrangian term and a penalty term to our original objective function, thereby attempting to find a an optimal solution $\boldsymbol{Y}^*$ that approximates the rank-1 constrained auxiliary variable $\boldsymbol{Z}$. These two additional terms in our objective function can be expressed as a single term by completing the square between the Lagrangian and penalty terms, as we have shown above. The parameter $\rho \in \mathbb{R}_+$ is a tuning parameter and $\boldsymbol{\Lambda} \in \mathbb{R}^{(n+1)\times(n+1)}$ is the dual variable corresponding to our self-imposed constraint $\boldsymbol{Y} = \boldsymbol{Z}$. Through this method, we transfer the non-convexity of the problem to the variable $\boldsymbol{Z}$.

The ADMM algorithm iterates through the following steps:

1) Convex $\boldsymbol{Y}$-minimization: $\quad \boldsymbol{Y}^{k+1} = \underset{Y,\theta \in \mathcal{S}_1}{\text{argmin}}\, L_1(\boldsymbol{Y}, \boldsymbol{Z}^k, \boldsymbol{\Lambda}^k)$,

2) Non-convex $\boldsymbol{Z}$-minimization: $\quad \boldsymbol{Z}^{k+1} = \underset{\text{rank}(\boldsymbol{Z}) \leq 1}{\text{argmin}}\, L_1(\boldsymbol{Y}^{k+1}, \boldsymbol{Z}, \boldsymbol{\Lambda}^k)$,

3) Dual variable $\boldsymbol{\Lambda}$-update: $\quad \boldsymbol{\Lambda}^{k+1} = \boldsymbol{\Lambda}^k + \rho(\boldsymbol{Y}^{k+1} - \boldsymbol{Z}^{k+1})$.

The $\boldsymbol{Y}$-minimization step minimizes L_1 subject to the convex set $\mathcal{S}_1$. For clarity, we present the complete convex $\boldsymbol{Y}$-minimization step

$$
\begin{aligned}
\min_{Y,\theta,\zeta} \quad & \mathrm{Tr}(\boldsymbol{QY}) + \frac{\rho}{2}\zeta \\
\text{s.t.} \quad & 0 \leq (1+c)\theta - \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) \quad \text{for } i = 1, ..., n, \\
& 0 \leq \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) - (1-c)\theta \quad \text{for } i = 1, ..., n, \\
& 1 = \sum_{i=1}^{n} Y_{i,n+1}, \\
& 1 = Y_{n+1,n+1}, \\
& 0 \preceq \begin{bmatrix} \boldsymbol{I} & \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k)) \\ \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k))^T & \zeta \end{bmatrix}, \\
& 0 \preceq \boldsymbol{Y},
\end{aligned}
\tag{8}
$$

where $\mathrm{vec}(\cdot)$ is an operator that vectorizes a matrix, $\boldsymbol{I}$ is an identity matrix of appropriate size, and the auxiliary variable $\zeta \in \mathbb{R}$ serves to write the square of the Frobenius norm term in the objective as a SDP constraint through its Schur complement

$$
||\boldsymbol{Y} - (\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda})||_F^2 \leq \zeta \iff \begin{bmatrix} \boldsymbol{I} & \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda})) \\ \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda}))^T & \zeta \end{bmatrix} \succeq 0.
$$

The $\boldsymbol{Z}$-minimization step consists of projecting the matrix $\boldsymbol{Y}^{k+1} + \frac{1}{\rho}\boldsymbol{\Lambda}^k$ onto a rank constrained set, i.e., projecting it onto a non-convex set. However, rank-constrained problems can be solved analytically through the Eckart–Young–Mirsky theorem [3, 33]. We can find the optimal rank-1 approximation by singular value decomposition (SVD)

$$
\boldsymbol{Z}^{k+1} = s_1 v_1 v_1^T,
$$

where $s_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^{n+1}$ are the top singular value and vector of the matrix $\boldsymbol{Y}^{k+1} + \frac{1}{\rho}\boldsymbol{\Lambda}^k$. Thus, the structure of this rank-constrained approach allows us to efficiently solve the otherwise difficult non-convex optimization step. Finally, the $\boldsymbol{\Lambda}$-update step is straightforward and can be carried out as shown above.

The SDP relaxation can be tightened towards a rank-1 solution by repeatedly iterating through the three steps outlined above. Once we converge to a rank-1 solution, the SDP is no longer considered a relaxation. Any optimal rank-1 solution will also be within the original feasible set, i.e., this solution will respect the risk dispersion constraints. Our motivation behind this algorithm is the manner in which the optimal solution is attained. Starting with a relaxed model allows us to find a lower bound to our optimal solution. Although a lower value is more desirable, in practice this lower bound is seldom feasible. Nevertheless, after identifying the lower bound, we can proceed sequentially towards the first available feasible solution. By design, this attempts to find a feasible solution that is as close as possible to the lower bound. This solution may be the global optimum, or, at the very least, we often find this a high quality optimal solution.

We note that the generalized risk parity framework and its ADMM algorithm are able to accommodate additional convex constraints, provided that these constraints still allow us to reformulate the problem as a SDP. Thus, practical considerations such as specific limits on short selling or weight concentration limits can be easily incorporated into our framework. However, the implementation of these variants is beyond the scope of this paper.

4.1 ADMM steps: Robust generalized risk parity

We can follow the same approach and relax the robust generalized risk parity framework from Model (5) a SDP. The objective function of the SDP relaxation uses the same decision variable $\boldsymbol{Y}$ and input parameter $\boldsymbol{Q}$, with the addition of the robust term that captures the ellipsoidal uncertainty around the estimated portfolio return. The corresponding augmented Lagrangian of this model is

$$L_2(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\Lambda}) = \mathrm{Tr}(\boldsymbol{QY}) + \lambda\omega ||\boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1}||_2 + \frac{\rho}{2}\big|\big|\boldsymbol{Y} - \big(\boldsymbol{Z} - \frac{1}{\rho}\boldsymbol{\Lambda}\big)\big|\big|_F^2.$$

