SIAM J. FINANCIAL MATH.
Vol. 14, No. 2, pp. 525–556

Geometric Characterization of Maximum Diversification Return Portfolio via Rao's Quadratic Entropy*

Hou-Duo Qi†

Abstract. Diversification return has been well studied in finance literature, mainly focusing on the various sources from which it may be generated. The maximization of diversification return, in its natural form, is often handed over to convex quadratic optimization for its solution. In this paper, we study the maximization problem from the perspective of Rao's quadratic entropy (RQE), which is closely related to the Euclidean distance matrix and hence has deep geometric implications. This new approach reveals a fundamental feature that the maximum diversification return portfolio (MDRP) admits a spherical embedding with the hypersphere having the least volume. This important characterization extends to the maximum volatility portfolio, the long-only MDRP, and the ridge-regularized MDRP. RQE serves as a unified formulation for diversification return related portfolios and generates new portfolios that are worth further investigation. As an application of this geometric characterization, we develop a computational formula for measuring the distance between a new asset and an existing portfolio that has the hyperspherical embedding. Numerical experiments demonstrate the developed theory.

Key words. maximum diversification return, Euclidean distance matrix, Rao's quadratic entropy, ridge regularization, long-only portfolio, spherical embedding

MSC codes. 90C20, 90C90, 90B50, 91B28

DOI. 10.1137/22M1492313

1. Introduction.

Diversification Return (DR) of a portfolio, a term coined by Booth and Fama [2] and also known as the excess growth rate [20], has been extensively studied in finance literature. One of the focuses is how DR is generated. Attributes include diversification through variance reduction [10] and rebalancing [32]. More discussions and debates on those two attributes among others can be found in [14, 27, 4, 5, 25, 18]. The extensive empirical results conducted in [20] confirm that maximizing the diversification return leads to attractive alternatives to competing smart portfolios. Unlike the Markovitz mean-variance model, however, the maximum diversification return portfolio (MDRP) lacks geometric interpretation. The main purpose of this paper is to establish an elegant characterization that MDRP embeds its assets on a hypersphere centered at the origin having the minimal volume with the origin being the MDRP-weighted average of the embedding points. This characterization explains why MDRP often yields negative exposures to some assets and positive exposure to others as observed in [20]. This characterization also extends to other portfolios such as the maximum

*Received by the editors April 24, 2022; accepted for publication (in revised form) February 3, 2023; published electronically May 24, 2023.
https://doi.org/10.1137/22M1492313
Funding: This work was partially supported by the departmental project P0044200.
†Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong (houduo.qi@polyu.edu.hk).

volatility portfolio [22, 20], the long-only MDRP (ℓMDRP) [20], and the ridge-regularized MDRP. In the following, we first explain the mathematical formulation of MDRP. We then introduce our approach and describe our main results.

1.1. Maximum diversification return portfolio. Suppose there are n risky assets S_i, $i = 1, \ldots, n$. We denote by V the covariance matrix of the returns of the n assets. Let $\mathbf{w} = (w_1, \ldots, w_n)^\top$ (column vector) be a portfolio that satisfies the budget constraint $\mathbf{1}_n^\top \mathbf{w} = 1$, where $\mathbf{1}_n$ is the column vector of all ones of dimension n and $\mathbf{1}_n^\top$ is its transpose (row vector). The diversification return of the portfolio $\mathbf{w}$ is given by

$$\mathrm{DR}(\mathbf{w}) = \frac{1}{2}\left(\sum_{i=1}^n w_i \sigma_i^2 - \mathbf{w}^\top V \mathbf{w}\right), \tag{1.1}$$

where σ_i^2 is the variance of the asset S_i. There are several ways of interpreting $\mathrm{DR}(\mathbf{w})$. Booth and Fama [2] treated it as an approximation to the difference between the compound rate of return of the portfolio and the averaged compound rate of returns of individual assets. Willenbrock [32] argued that the geometric rate of return is more suitable. Maeso and Martellini [20] derived it as the excess growth rate of a stochastic portfolio.

The MDRP is defined as

$$\mathbf{w}^* := \arg\max\ \mathrm{DR}(\mathbf{w}), \qquad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1, \tag{1.2}$$

where ":=" means "define." Let $\eta := (\sigma_1^2, \ldots, \sigma_n^2)^\top$ be the variance vector. The problem has a closed-form solution if V is nonsingular. If η is proportional to the expected return vector $\mathbf{r}$, i.e., $\eta = k\mathbf{r}$ for some constant $k > 0$, then MDRP (1.2) is obviously a special case of the Markovitz mean-variance model and hence it is on the efficient frontier. Otherwise, MDRP is not efficient in the sense of Markovitz. However, the extensive numerical results in [20] show that MDRP in many scenarios outperforms several popular smart strategies. Motivated by the Markovitz model, the first question we would like to ask is what geometric optimality property MDRP may enjoy. The answer lies with casting (1.2) as Rao's quadratic entropy maximization problem, which is introduced in the following subsection.

The second question we would like to address is under what circumstances MDRP is a long-only portfolio, i.e., $w_i^* \geq 0$ for all $i = 1, \ldots, n$. In [20], the long-only constraint $\mathbf{w} \geq 0$ was added to (1.2) to get the (ℓMDRP):

$$\mathbf{w}_L^* := \arg\max\ \mathrm{DR}(\mathbf{w}), \qquad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1, \ \mathbf{w} \geq 0. \tag{1.3}$$

This problem does not have a closed-form solution anymore. In [22], it regularizes the positive part of $\mathbf{w}^*$ to get a long-only portfolio by

$$\mathbf{w}_+^* := \frac{1}{\sum_{i=1}^n \max(0, w_i^*)} \max(0, \mathbf{w}^*),$$

where $\max(0, \mathbf{w}^*)$ is the positive part of $\mathbf{w}^*$. One common drawback among the two strategies is that the generated portfolio has relatively a small number of active assets (portfolio concentration issue). We propose a new portfolio that is based on the ridge regularization:

$$\mathbf{w}_\rho^* = \arg\max\ \mathrm{DR}(\mathbf{w}) - \frac{\rho}{2}\|\mathbf{w}\|^2, \quad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1, \tag{1.4}$$

where $\rho \geq 0$ is a regularization parameter and $\|\mathbf{w}\|$ is the Euclidean norm. It is known [8] that the ridge regularization is equivalent to the norm $\|\mathbf{w}\| \leq \tau$ with properly defined τ. When $\tau = 1/n$, the normal constraint yields the equal-weight portfolio: $w_i = 1/n$, $i = 1, \ldots, n$. It is also known that norm-regularization (e.g., ℓ_1 and ℓ_2 norm) also leads to robust and structural portfolios (see, e.g., [17, 33]). We will see the geometric characterization for MDRP also extends to the problem (1.4). Moreover, we will establish a computable lower bound ρ_0 such that whenever $\rho \geq \rho_0$, we always have $\mathbf{w}_\rho^* > 0$. This result gives us the flexibility to control the number of active long-only assets in the resulting portfolio.

Our third question is to address a practical scenario when a new risky asset S_{n+1} is available and we try to measure its distance to the MDRP of the existing n risky assets. It turns out that this question is closely related to the problem of adding a new point to an existing vector diagram [11] and the landmark multidimensional scaling problem [7]. We will develop a computational formula for measuring the distance. In the following, we explain how we will resolve those questions.

1.2. Rao's quadratic entropy and main results. Our departing point from existing research is to cast $\mathrm{DR}(\mathbf{w})$ as an instance of Rao's quadratic entropy (RQE) [29, 28], which was initially developed for measuring bio-diversities; see [30] for its comparison to other diversity measures including Shannon entropy. We briefly describe it in terms of the n risky assets S_i, $i = 1, \ldots, n$. Let D_{ij} be a dissimilarity measure between S_i and S_j and $D_{ij} = D_{ji}$. For any long-only portfolio $\mathbf{w}$, RQE is defined as

$$q_D(\mathbf{w}) := \frac{1}{2}\mathbf{w}^\top D \mathbf{w} = \frac{1}{2}\sum_{i,j} w_i w_j D_{ij} \qquad \text{with} \quad \mathbf{w} \geq 0, \quad \mathbf{1}_n^\top \mathbf{w} = 1,$$

where $D := (D_{ij})_{i,j=1}^n$ is the dissimilarity matrix. The quantity $q_D(\mathbf{w})$, also known as Rao's diversity measure, summarizes the total dissimilarity among the n risky assets. A fundamental requirement for being a legitimate RQE is that it preserves the following principle: the diversity of a simple mixing of two portfolios should be higher than the simple mix of individual diversities:

$$\left\{ \begin{array}{l} q_D((1-\lambda)\mathbf{w}_1 + \lambda \mathbf{w}_2) \geq (1-\lambda) q_D(\mathbf{w}_1) + (1-\lambda) q_D(\mathbf{w}_2), \\ \forall\ \mathbf{w}_i,\ \mathbf{1}_n^\top \mathbf{w}_i = 1, \quad \mathbf{w}_i \geq 0, \quad i = 1, 2, \quad \text{and} \quad 0 \leq \lambda \leq 1. \end{array} \right. \tag{1.5}$$

We refer to the important reference [3] for more applications of RQE to portfolio constructions. A key observation is

$$\mathrm{DR}(\mathbf{w}) = q_D(\mathbf{w}) \quad \text{for} \quad D = D_V := \frac{1}{2}(\mathbf{1}_n \eta^\top + \eta \mathbf{1}_n^\top) - V \tag{1.6}$$

by using the fact $\mathbf{1}_n^\top \mathbf{w} = 1$. We will show in Lemma 2.3 that D_V satisfies the criterion (1.5). (when no confusion is caused, we drop its dependence on V). Therefore, maximizing $\mathrm{DR}(\mathbf{w})$ in (1.2) is equivalent to

$$\max\ q_D(\mathbf{w}) \quad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1. \tag{1.7}$$

The reformulation (1.7) is significant in several aspects.

(i) First, it is easy to see $D_{ij} = \frac{1}{2}(\sigma_i^2 + \sigma_j^2) - \sigma_{ij} \geq 0$. This immediately implies that $q_D(\mathbf{w}) \geq 0$ for any long-only portfolio $\mathbf{w} \geq 0$, recovering the result proved in [20, App. A]. Being nonnegative and assuming it is irreducible, D has a positive eigenvector $\mathbf{p}$ corresponding to its largest eigenvalue and this vector is known as the Perron–Frobenius (PF) eigenvector [24]. Hence, it naturally leads to a long-only portfolio (PF portfolio). Moreover, if $\mathbf{1}_n$ is a PF eigenvector of D, we prove in Corollary 3.4 that the equal-weight portfolio is MDRP. This happens when D corresponds to regular figures in graph theory [15].