The ADMM algorithm requires the same three steps as before. The inclusion of the robust term in the augmented Lagrangian L_2 requires that the first step, namely the $\boldsymbol{Y}$-minimization step, is revised to the following SDP

$$\begin{aligned}
\min_{Y,\theta,\tau,\zeta} \quad & \mathrm{Tr}(\boldsymbol{QY}) + \lambda\omega\tau + \frac{\rho}{2}\zeta \\
\text{s.t.} \quad & 0 \le (1+c)\theta - \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) \quad \text{for } i = 1, ..., n, \\
& 0 \le \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) - (1-c)\theta \quad \text{for } i = 1, ..., n, \\
& 1 = \sum_{i=1}^{n} Y_{i,n+1}, \\
& 1 = Y_{n+1,n+1}, \\
& 0 \preceq \begin{bmatrix} \tau\boldsymbol{I} & \boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1} \\ (\boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1})^T & \tau \end{bmatrix}, \\
& 0 \preceq \begin{bmatrix} \boldsymbol{I} & \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k)) \\ \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k))^T & \zeta \end{bmatrix}, \\
& 0 \preceq \boldsymbol{Y},
\end{aligned} \tag{9}$$

where the auxiliary variable $\tau \in \mathbb{R}$ serves to write the Euclidean norm corresponding to the robust term from from Model (5) as a SDP constraint through its Schur complement

$$||\boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1}||_2 \le \tau \iff \begin{bmatrix} \tau\boldsymbol{I} & \boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1} \\ (\boldsymbol{\Omega}^{1/2}\mathbf{Y}_{1:n,n+1})^T & \tau \end{bmatrix} \succeq 0.$$

The $\boldsymbol{Z}$-minimization and $\boldsymbol{\Lambda}$-update steps pertain to the sequential approximation of a rank-1 solution, and are unaffected by the inclusion of the robust term in the in the augmented Lagrangian L_2. Thus, these two steps remain the same as before. For reference, we will refer to the convex set described by the constraints in Model (9) as $\mathcal{S}_2$.

4.2 ADMM steps: Lowest variance risk parity

As discussed in Section 3.2, the lowest variance risk parity from Model (6) can be formulated as a special case of the generalized risk parity framework with the trade-off coefficient and the risk diversification coefficient set to zero. In turn, the corresponding ADMM steps can be set by following the same methodology as before, using Model (8) as the $\boldsymbol{Y}$-minimization step with $\lambda = c = 0$. However, we can take advantage of these conditions to simplify our problem. First, we describe a new input parameter

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)},$$

aligning the covariance matrix with the dimensions of the decision variable $\boldsymbol{Y}$. Next, we present the corresponding $\boldsymbol{Y}$-minimization step. Recasting the SDP in this form is beneficial because we take advantage of the fact that $\lambda = c = 0$ to simplify the problem. The lowest variance risk parity $\boldsymbol{Y}$-minimization step is

$$\begin{aligned} \min_{Y,\theta,\zeta} \quad & \mathrm{Tr}(\boldsymbol{MY}) + \frac{\rho}{2}\zeta \\ \text{s.t.} \quad & \theta = \mathrm{Tr}(\boldsymbol{B}_i\boldsymbol{Y}) \quad \text{for } i = 1, ..., n, \\ & 1 = \sum_{i=1}^{n} Y_{i,n+1}, \\ & 1 = Y_{n+1,n+1}, \\ & 0 \preceq \begin{bmatrix} \boldsymbol{I} & \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k)) \\ \mathrm{vec}(\boldsymbol{Y} - (\boldsymbol{Z}^k - \frac{1}{\rho}\boldsymbol{\Lambda}^k))^T & \zeta \end{bmatrix}, \\ & 0 \preceq \boldsymbol{Y}. \end{aligned} \tag{10}$$

As in Section 4.1, the changes brought on by the lowest variance risk parity scenario have no impact on the $\boldsymbol{Z}$-minimization and $\boldsymbol{\Lambda}$-update steps, meaning these remain unchanged. As before, we will refer to the convex set described by the constraints of Model (10) as $\mathcal{S}_3$.

4.3 ADMM algorithm

Here we describe a tractable implementation of the ADMM algorithm. To do so, we introduce several new parameters pertaining to the tuning of the iteration step size. The selection of parameters in ADMM is the source of ongoing research (e.g., see [3, 13, 30]), and there is no clear consensus on how to set them precisely.

To set the algorithm stopping criteria and the penalty parameter *rho*, we borrow some insight from Boyd et al. [3]. The stopping criteria can be determined by the sizes of the primal and dual residuals. In the context of ADMM, we have

$$\begin{aligned} \varepsilon_p^{k+1} &= ||\boldsymbol{Y}^{k+1} - \widehat{\boldsymbol{Z}}^{k+1}||_F, \\ \varepsilon_d^{k+1} &= \rho_k||\widehat{\boldsymbol{Z}}^{k+1} - \widehat{\boldsymbol{Z}}^{k}||_F, \end{aligned}$$

where the primal residual ε_p measures the difference between our solution to the convex step $\boldsymbol{Y}$ and its rank-1 approximation $\boldsymbol{Z}$, and the dual residual ε_d measures the progress attained by $\boldsymbol{Z}$ between each iteration. We can use the residuals to measure the quality of our convergence and to determine whether to stop our search for a rank-1 solution. A desirable tolerance for the residuals can be set by the user. In our case, we found that having $\varepsilon_p^k \leq 10^{-6}$ was sufficient to guarantee a good rank-1 approximation.

We proceed to discuss the penalty parameter ρ. At every iteration, the penalty parameter serves to measure the trade-off between the primal and dual residuals. Intuitively, a large penalty emphasizes our search for a solution that satisfies the rank-1 condition, reducing our primal residual. However, this has the effect of increasing our dual residual. From a practical standpoint, an aggressive penalty has the additional drawback of forcing the algorithm to take large steps in pursuing a rank-1 solution, limiting our search of an optimal portfolio to a smaller space within the feasible set. Thus, we initialize the algorithm with a small penalty of $\rho_0 = 0.005$, and we update the penalty parameter after each iteration using the scheme outlined in [3]

$$\rho_{k+1} = \begin{cases} \gamma\rho_k & \text{if } \varepsilon_p^{k+1} > \eta\varepsilon_d^{k+1} \\ \rho_k/\gamma & \text{if } \varepsilon_d^{k+1} > \eta\varepsilon_p^{k+1} \\ \rho_k & \text{otherwise} \end{cases}$$

where $\gamma \geq 1$ and $\eta \geq 1$ are two additional parameters that allow us to quantify the permissible level of variation of our penalty parameter. To provide a smooth change per iteration, we have found that the parameters $\gamma = 1.05$ and $\eta = 10$ work well in practice.