(ii) Second, D is in fact a Euclidean distance matrix (EDM). That is, there exist a set of points $\mathbf{x}_i \in \Re^k$, $i = 1, \ldots, n$, such that $\|\mathbf{x}_i - \mathbf{x}_j\|^2 = D_{ij}$, where k is the embedding dimension of D. By applying Gower's fundamental result on EDM [13, Thm. 3], we show that MDRP corresponds to a set of embedding points $\{\mathbf{x}_i\}$ on a hypersphere with the following properties (see Theorem 3.1):

   (a) The center of the hypersphere is the origin.

   (b) The origin is the MDRP-weighted average of the embedding points:

$$\sum_{i=1}^{n} w_i^* \mathbf{x}_i = 0. \tag{1.8}$$

   (c) The hypersphere has the smallest radius among all possible spherical embeddings.

   In recognition of Gower's contribution, we denote the radius by R_G and the hypersphere by Gower Sphere(R_G). Since k is the smallest embedding dimension, the Gower sphere has the minimal volume. Moreover, it is shown in Corollary 3.2 that any feasible portfolio $\mathbf{w}$ corresponds to a spherical embedding with properties that are different from those in (a)–(c) above. Example 3.1 shows a scenario where the equal-weight portfolio leads to a spherical embedding of the same volume as of the Gower sphere, but it is not MDRP. The equal-weight portfolio does not satisfy the property (b) in (1.8). In other words, out of infinitely many spherical embeddings, MDRP looks for the one that satisfies the properties (a)–(c). When the origin is within the convex hull of its embedding points, MDRP is in fact a long-only portfolio. However, it is difficult to know when this would happen as we do not know those $\mathbf{x}_i$ a priori.

(iii) The power of the RQE reformulation is further reflected by the fact that the three MDRP related portfolios (the maximum volatility portfolio (maxVP), the ℓMDRP (1.3), and the regularized MDRP (1.4)) each has a RQE representation via a properly defined Euclidean distance matrix. Hence, those portfolios all have a spherical embedding that achieve a minimum volume among all possible embeddings. In particular, for the latter, we are able to prove in Theorem 3.7 that

$$\mathbf{w}_\rho^* \geq 0 \quad \text{when} \quad \rho \geq \left(1 + (n-1)\sqrt{n}\right) \max_i \sum_{j=1}^{n} D_{ij}.$$

   In fact, the ridge regularization acts like a shrinkage operator that pulls the weights toward the equal-weight portfolio. Extension to maxVP and ℓMDRP will be, respectively, discussed in subsections 3.2 and 3.4.

Last, the minimal spherical characterization allows us to project a newly available asset to Sphere(R_G) and we can use the error to measure the distance between the new asset and the MDRP portfolio. A fast computational formula is developed in Proposition 4.1. Therefore, this paper provides a complete mathematical study of MDRP and three of its closely related portfolios and lays a foundation for future research on other diversification maximization problems.

1.3. Organization.

The paper is organized as follows. We review some background on the Euclidean distance matrix in section 2. The geometric characterization is done in section 3 with its implications to the maximum volatility portfolio. Both the regularized problem (1.4) and the ℓMDRP (1.3) are, respectively, studied in subsections 3.3 and 3.4. As an application, section 4 addresses the issue of measuring the distance between a newly available asset and MDRP. In section 5, we conduct some preliminary numerical experiments to verify the theoretical results. We conclude the paper in section 6.

Notation. A (column) vector is often denoted by a boldfaced letter such as $\mathbf{s}$ with s_i being its elements. Diag$(\mathbf{s})$ is the diagonal matrix whose diagonals are given by $\mathbf{s}$ and diag(A) is the diagonal vector of a squared matrix A. We let I_n be the identity matrix of size n, and it is often abbreviated as I when the dimension n is obvious. For a set of risky assets S_i, $i=1,\ldots,n$, σ_i^2 is the variance of the return of asset S_i, and ρ_{ij} is the correlation between S_i and S_j. Their covariance is denoted as σ_{ij}, which also satisfies $\sigma_{ij}=\rho_{ij}\sigma_i\sigma_j$. The vector η is reserved for the variance vector with its ith component being $\eta_i=\sigma_i^2$. For a set of points $\mathbf{x}_i \in \Re^k$, $i=1,\ldots,n$, and a portfolio $\mathbf{w}\in\Re^n$ satisfying $\mathbf{1}_n^\top \mathbf{w}=1$, the $\mathbf{w}$-weighted average is $\sum_{i=1}^n w_i \mathbf{x}_i$. The vector $\mathbf{w}_{1/n}$ denotes the equally weighted portfolio with each weight being $1/n$. For a given square matrix D, D^- denotes a generalized inverse that satisfies the property $DD^-D=D$ and $D^-DD^-=D^-$. For a square matrix A, $A\succeq 0$ means A is positive semidefinite and $A\succeq B$ means $(A-B)\succeq 0$. Furthermore, $A\preceq 0$ means $(-A)\succeq 0$.

2. Background on Euclidean distance matrix.

We recall an $n\times n$ matrix $D=(D_{ij})_{i,j=1}^n$ is called EDM if there exists a set of points $\mathbf{x}_i\in\Re^k$, $i=1,\ldots,n$ such that the squared Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ reproduces D_{ij}, i.e., $\|\mathbf{x}_i-\mathbf{x}_j\|^2=D_{ij}$. Note that there exist many such embedding points. The smallest dimension k in which those embedding points live is called the embedding dimension of D. There are many useful references about EDM (see, e.g., [6, 26, 9, 1]). We only describe some key properties that are to be used in this paper.

A first major characterization of EDM D was due to Schoenberg [31]:

$$\text{(2.1)} \qquad \mathrm{diag}(D)=0 \qquad \text{and} \qquad \mathbf{h}^\top D\mathbf{h}\le 0 \ \ \forall\ \mathbf{h}\in\Re^n \text{ satisfying } \mathbf{1}_n^\top\mathbf{h}=0.$$

This characterization is equivalent to

$$\text{(2.2)} \qquad \mathrm{diag}(D)=0 \qquad \text{and} \qquad JDJ\preceq 0 \ \text{ with } \ J:=I_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top.$$

We note that the matrix J is the orthogonal projection operator to the subspace $\mathbf{1}_n^\perp$, which contains all the vectors orthogonal to $\mathbf{1}_n$. Gower generalized the characterization (2.2) to a class of projection operators.

Lemma 2.1 (see [12, Thm. 2]). *D is EDM if and only if* $\mathrm{diag}(D)=0$ *and* $J_{\mathbf{s}}^\top D J_{\mathbf{s}} \preceq 0$ *for any* $\mathbf{s}\in\Re^n$ *satisfying* $\mathbf{1}_n^\top \mathbf{s}=1$. *Here* $J_{\mathbf{s}} := I_n - \mathbf{s}\mathbf{1}_n^\top$.

The original result of Gower has an extra condition on $\mathbf{s}$: $D\mathbf{s}\neq 0$. This condition was later proved superfluous [13, p. 83]. Given EDM D, a set of embedding points $\mathbf{x}_i$ can be generated as follows. Suppose $\mathbf{s}\in\Re^n$ satisfies $\mathbf{1}_n^\top\mathbf{s}=1$. Since the matrix $(-J_{\mathbf{s}}^\top D J_{\mathbf{s}})$ is positive semidefinite, it admits the following decomposition:

$$B := -\frac{1}{2} J_{\mathbf{s}}^\top D J_{\mathbf{s}} = X^\top X \qquad \text{with} \quad X := [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n], \tag{2.3}$$

where $\mathbf{x}_i\in\Re^k$ are embedding points of D, i.e., $\|\mathbf{x}_i-\mathbf{x}_j\|^2 = D_{ij}$, $i,j=1,\ldots,n$, k, is the embedding dimension, and $k=\mathrm{rank}(J_{\mathbf{s}}^\top D J_{\mathbf{s}})$. We note that with a different choice of $\mathbf{s}$, the decomposition would generate different sets of embedding points.

We say EDM D is spherical if there exists a set of embedding points that lie on a hypersphere. We note that the center of the hypersphere is not necessarily at zero (the origin of the coordinate system). For example, [1, Thm. 4.1] gave a constructive way to compute the center of such a hypersphere for a set of embedding points corresponding to $\mathbf{s}=(1/n)\mathbf{1}_n$. We will use Gower's construction.

Lemma 2.2. *Let D be nonzero EDM and D^- be any generalized inverse of D (i.e., $DD^-D = D$ and $D^-DD^- = D^-$). We have the following results.*

(i) [13, Thm. 2] *It holds that* $\mathbf{1}_n^\top D^- D = \mathbf{1}_n^\top$ *and* $DD^-\mathbf{1}_n = \mathbf{1}_n$. *Moreover, the quantity* $\mathbf{1}_n^\top D^-\mathbf{1}_n$ *is the same for any choice of* D^-. *In other words, it does not depend on the choice of* D^-.

(ii) [13, Thm. 3] *D is spherical if and only if* $\mathbf{1}_n^\top D^-\mathbf{1}_n\neq 0$. *In this case, there exists a hypersphere centered at the origin with the radius R_G given by*

$$R_G^2 = \frac{1}{2(\mathbf{1}_n^\top D^- \mathbf{1}_n)}.$$

A set of embedding points sitting on the sphere are generated by (2.3) *corresponding to* $\mathbf{s}=D^-\mathbf{1}_n/\mathbf{1}_n^T D^-\mathbf{1}_n$.

(iii) [13, Thm. 6] *Let* $\mathrm{rank}(D)=r$. *Then the embedding dimension* $k=r-1$ *if and only if* $\mathbf{1}_n^\top D^-\mathbf{1}_n\neq 0$ *and* $k=r-2$ *if and only if* $\mathbf{1}_n^\top D^-\mathbf{1}_n=0$.

We note that Gower [13] called $(-D/2)$ EDM. Therefore, Lemma 2.2 is a restatement of Gower's result. We now prove that the RQE dissimilarity matrix is actually EDM.

Lemma 2.3. *The following hold.*

(i) *A matrix D is EDM if and only if* $\mathrm{diag}(D)=0$ *and the RQE condition* (1.5) *holds.*

(ii) *Let $D=D_V$ in* (1.6). *Then D is EDM. Furthermore, D is nonsingular if V is nonsingular.*

Proof. (i) Simple calculation verifies that the RQE condition (1.5) is actually equivalent to

$$(\mathbf{w}_2-\mathbf{w}_1)^\top D(\mathbf{w}_2-\mathbf{w}_1)\le 0 \quad \forall\ \mathbf{w}_i\ge 0,\ \ \mathbf{1}_n^\top\mathbf{w}_i=1,\ i=1,2. \tag{2.4}$$

We use (2.1) for EDM characterization. Now suppose the condition (2.1) holds. Choosing $\mathbf{h} = \mathbf{w}_2 - \mathbf{w}_1$ leads to the condition (2.4). Hence, when D is EDM, it must be a RQE dissimilarity matrix. We now prove the converse part.