In an effort to improve computational performance, we also implement the 'fast ADMM' algorithm proposed in [16] and [13]. The 'fast ADMM' algorithm is applicable to convex problems, but can be applied to our problem given that our non-convex $\boldsymbol{Z}$-minimization step can be carried out exactly and efficiently. The improvements outlined by this algorithm are based on Nesterov's gradient decent method [27]. This variant of gradient descent is accelerated through an over-relaxation step. The fast ADMM algorithm applies a similar technique, and prescribe relaxation of the $\boldsymbol{Z}$-minimization and $\boldsymbol{\Lambda}$-update steps. For clarity, we outline the relaxation parameters α and β from [13]

$$\begin{aligned} \beta^{k+1} &= \frac{1 + \sqrt{1 + 4(\beta^k)^2}}{2} \\ \alpha^{k+1} &= \begin{cases} 1 + \frac{\beta^k - 1}{\beta^{k+1}} & \text{if } \frac{\max(||\varepsilon_p^k||, ||\varepsilon_d^k||)}{\max(||\varepsilon_p^{k-1}||, ||\varepsilon_d^{k-1}||)} < 1, \\ 1 & \text{otherwise.} \end{cases} \end{aligned}$$

These parameters serve to relax the our ADMM steps

$$\begin{aligned} \widehat{\boldsymbol{Z}}^{k+1} &= \alpha^{k+1}\boldsymbol{Z}^{k+1} + (1 - \alpha^{k+1})\boldsymbol{Z}^k, \\ \widehat{\boldsymbol{\Lambda}}^{k+1} &= \alpha^{k+1}\boldsymbol{\Lambda}^{k+1} + (1 - \alpha^{k+1})\boldsymbol{\Lambda}^k. \end{aligned}$$

The implementation of the fast ADMM algorithm is not necessary for convergence. However, it improves the convergence rate of the algorithm, which we found particularly desirable when dealing with large-scale problems.

Finally, we present the complete ADMM algorithm below in Algorithm 1. Depending on the user's choice of problem, we can select to solve (a) the generalized risk parity, (b) the robust generalized risk parity, or (c) the lowest variance risk parity. Moreover, if a user does not wish to use the fast ADMM algorithm, the corresponding steps can be ignored.

Algorithm 1 Fast ADMM for generalized risk parity

1. Initialize ADMM algorithm
   - 1.1 Set the counter: $k = 0$,
   - 1.2 Set the penalty parameters: $\rho_0 = 0.01$, $\gamma = 1.1$, $\eta = 5$.
   - 1.3 Set the primal and dual residuals: $\varepsilon_p^0 = \varepsilon_d^0 = 1$.
   - 1.4 Set the fast ADMM parameters $\alpha^0 = \beta^0 = 1$.
   - 1.5 Select the appropriate $\boldsymbol{Y}$-minimization step:

     (a) Generalized risk parity: $\boldsymbol{Y}^0 = \underset{Y,\theta,\zeta \in \mathcal{S}_1}{\text{argmin}}\ \text{Tr}(\boldsymbol{QY})$

     (b) Robust gen. risk parity: $\boldsymbol{Y}^0 = \underset{Y,\theta,\zeta,\tau \in \mathcal{S}_2}{\text{argmin}}\ \text{Tr}(\boldsymbol{QY}) + \lambda\omega\tau$

     (c) Lowest variance: $\boldsymbol{Y}^0 = \underset{Y,\theta,\zeta \in \mathcal{S}_3}{\text{argmin}}\ \text{Tr}(\boldsymbol{MY})$

   - 1.6 $\boldsymbol{Z}$-minimization: $\boldsymbol{Z}^0 = s_1 \boldsymbol{v}_1 \boldsymbol{v}_1^T$ through the SVD of $\boldsymbol{Y}^0$
   - 1.7 $\boldsymbol{\Lambda}$-update: $\boldsymbol{\Lambda}^0 = \rho_0(\boldsymbol{Y}^0 - \boldsymbol{Z}^0)$.
   - 1.8 Set the fast ADMM variables $\widehat{\boldsymbol{Z}}^0 = \boldsymbol{Z}^0$ and $\widehat{\boldsymbol{\Lambda}}^0 = \boldsymbol{\Lambda}^0$.

2. While $\varepsilon_p^k \geq 10^{-6}$
   - 2.1 Select the appropriate $\boldsymbol{Y}$-minimization step:
     - (a) Find $\boldsymbol{Y}^{k+1}$ by solving Model (8) with $\widehat{\boldsymbol{Z}}^k$ and $\widehat{\boldsymbol{\Lambda}}^k$
     - (b) Find $\boldsymbol{Y}^{k+1}$ by solving Model (9) with $\widehat{\boldsymbol{Z}}^k$ and $\widehat{\boldsymbol{\Lambda}}^k$
     - (c) Find $\boldsymbol{Y}^{k+1}$ by solving Model (10) with $\widehat{\boldsymbol{Z}}^k$ and $\widehat{\boldsymbol{\Lambda}}^k$
   - 2.2 Find $\boldsymbol{Z}^{k+1} = s_1 v_1 v_1^T$ through the SVD of $\left(\boldsymbol{Y}^{k+1} + \frac{1}{\rho_k}\widehat{\boldsymbol{\Lambda}}^k\right)$.
   - 2.3 $\boldsymbol{\Lambda}^{k+1} = \widehat{\boldsymbol{\Lambda}}^k + \rho_k(\boldsymbol{Y}^{k+1} - \boldsymbol{Z}^{k+1})$.
   - 2.4 Update the Fast ADMM parameters and variables
     - 2.4.1 $\beta^{k+1} = \frac{1+\sqrt{1+4(\beta^k)^2}}{2}$.
     - 2.4.2 $\alpha^{k+1} = \begin{cases} 1 + \frac{\beta^k - 1}{\beta^{k+1}} & \text{if } \frac{\max(\|\varepsilon_p^k\|, \|\varepsilon_d^k\|)}{\max(\|\varepsilon_p^{k-1}\|, \|\varepsilon_d^{k-1}\|)} < 1, \\ 1 & \text{otherwise.} \end{cases}$.
     - 2.4.3 $\widehat{\boldsymbol{Z}}^{k+1} = \alpha^{k+1}\boldsymbol{Z}^{k+1} + (1 - \alpha^{k+1})\boldsymbol{Z}^k$.
     - 2.4.4 $\widehat{\boldsymbol{\Lambda}}^{k+1} = \alpha^{k+1}\boldsymbol{\Lambda}^{k+1} + (1 - \alpha^{k+1})\boldsymbol{\Lambda}^k$.
   - 2.5 $\varepsilon_p^{k+1} = \|\boldsymbol{Y}^{k+1} - \widehat{\boldsymbol{Z}}^{k+1}\|_F$.
   - 2.6 $\varepsilon_d^{k+1} = \rho_k \|\widehat{\boldsymbol{Z}}^{k+1} - \widehat{\boldsymbol{Z}}^k\|_F$.