Suppose the condition (1.5) holds. Let $0 \neq \mathbf{h}$ such that $\mathbf{1}_n^\top \mathbf{h} = 0$. Define $\mathbf{h}_+ := \max\{\mathbf{h}, 0\}$ and $\mathbf{h}_- := \max\{-\mathbf{h}, 0\}$. We then have

$$\mathbf{h} = \mathbf{h}_+ - \mathbf{h}_- \quad \text{and} \quad \mathbf{1}_n^\top \mathbf{h}_+ = \mathbf{1}_n^\top \mathbf{h}_- =: c > 0.$$

Define $\mathbf{w}_1 := \mathbf{h}_+/c$ and $\mathbf{w}_2 := \mathbf{h}_-/c$. It follows from (2.4) that

$$(\mathbf{w}_2 - \mathbf{w}_1)^\top D(\mathbf{w}_2 - \mathbf{w}_1) = \frac{1}{c^2}(\mathbf{h}_+ - \mathbf{h}_1)^\top D(\mathbf{h}_+ - \mathbf{h}_-) = \frac{1}{c^2}\mathbf{h}^\top D\mathbf{h} \leq 0,$$

establishing the condition (2.1).

(ii) An immediate consequence of the characterization (2.2) is that the matrix D_V in (1.6) is EDM because

$$\operatorname{diag}(D_V) = 0 \qquad \text{and} \qquad JD_VJ = \underbrace{J\left(\mathbf{1}_n\eta^\top + \eta\mathbf{1}_n^\top\right)J}_{=0} - JVJ = -JVJ \preceq 0,$$

where we used the fact $J\mathbf{1}_n = 0$. If V is nonsingular, then the embedding dimension k is given by

$$k = \operatorname{rank}(JD_VJ) = \operatorname{rank}(JVJ) = n-1,$$

because $\operatorname{rank}(J) = n-1$. Therefore, Lemma 2.2(iii) implies that $\operatorname{rank}(D_V)$ is either $(k+1) = n$ or $(k+2) = n+1$ (impossible). Hence, $\operatorname{rank}(D_V) = n$. That is, D_V is nonsingular if V is so. ■

It is possible that D_V is nonsingular even if V is singular.

3. Characterization of minimal spherical embedding.

In this section, we study the geometric characterization of MDRP discussed in the introduction and extend the characterization to the maximum volatility portfolio, the ridge-regularized MDRP, and the ℓMDRP. The study makes heavy use of EDM properties.

Throughout, we assume that MDRP $\mathbf{w}^*$ in (1.2) exists and $D \neq 0$. This assumption is satisfied if the covariance matrix V is nonsingular. In particular, when $D = 0$, V must be a rank-2 matrix:

$$V = \frac{1}{2}\left(\mathbf{1}_n^\top\eta + \eta\mathbf{1}_n^\top\right).$$

In other words, the covariance σ_{ij} satisfies

$$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j = (\sigma_i^2 + \sigma_j^2)/2.$$

Using the inequality $(\sigma_i^2 + \sigma_j^2) \geq 2\sigma_i\sigma_j$, the above equation means $\rho_{ij} = 1$ and $\sigma_i = \sigma_j$ for all i, j. It is equivalent to say that the n risky assets are perfectly correlated and share the same variance. They can be regarded as n copies of the same asset. Our assumption on $D \neq 0$ removes this possibility from our study.

It is also worth pointing out that when V is nonsingular, the solution of MDRP (1.2) is given as follows in terms of the inverse of V:

$$
\begin{aligned}
\mathbf{w}^* &= \left(1 - \frac{\mathbf{1}_n^\top V^{-1}\eta}{2}\right)\frac{V^{-1}\mathbf{1}_n}{\mathbf{1}_n^\top V^{-1}\mathbf{1}_n} + \frac{1}{2}V^{-1}\eta \\
&= \left(1 - \frac{\mathbf{1}_n^\top V^{-1}\eta}{2}\right)\underbrace{\frac{V^{-1}\mathbf{1}_n}{\mathbf{1}_n^\top V^{-1}\mathbf{1}_n}}_{=:\mathbf{w}_{\mathrm{mvp}}} + \frac{\mathbf{1}_n^\top V^{-1}\eta}{2} \times \underbrace{\frac{V^{-1}\eta}{\mathbf{1}_n^\top V^{-1}\eta}}_{=:\mathbf{w}_{\mathrm{maxvp}}},
\end{aligned} \tag{3.1}
$$

where $\mathbf{w}_{\mathrm{mvp}}$ is the classical minimum variance portfolio and $\mathbf{w}_{\mathrm{maxvp}}$ is the maximum volatility portfolio. Hence, MDRP is an affine combination of MVP and maxVP. This formula has an explicit assumption that the quantity $\mathbf{1}_n^\top V^{-1}\eta \neq 0$. However, the quantity may be zero even V is nonsingular; see an example in subsection 3.2.

3.1. Maximum diversification return portfolio. In this part, we first present our main result, Theorem 3.1, which states that the MDRP $\mathbf{w}^*$ corresponds to the Gower sphere. We then construct an example to show that other portfolios may also have a spherical embedding with the same volume as the Gower sphere, but the weighted average of the embedding point does not lie at the center of the sphere (see Example 3.1). We further study a class of risky assets, whose MDRP is the equal-weight portfolio (see Corollary 3.4).

Theorem 3.1. *Let V be the covariance matrix of n risky assets and $\eta = \mathrm{diag}(V)$. Let the dissimilarity matrix D be defined by $D = D_V = \frac{1}{2}(\mathbf{1}_n\eta^\top + \eta\mathbf{1}_n^\top) - V$. Consider the MDRP $\mathbf{w}^*$ in (1.2). The following hold.*

(i) *D is EDM and $\mathbf{1}_n^\top D^- \mathbf{1}_n > 0$ for any generalized inverse D^-.*

(ii) *The n risky assets can be represented on the Gower sphere centered at the origin with the radius R_G given by $R_G^2 = 1/(2\mathbf{1}_n^\top D^-\mathbf{1}_n)$. The corresponding n embedding points on the sphere are generated by (2.3) with $\mathbf{s}$ being the maximum diversification portfolio $\mathbf{w}^* = \frac{D^-\mathbf{1}_n}{\mathbf{1}_n^\top D^-\mathbf{1}_n}$. Furthermore, the maximum diversification return is $\mathrm{DR}(\mathbf{w}^*) = R_G^2$.*

(iii) *Suppose D can be embedded on another sphere with radius R and its center is allowed to be arbitrarily chosen. We must have $R \geq R_G$.*

Proof. (i) That D is EDM has been proved in Lemma 2.3(ii). The characterization (1.5) of D implies that $q_D(\mathbf{w})$ is concave over the feasible region of (1.2). Therefore, the optimality condition holds at $\mathbf{w}^*$:

$$
D\mathbf{w}^* = \lambda\mathbf{1}_n \quad \text{and} \quad \mathbf{1}_n^\top \mathbf{w}^* = 1 \tag{3.2}
$$

for some $\lambda \in \Re$. Multiplying both sides of (3.2) by $\mathbf{w}^*$ yields

$$
2\mathrm{DR}(\mathbf{w}^*) = 2q_D(\mathbf{w}^*) = (\mathbf{w}^*)^\top D\mathbf{w}^* = \lambda\mathbf{1}_n^\top\mathbf{w}^* = \lambda.
$$

Since $D \neq 0$, there exists $D_{ij} > 0$ for some (i,j). Let $w_i = w_j = 1/2$. Then $DR(\mathbf{w}^*) \geq 2w_iw_jD_{ij} > 0$, implying $\lambda > 0$. Premultiplying $\mathbf{1}_n^T D^-$ on both sides of the first equation in (3.2) and using the fact $\mathbf{1}_n^\top D^- D = \mathbf{1}_n^\top$ in Lemma 2.2(i) leads to

$$
\lambda\mathbf{1}_n^\top D^-\mathbf{1}_n = \mathbf{1}_n^\top D^- D\mathbf{w}^* = \mathbf{1}^\top\mathbf{w}^* = 1.
$$

The fact that $\lambda > 0$ implies $\mathbf{1}_n^\top D^- \mathbf{1}_n > 0$.

Since $\mathbf{1}_n^\top D^- \mathbf{1}_n > 0$, (ii) follows from Lemma 2.2(ii). In particular, [13, Thm. 3] proved that $\mathbf{s} = \frac{D^- \mathbf{1}_n}{\mathbf{1}_n^\top D^- \mathbf{1}_n}$ satisfies the optimality condition (3.2). Hence $\mathbf{w}^* = \mathbf{s}$ and

$$\mathrm{DR}(\mathbf{w}^*) = q_D(\mathbf{w}^*) = \frac{1}{2} \frac{\mathbf{1}_n^\top D^- D D^- \mathbf{1}_n}{(\mathbf{1}_n^\top D^- \mathbf{1}_n)^2} = \frac{1}{2} \frac{1}{\mathbf{1}_n^\top D^- \mathbf{1}_n} = R_G^2.$$

(iii) Assume that D can be embedded on a sphere with radius $R > 0$ but with an unknown center. We treat this unknown center as the $(n+1)$th point. The n embedding points will have the same distance R from the $(n+1)$th point. In other words, the matrix $\overline{D}$ below is Euclidean:

$$\overline{D} := \begin{bmatrix} D & R^2 \mathbf{1}_n \\ R^2 \mathbf{1}_n^T & 0 \end{bmatrix}. \tag{3.3}$$

By Lemma 2.1, $J_{\mathbf{s}}^\top \overline{D} J_{\mathbf{s}} \preceq 0$ with $\mathbf{s} = \mathbf{e}_{n+1}$ being the vector of zeros except the last element being 1 (i.e., $\mathbf{e}_{n+1}$ is the $(n+1)$th unit vector in $\Re^{n+1}$). We compute

$$\begin{aligned} 0 \succeq J_{\mathbf{e}_{n+1}}^\top \overline{D} J_{\mathbf{e}_{n+1}} &= \left(I_{n+1} - \mathbf{1}_{n+1} \mathbf{e}_{n+1}^T \right) \overline{D} \left(I_{n+1} - \mathbf{e}_{n+1} \mathbf{1}_{n+1}^T \right) \\ &= \begin{bmatrix} D & R^2 \mathbf{1}_n \\ R^2 \mathbf{1}_n^T & 0 \end{bmatrix} - \begin{bmatrix} 2R^2 \mathbf{1}_n \mathbf{1}_n^T & R^2 \mathbf{1}_n \\ R^2 \mathbf{1}_n^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} D - 2R^2 \mathbf{1}_n \mathbf{1}_n^T & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