2.7 $\rho_{k+1} = \begin{cases} \gamma\rho_k & \text{if } \varepsilon_p^{k+1} > \eta\varepsilon_d^{k+1} \\ \rho_k/\gamma & \text{if } \varepsilon_d^{k+1} > \eta\varepsilon_p^{k+1} \\ \rho_k & \text{otherwise} \end{cases}.$

2.8 $k = k + 1$.

5 Numerical experiments

This section presents the results of multiple numerical experiments to test the generalized risk parity framework and the two extensions proposed, namely the robust generalized risk parity and lowest variance risk parity frameworks. There are three different experiments, one for each of the risk parity frameworks. Each experiment is composed of four individual trials.

The three experiments share the following structure. There are four trials per experiment, with the number of assets increasing per trial: $n = 33$, $n = 50$, $n = 75$, and $n = 100$. The first two trials, where $n = 33$ and $n = 50$, uses historical data from a pool of U.S. stocks belonging to the S&P 500 index, with stocks from each of the eleven *Global Industry Classification Standard* (GICS) sectors. A list of these stocks is presented in Table 1. When $n = 33$ we use the stocks listed in the first three columns of the table. When $n = 50$ we use all the stocks listed in the table. The relevant input parameters, namely $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, were estimated using weekly excess returns from 01-Jan-2007 to 31-Dec-2009. The historical stock prices were obtained from [29]. The third and fourth trials, where $n = 75$ and $n = 100$, use randomly generated synthetic data to construct the input parameters. The only condition imposed on these simulations is that the asset expected returns are positive and the corresponding covariance matrix is PSD.

Table 1 List of assets

GICS Sector	**Company Tickers**				
Consumer Disc.	DIS	F	MCD	GPS	NKE
Consumer Staples	WMT	KO	KR	PG	CL
Financials	JPM	BAC	C	AON	WFC
Healthcare	BMY	PFE	JNJ	LLY	HUM
Industrials	BA	CAT	GE	LMT	MMM
Information Tech.	AAPL	HPQ	IBM	ORCL	QCOM
Materials	IP	MOS	NEM	PPG	PX
Energy	XOM	HAL	OXY	MRO	
Utilities	CNP	DTE	DUK	ED	
Real Estate	HCP	REG	UDR	WY	
Telecom.	S	T	VZ		

The numerical results are evaluated by three measures. First, we compare their objective values after optimization, with a lower number considered better since these are minimization problems. We note that the objective value is based on the risk–return profile alone, excluding the value of the Lagrangian penalty term brought on by the augmented Lagrangian method. In the case of the robust framework, we consider only the robust risk–return profile. For the lowest variance framework, we only consider the portfolio variance as the objective value.

The second measure of performance is the coefficient of variation (CV) of the resulting portfolios, which quantifies our risk-based diversification. The CV is calculated by dividing the standard deviation of the risk contributions by their average, i.e.,

$$\mathrm{CV} = \frac{\mathrm{SD}(\boldsymbol{x} \odot (\boldsymbol{\Sigma x}))}{\frac{1}{n}\boldsymbol{x}^T \boldsymbol{\Sigma x}},$$

where '$\odot$' is the element-wise multiplication operator and $\mathrm{SD}(\cdot)$ computes the standard deviation of the corresponding vector. A lower CV indicates a higher degree of risk-based diversification. In theory, the solution to the nominal risk parity problem from Model (2) should yield a CV of zero. The risk diversification coefficient c controls the dispersion of the asset risk contributions, approximately bounding the CV. For example, if we have $c = 0.2$, we expect the CV of the resulting portfolios to be $CV \approx 0.2$.

Finally, our third measure is the model runtime. This is a relative measure to compare the speed at which the different optimization models are able to produce an optimal solution. We consider this a relative measure due to the availability of our computing power. Thus, it is only useful in the context of the different models tested during these experiments.

All computations were performed on an Apple MacBook Pro computer (2.8 GHz Intel Core i7, 16 GB 2133 MHz DDR3 RAM) running macOS 'Mojave'. The computer script was written in the Julia programming language (version 1.1.0) using the optimization modeling language JuMP [8] with MOSEK (version 8.1.0.45) as the optimization solver for the SDPs and IPOPT (version 3.12.6) as the solver for the non-convex problems.

5.1 Generalized risk parity: numerical results

We conducted four trials to evaluate the generalized risk parity framework, with $n = 33$, $n = 50$, $n = 75$ and $n = 100$ as the number of assets per trial. The trials were conducted with a risk–return trade-off coefficient $\lambda = 0.1$ and with different values of the risk diversification coefficient c, as noted in the headers of Table 2. For comparison, we tested the following competing optimization models:

i) Nominal MVO: this model is a relaxation of Model (4), where we maintain the investor's risk–return profile as the objective, but remove the $2n$ risk-based diversification constraints.

ii) SDP: the SDP from Model (7), which is the convex relaxation of the generalized risk parity framework.

iii) Non-convex: the original non-convex formulation of the generalized risk parity framework, corresponding to Model (4).

iv) NC warm: the same as the non-convex Model (4), except we warm-start the problem with the solution to the SDP relaxation.

v) ADMM: the generalized risk parity framework prescribed by Algorithm 1 (a).

Table 2 presents a summary of the results for all five optimization models. The objective value of the MVO model serves as a benchmark, i.e., it is the optimal value we would attain if we did not impose any risk-based diversification conditions. From our perspective, a good generalized risk parity framework should come as close as possible to this value. Not surprisingly, the objective value of the SDP relaxation comes as close as possible to that of the MVO model, with the caveat that it violates the risk diversification constraints. Nevertheless, the SDP relaxation serves to set a lower bound on the generalized risk parity framework. The non-convex model is able to find a local optimal solution that complies with the risk dispersion constraints, but it appears to be far from the global optimal solution. This becomes apparent when we compare it to the warm-start non-convex model. However, the ADMM algorithm attains the most desirable objective value while still respecting the risk dispersion bounds. While we are not able to provide a theoretical guarantee on global optimality, we note that the ADMM algorithm is able to provide a better solution than its non-convex counterparts. We note that this conclusion holds for all three experiments.