This is equivalent to $D \preceq 2R^2 \mathbf{1}_n \mathbf{1}_n^T$. Consequently, we have

$$\begin{aligned} 2R^2 = 2R^2 (\mathbf{w}^*)^\top \mathbf{1}_n \mathbf{1}_n^\top \mathbf{w}^* &\geq (\mathbf{w}^*)^\top D \mathbf{w}^* \\ &= \frac{1}{(\mathbf{1}_n^\top D^- \mathbf{1}_n)^2} \mathbf{1}_n^\top D^- D D^- \mathbf{1}_n = \frac{1}{\mathbf{1}_n^\top D^- \mathbf{1}_n}, \end{aligned}$$

which implies

$$R^2 \geq \frac{1}{2 \mathbf{1}_n^\top D^- \mathbf{1}_n} = R_G^2.$$

Hence, R_G is the smallest radius of all spheres that can contain embedding points of D. ■

Corollary 3.2.

(i) *The n risky assets can be embedded on a sphere of any radius $R \geq R_G$.*

(ii) *For any portfolio $\mathbf{w}$ satisfying $\mathbf{1}_n^\top \mathbf{w} = 1$, the n risky assets can be embedded on a sphere of the radius R_G such that the $\mathbf{w}$-weighted average of the embedding points is the origin. In this case, the origin may not be the center of the sphere.*

Proof. (i) Let $R \geq R_G$ be given. Theorem 3.1 proved that the n risky assets can be embedded on the Gower sphere. Therefore, the matrix

$$D_g := \begin{bmatrix} D & R_G^2 \mathbf{1}_n \\ R_G^2 \mathbf{1}_n^T & 0 \end{bmatrix}$$

is Euclidean. For (i), it is enough to prove that the matrix $\overline{D}$ in (3.3) is EDM. We note that $\overline{D}$ has the following decomposition:

$$\overline{D} = D_g + \frac{R^2 - R_g^2}{2}\left(\mathbf{e}_{n+1}\mathbf{1}_{n+1}^\top + \mathbf{1}_{n+1}\mathbf{e}_{n+1}^\top\right) - (R^2 - R_g^2)\mathbf{e}_{n+1}\mathbf{e}_{n+1}^\top.$$

Choosing $\mathbf{s} = \mathbf{e}_{n+1}$, it is easy to verify that

$$J_\mathbf{s}^\top \overline{D} J_\mathbf{s} = \underbrace{J_\mathbf{s}^\top D_g J_\mathbf{s}}_{\preceq 0} - (R^2 - R_g^2)\underbrace{J_\mathbf{s}^\top \mathbf{e}_{n+1}\mathbf{e}_{n+1}^\top J_\mathbf{s}}_{\succeq 0} \preceq 0,$$

where we used the fact $J_\mathbf{s}^\top \mathbf{1}_{n+1} = 0$ and D_g is Euclidean. Therefore, $\overline{D}$ is Euclidean by Lemma 2.1. Consequently, the n risky assets can be embedded on a sphere with any radius $R \ge R_G$.

(ii) We give a constructive proof. Suppose $\mathbf{x}_i$, $i = 1\ldots,n$, are the embedding points of the n risky assets on the Gower sphere. Hence, we have $\|\mathbf{x}_i\| = R_G$. For a given portfolio $\mathbf{w}$, let $\bar{\mathbf{x}}_\mathbf{w}$ be the average of the embedding points by the portfolio:

$$\bar{\mathbf{x}}_\mathbf{w} := \sum_{i=1}^n w_i \mathbf{x}_i.$$

We define $\mathbf{y}_i := \mathbf{x}_i - \bar{\mathbf{x}}_\mathbf{w}$, $i = 1,\ldots,n$. It follows that

$$\|\mathbf{y}_i - (-\bar{\mathbf{x}}_\mathbf{w})\| = \|\mathbf{y}_i + \bar{\mathbf{x}}_\mathbf{w}\| = \|\mathbf{x}_i\| = R_G, \qquad i = 1,\ldots,n.$$

Therefore the points $\mathbf{y}_i$ lie on the sphere of radius R_G with $(-\bar{\mathbf{x}}_\mathbf{w})$ being its center. The $\mathbf{w}$-average of the new embedding points $\{\mathbf{y}_i\}$ is

$$\bar{\mathbf{y}}_\mathbf{w} := \sum_{i=1}^n w_i \mathbf{y}_i = \sum_{i=1}^n w_i \mathbf{x}_i - \Big(\sum_{i=1}^n w_i\Big)\bar{\mathbf{x}}_\mathbf{w} = \bar{\mathbf{x}}_\mathbf{w} - \bar{\mathbf{x}}_\mathbf{w} = 0.$$

It is important to note that $\bar{\mathbf{x}}_\mathbf{w} \neq \bar{\mathbf{y}}_\mathbf{w}$ in general. This proves the result. ∎

Motivated by Corollary 3.2(ii), we propose a concept of portfolio centrality.