Table 2 Summary of results for the generalized risk parity framework

	$n = 33$, $c = 0.25$, $\lambda = 0.1$				
	MVO	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.0040	0.0157	0.0785	6.810	0.0376
CV	2.59	1.15	0.249	0.256	0.254
Runtime (sec)	0.003	0.023	0.045	0.055	51.04
	$n = 50$, $c = 0.15$, $\lambda = 0.1$				
	MVO	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.0301	-0.0171	0.0977	0.0708	0.0188
CV	2.93	1.36	0.153	0.151	0.140
Runtime (sec)	0.006	0.054	0.090	0.085	1,376.0
	$n = 75$, $c = 0.2$, $\lambda = 0.1$				
	MVO	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.545	-0.483	-0.354	-0.417	-0.451
CV	1.62	0.525	0.202	0.200	0.198
Runtime (sec)	0.023	0.171	0.238	0.266	807.8
	$n = 100$, $c = 0.2$, $\lambda = 0.1$				
	MVO	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.631	-0.579	1.47	-0.428	-0.545
CV	1.21	0.472	0.202	0.199	0.203
Runtime (sec)	0.029	0.379	0.455	0.597	2,577.0

Improving the optimal objective value through the ADMM algorithm comes at an increased computational cost. The runtime of the ADMM algorithm is significantly higher than the other models, which is due to two reasons. First, the ADMM algorithm requires multiple iterations to converge to a desirable solution, meaning we must solve a sequence of convex optimization problems. Second, each iteration requires the solution to a SDP, which cannot be solved as efficiently as quadratic programs or second-order cone programs. This is particularly true for large-scale SDPs, as shown by the runtimes in Table 2.

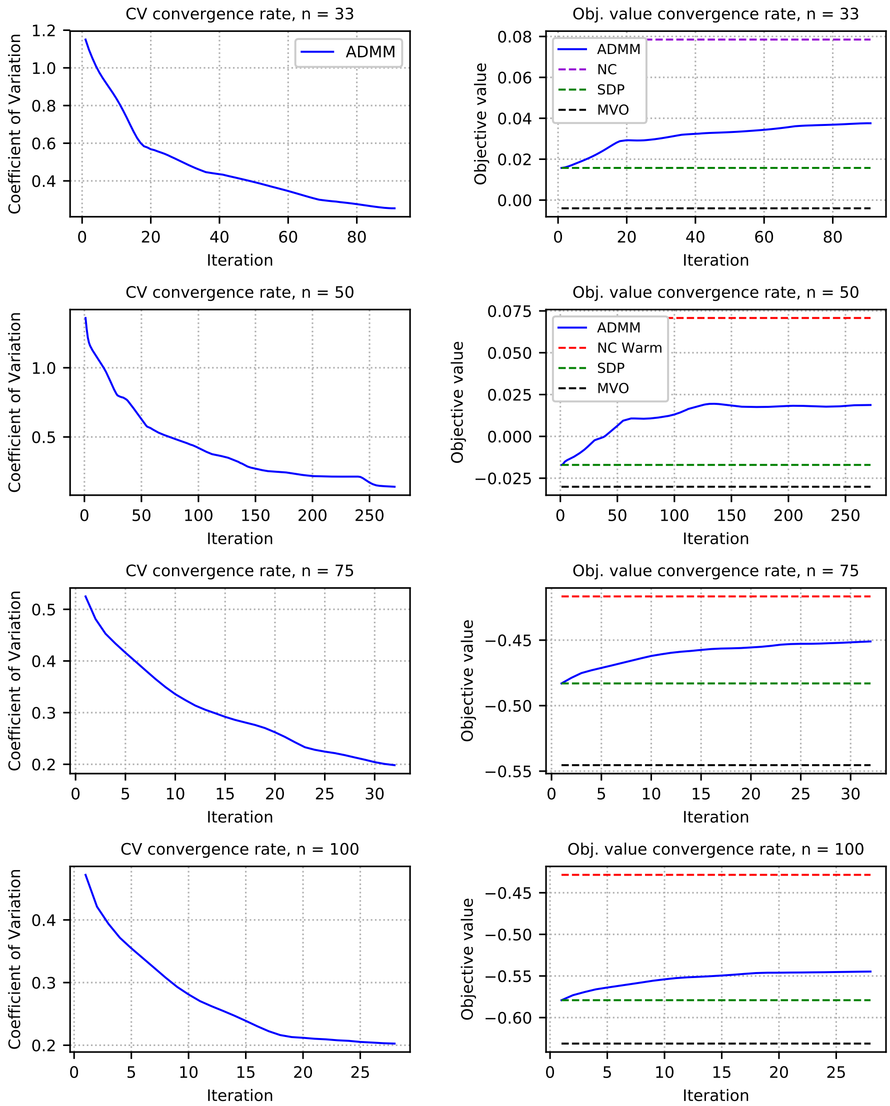

Fig. 1 Convergence plots for the generalized risk parity framework

Figure 1 presents the convergence rate of the ADMM algorithm for all four trials. The plots on the left-hand side show the convergence of the CV. For all four trials, we can see how the CV converges towards the desired risk dispersion limit c, where the value of c is shown in Table 2 for each trial. The plots on the right-hand side display the objective value of the ADMM algorithm after each iteration. For comparison, these plots also show the optimal values of the MVO model, the SDP relaxation, and either the non-convex model or the warm-start non-convex model. Between the two non-convex models, only the one with the most desirable objective value is shown. As a reminder, we note that the objective value of the ADMM algorithm is based on the risk–return profile objective, excluding any impact from the augmented Lagrangian terms.

The plots on the right-hand side of Figure 1 show the objective value of the ADMM algorithm after each iteration. These plots show the ADMM algorithm initially departs from the same point as the SDP relaxation. However, the objective value increases as we tighten the solution around a rank-1 constraint. Nevertheless, we can also see that the algorithm converges to a more desirable point than either of the non-convex models. We note that during the first trial, with $n = 33$, the warm-start non-convex model actually converges to a worse local optimum when compared to the regular non-convex model. This, in turn, highlights that a warm start will not always lead to a better solution. While the ADMM algorithm cannot guarantee global optimality, the path taken by algorithm shows we start from a theoretical lower bound, sequentially sacrificing optimality while seeking feasibility in the original problem. The ADMM algorithm will stop as soon as it finds the first rank-1 solution, i.e., the first solution that satisfies feasibility of the original problem. Since we approach feasibility from the theoretical lower bound, this method often converges to a high quality, if not the global, optimal solution.

5.2 Robust generalized risk parity: numerical results

Our second experiment evaluates the robust generalized risk parity framework. We construct the robust structure with a 90% confidence interval around the synthetic asset expected returns. As with the previous experiment, we tested five different optimization models:

i) Robust MVO: this model is a relaxation of Model (5), removing the $2n$ risk-based diversification constraints.

ii) SDP: the SDP from Model (9), which is the relaxation of the robust generalized risk parity framework.

iii) Non-convex: the non-convex formulation from Model (5).

iv) NC warm: warm-start Model (5) with the solution to the SDP relaxation.

v) ADMM: the robust ADMM algorithm prescribed by Algorithm 1 (b).

As noted in Table 3, all four trials were conducted with $\lambda = 0.1$ and with varying values of c. The results shown in Table 3 and Figure 2 follow the same pattern as those shown in Section 5.1. Once again, we can see the ADMM algorithm is able to attain a higher quality optimal solution when compared to the competing non-convex models. However, we also find that the additional complexity from the

Table 3 Summary of results for the robust generalized risk parity framework

	$n = 33$, $c = 0.25$, $\lambda = 0.1$				
	Robust	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	0.099	0.110	0.536	0.203	0.146
CV	1.72	1.06	0.258	0.253	0.253
Runtime (sec)	0.026	0.026	0.049	0.061	103.1
	$n = 50$, $c = 0.15$, $\lambda = 0.1$				
	Robust	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	0.094	0.103	0.223	0.259	0.141
CV	1.75	1.12	0.150	0.150	0.150
Runtime (sec)	0.040	0.253	0.101	0.103	1,376.3
	$n = 75$, $c = 0.2$, $\lambda = 0.1$				
	Robust	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.271	-0.265	7.58	0.697	-0.241
CV	0.976	0.724	0.205	0.204	0.198
Runtime (sec)	0.055	0.177	0.260	0.264	3,572.0
	$n = 100$, $c = 0.2$, $\lambda = 0.1$				
	Robust	SDP	Non-Convex	NC (Warm)	ADMM
Obj. Value	-0.281	-0.27	0.244	0.219	-0.233
CV	1.130	0.847	0.202	0.201	0.205
Runtime (sec)	1.155	0.536	0.610	0.547	13,770.2

robust formulation increases the runtime of the ADMM algorithm when compared to the non-robust ADMM algorithm from before. We found that the ADMM algorithm necessitated more iterations than the non-robust version, and it also took longer per iteration. Nevertheless, it is still able to deliver a high quality solution within reasonable time.

5.3 Lowest variance risk parity: numerical results

The third experiment evaluates the lowest variance risk parity framework. As before, we tested five different optimization models:

i) Risk parity: this is the nominal risk parity framework from Model (2), and it imposes the long-only restriction. The objective of this model is to equalize risk contributions, and does not attempt to minimize risk explicitly.

ii) SDP: the SDP from Model (10), which is the relaxation of the lowest variance risk parity framework.

iii) Non-convex: the non-convex formulation from Model (6).

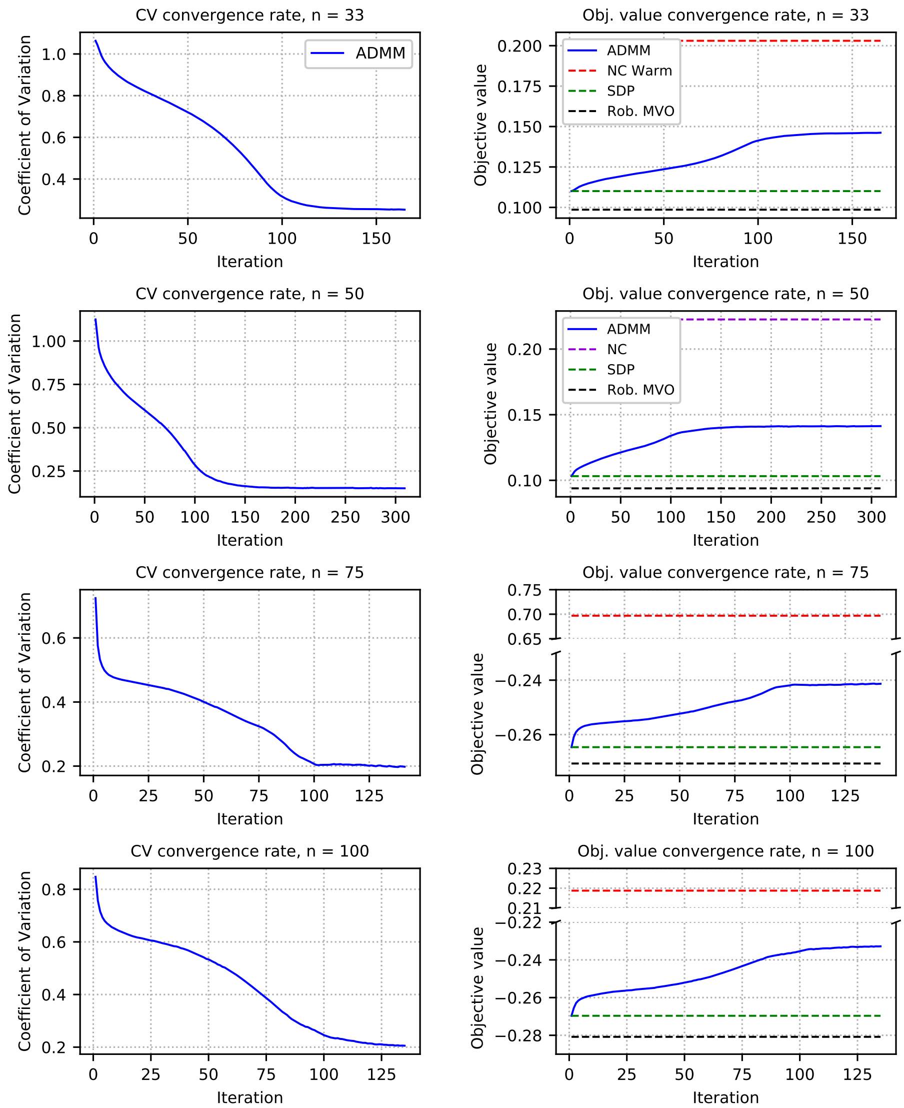

Fig. 2 Convergence plots for the robust generalized risk parity framework

iv) NC warm: warm-start Model (6) with the solution to the SDP relaxation.

v) ADMM: the algorithm prescribed by Algorithm 1 (c).

Unlike our previous experiments, the nominal risk parity model does not share the same objective as the rest of the competing models. Given that all other competing models seek to minimize variance, the nominal risk parity model is not expected to provide a comparable optimal objective value. Instead, its purpose is to show the portfolio variance attained if we attempt to equalize the asset risk contributions subject to long-only constraints, i.e., it highlights the opportunity forfeited when we disallow short selling.

Table 4 Summary of results for the lowest variance risk parity framework

	$n = 33$				
	Risk parity	SDP	Non-Convex	NC (Warm)	ADMM
Variance	0.115	0.0465	0.0738	0.0699	0.067
CV	2e-12	1.23	0.0005	0.0005	0.0091
Runtime (sec)	0.007	0.016	0.044	0.035	72.51
	$n = 50$				
	Risk parity	SDP	Non-Convex	NC (Warm)	ADMM
Variance	0.105	0.0322	0.0601	0.0597	0.0569
CV	2e-15	1.76	0.0008	0.0008	0.0087
Runtime (sec)	0.816	0.032	0.634	0.089	786.8
	$n = 75$				
	Risk parity	SDP	Non-Convex	NC (Warm)	ADMM
Variance	0.00539	0.00397	0.00565	0.00560	0.00506
CV	1e-13	1.01	0.0132	0.0133	0.0098
Runtime (sec)	0.025	0.076	0.309	0.6774	3,854.0
	$n = 100$				
	Risk parity	SDP	Non-Convex	NC (Warm)	ADMM
Variance	0.00387	0.00295	0.00358	0.00358	0.00356
CV	3e-13	0.682	0.028	0.028	0.010
Runtime (sec)	0.082	0.160	0.466	0.368	10,530.3

A summary of the numerical results is given in Table 4. Unlike previous experiments, this table displays the portfolio variance instead of the objective value. This serves to allow for a comparison between the nominal risk parity model and the lowest variance risk parity models. By design, we expect the lowest variance risk parity models to have a lower or at least equal variance when compared to the nominal. However, the trial with $n = 75$ exemplifies the weakness of the two non-convex models. The non-convex model is able to find a local optimum to the long-short risk parity problem, but the non-convexity of the feasible set means it converges to a portfolio with higher variance. As shown in the results, warm starting the problem marginally improves our optimal solution, but it cannot reliably remedy this weakness.

On the other hand, the numerical results show that the lowest variance risk parity ADMM algorithm can consistently deliver risk parity portfolios with lower variance when compared to the remaining models. By definition, the SDP relaxation establishes a lower bound on the lowest variance attainable. Figure 3 shows the how the algorithm sacrifices optimality while seeking to satisfy the risk parity condition. Initially departing from the lower bound, the portfolio variance increases as it seeks to find a rank-1 solution where the asset risk contributions are equalized.

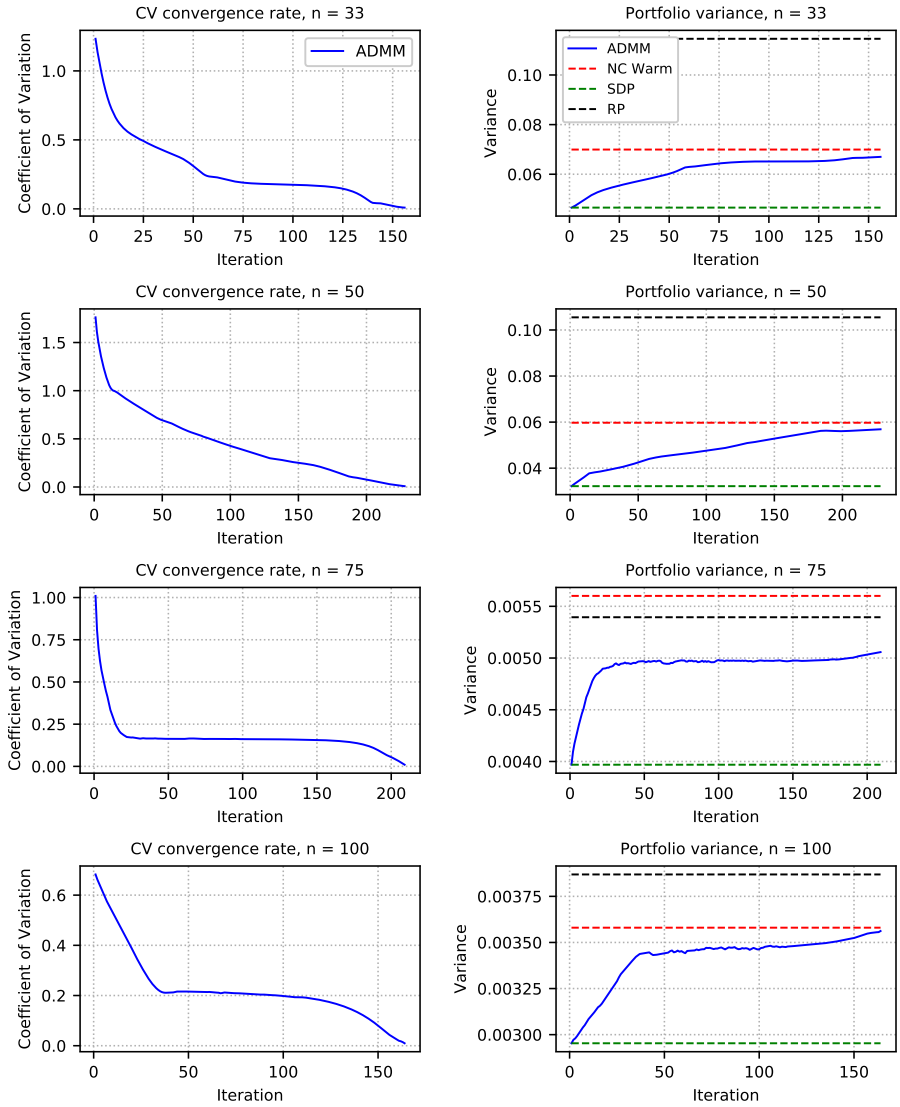

Fig. 3 Convergence plots for the lowest variance risk parity framework

6 Conclusion

This paper introduced a generalized risk parity framework that allows an investor to optimize a given risk–return profile while maintaining a desirable degree of risk-based diversification. The proposed framework also allows for short selling, thereby giving an investor increased flexibility when compared to traditional risk parity frameworks. This combines the benefits of typical portfolio optimization strategies, such as MVO, while relaxing the tight bounds on risk dispersion imposed by risk parity. However, imposing the constraints necessary to limit the risk dispersion means this is a non-convex optimization problem.

We propose an ADMM algorithm designed to handle the non-convexity of our generalized risk parity framework. We begin by relaxing the non-convex into a SDP. The algorithm operates by partitioning our original problem into two subproblems: a convex SDP and a non-convex rank-constrained problem. The structure of the latter problem allows us to find an analytic solution, thereby avoiding both the theoretical and numerical drawbacks associated with non-convexity. The algorithm proceeds to solve the problem sequentially through multiple iterations, with each iteration tightening the SDP relaxation towards a rank-1 solution. A rank-1 solution is no longer considered a relaxation, providing an optimal solution that complies with the risk dispersion constraints.

In addition, we proposed two extensions to our framework. First, we introduced an ellipsoidal uncertainty structure around the portfolio expected returns to formulate a robust generalized risk parity framework. Second, we addressed the one specific scenario of risk parity, where we seek the risk parity solution with the lowest variance. We note that our proposed framework is able to accommodate additional convex constraints, provided these can be reformulated into a SDP.

The numerical experiments show the proposed ADMM algorithm is able to deliver a higher quality optimal solution when compared to the original non-convex problem, even after we warm-start the non-convex problem. Although the algorithm does require additional runtime, it is still able to converge within reasonable time. Since this algorithm is a heuristic, we are unable to make any claims regarding global optimality. Nevertheless, the design of the algorithm guarantees we start from a theoretical lower bound, slowly climbing upwards towards the nearest feasible solution. This, in turn, suggests we should find a high quality, if not global, optimal solution.

The proposed generalized risk parity framework and the implementation of the ADMM algorithm provide an investor with greater flexibility when considering portfolio optimization tools. Indeed, we believe this framework combines the most appealing aspects of both MVO and risk parity. Moreover, the ADMM algorithm, while unable to guarantee global optimality, is still able to deliver a high quality optimal solution. Finally, the selection of the ADMM parameters and, more generally, the convergence rate of our proposed ADMM algorithm are still the subject of ongoing research.

References

1. Bai, X., Scheinberg, K., Tütüncü, R.H.: Least-squares approach to risk parity in portfolio selection. Quantitative Finance **16**(3), 357–376 (2016)

2. Best, M.J., Grauer, R.R.: On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. The review of financial studies **4**(2), 315–342 (1991)
3. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning **3**(1), 1–122 (2011)
4. Broadie, M.: Computing efficient frontiers using estimated parameters. Annals of Operations Research **45**(1), 21–58 (1993)
5. Bruder, B., Roncalli, T., et al.: Managing risk exposures using the risk budgeting approach. Tech. rep., University Library of Munich, Germany (2012)
6. Chopra, V.K., Ziemba, W.T.: The effect of errors in means, variances, and covariances on optimal portfolio choice. Journal of Portfolio Management pp. 6–11 (1993)
7. Delage, E., Ye, Y.: Distributionally robust optimization under moment uncertainty with application to data-driven problems. Operations Research **58**(3), 595–612 (2010)
8. Dunning, I., Huchette, J., Lubin, M.: Jump: A modeling language for mathematical optimization. Society for Industrial and Applied Mathematics **59**(2), 295–320 (2017). DOI 10.1137/15M1020575
9. Fabozzi, F.J., Kolm, P.N., Pachamanova, D.A., Focardi, S.M.: Robust portfolio optimization. Journal of Portfolio Management **33**(3), 40 (2007)
10. Fama, E.F., French, K.R.: Common risk factors in the returns on stocks and bonds. Journal of Financial Economics **33**(1), 3–56 (1993)
11. Feng, Y., Palomar, D.P.: Scrip: Successive convex optimization methods for risk parity portfolio design. IEEE Transactions on Signal Processing **63**(19), 5285–5300 (2015)
12. Gabay, D., Mercier, B.: A dual algorithm for the solution of non linear variational problems via finite element approximation. Institut de recherche d'informatique et d'automatique (1975)
13. Ghadimi, E., Teixeira, A., Shames, I., Johansson, M.: Optimal parameter selection for the alternating direction method of multipliers (admm): quadratic problems. IEEE Transactions on Automatic Control **60**(3), 644–658 (2015)
14. Glowinski, R., Marroco, A.: Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique **9**(R2), 41–76 (1975)
15. Goldfarb, D., Iyengar, G.: Robust portfolio selection problems. Mathematics of Operations Research **28**(1), 1–38 (2003)
16. Goldstein, T., O'Donoghue, B., Setzer, S., Baraniuk, R.: Fast alternating direction optimization methods. SIAM Journal on Imaging Sciences **7**(3), 1588–1623 (2014)
17. Haugh, M., Iyengar, G., Song, I.: A generalized risk budgeting approach to portfolio construction. Journal of Computational Finance **21**(2), 29–60 (2017)
18. Jorion, P.: Bayes-stein estimation for portfolio analysis. Journal of Financial and Quantitative Analysis **21**(3), 279–292 (1986)
19. Kapsos, M., Christofides, N., Rustem, B.: Robust risk budgeting. Annals of Operations Research pp. 1–23 (2017)
20. Ledoit, O., Wolf, M.: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. Journal of Empirical Finance **10**(5), 603–621 (2003)
21. Lintner, J.: The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The Review of Economics and Statistics pp. 13–37 (1965)
22. Lobo, M.S., Boyd, S.: The worst-case risk of a portfolio. Technical report. Available from http://web.stanford.edu/∼boyd/papers/pdf/risk_bnd.pdf (2000)
23. Maillard, S., Roncalli, T., Teïletche, J.: The properties of equally weighted risk contribution portfolios. Journal of Portfolio Management **36**(4), 60–70 (2010)
24. Markowitz, H.: Portfolio selection. Journal of Finance **7**(1), 77–91 (1952)
25. Merton, R.C.: On estimating the expected return on the market: An exploratory investigation. Journal of financial economics **8**(4), 323–361 (1980)
26. Mossin, J.: Equilibrium in a capital asset market. Econometrica: Journal of the econometric society pp. 768–783 (1966)
27. Nesterov, Y.: Introductory lectures on convex optimization: A basic course, vol. 87. Springer Science & Business Media (2013)
28. Qian, E.: On the financial interpretation of risk contribution: risk budgets do add up. Journal of Investment Management **4**(4), 41–51 (2006)

29. Quandl.com: Wiki – various end-of-day stock prices (2017). URL https://www.quandl.com/databases/WIKIP/usage/export. [Online; accessed 07-Nov-2017]
30. Raghunathan, A.U., Di Cairano, S.: Alternating direction method of multipliers for strictly convex quadratic programs: Optimal parameter selection. In: American Control Conference (ACC), 2014, pp. 4324–4329. IEEE (2014)
31. Sharpe, W.F.: Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance **19**(3), 425–442 (1964)
32. Tütüncü, R.H., Koenig, M.: Robust asset allocation. Annals of Operations Research **132**(1-4), 157–187 (2004)
33. You, S., Peng, Q.: A non-convex alternating direction method of multipliers heuristic for optimal power flow. In: Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on, pp. 788–793. IEEE (2014)