Definition 3.3 (centrality of portfolio). *Suppose there are n risky assets with V being their covariance matrix. Let $\mathbf{x}_i$, $i = 1,\ldots,n$, be the embedding points on the Gower sphere determined by the distance matrix D_V. The centrality of portfolio $\mathbf{w}$ is defined as*

$$c(\mathbf{w}) := \|w_1\mathbf{x}_1 + w_2\mathbf{x}_2 + \cdots + w_n\mathbf{x}_n\|.$$

The embedding points $\{\mathbf{x}_i\}$ can be obtained via (2.3) with D replaced by D_V. It follows that

$$c^2(\mathbf{w}) = \|X\mathbf{w}\|^2 = \mathbf{w}^\top X^\top X\mathbf{w} = -\frac{1}{2}\mathbf{w}^\top J_\mathbf{s}^\top D_V J_\mathbf{s}\mathbf{w}.$$

This means that $c(\mathbf{w})$ does not depend on a particular set of embedding points $\{\mathbf{x}_i\}$ and is hence well defined for any portfolio $\mathbf{w}$. Corollary 3.2 says that portfolio $\mathbf{w}$ can be embedded on a sphere with the same radius as that of the Gower sphere and its center being the $\mathbf{w}$-weighted average of $\{-\mathbf{x}_i\}$. Hence, the centrality measures how far the center is from the origin, which corresponds to MDRP. In the numerical part, we will investigate the relationship between centrality and the diversification return. It turns out that the portfolios on the efficient frontier have a larger centrality than those studied in this paper. Other numerical observations call for more mathematical investigation on this concept.

The following example illustrates the key features in MDRP.

Example 3.1. Suppose there are three risky assets S_i, $i=1,2,3$, whose covariance matrix V, the corresponding distance matrix D, and its inverse D^{-1} are, respectively, given by

$$V = \frac{1}{9}\begin{bmatrix} 11 & 8 & 8 \\ 8 & 23 & -4 \\ 8 & -4 & 23 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}, \qquad D^{-1} = \frac{1}{6}\begin{bmatrix} -9 & 3 & 3 \\ 3 & -1 & 1 \\ 3 & 1 & -1 \end{bmatrix}.$$

We note that V is also nonsingular. It follows from Theorem 3.1 that the MDRP is $\mathbf{w}^* = (-1,1,1)^\top$. The corresponding embedding points are $\mathbf{x}_1 = (1,0)^\top$, $\mathbf{x}_2 = (1/2, \sqrt{3}/2)^\top$, and $\mathbf{x}_3 = (1/2, -\sqrt{3}/2)^\top$. They lie on the sphere of radius $R_G = 1$ centered at the origin. On the other hand, the equal-weight portfolio $\mathbf{w} = (1/3,1/3,1/3)^\top$ has the embedding points $\mathbf{y}_1 = (0,1/3)^\top$, $\mathbf{y}_2 = (-\sqrt{3}/2,-1/6)^\top$, $\mathbf{y}_3 = (\sqrt{3}/2,-1/6)^\top$. They lie on the sphere of radius $R=1$, but centered at $(0,-2/3)$. In other words, MDRP seeks an investment that places all assets on the sphere of minimal radius centered at the origin. If the origin is not within the convex hull of the embedding points, the portfolio may have to short certain assets (e.g., asset S_1 in this example). We note that MDRP has the diversification return $\mathrm{DR}(\mathbf{w}^*) = R_G^2 = 1$, while the equal-weight portfolio has $\mathrm{DR} = 5/9$. Both MDRP and the equally weighted portfolios are illustrated in Figure 1.

We note that Theorem 3.1 does not require V being nonsingular. It is known that for risky assets that share same variance the classical minimum variance portfolio of those risky assets is also MDRP [22, 20]. It is also known that when all risky assets share the same variance and have the same level of correlation, the equally weighted portfolio is MDRP. In fact, the result can be generalized to the following.

Corollary 3.4. *Consider n risky assets with V being its covariance matrix. Let $D = D_V$ be the corresponding distance matrix from V. If the vector $\mathbf{1}_n$ is an eigenvector of D, then the equal-weight portfolio $\mathbf{w}_{1/n}$ is MDRP.*

Proof. Since $\mathbf{1}_n$ is an eigenvector of D, there must exist a positive eigenvalue $\lambda_1 > 0$ (assuming $D \neq 0$) such that $D\mathbf{1}_n = \lambda_1 \mathbf{1}_n$. Suppose the rest of the $(n-1)$ eigenvalues are denoted by λ_i, $i=2,\ldots,n$, with the corresponding normalized eigenvectors $\mathbf{p}_i$, $i=2,\ldots,n$, which are orthogonal to $\mathbf{1}_n$, i.e., $\mathbf{1}_n^\top \mathbf{p}_i = 0$ for $i=2,\ldots,n$. We choose D^- to be the Moore–Penrose inverse:

$$D^- = \frac{1}{n\lambda_1}\mathbf{1}_n\mathbf{1}_n^\top + \sum_{i=2}^{n} \frac{1}{\lambda_i}\mathbf{p}_i\mathbf{p}_i^\top,$$

where $1/\lambda_i$ is taken to be 0 for $\lambda_i = 0$. Therefore,

$$D^-\mathbf{1}_n = \frac{1}{\lambda_1}\mathbf{1}_n \qquad \text{and} \qquad \mathbf{w}^* = \frac{D^-\mathbf{1}_n}{\mathbf{1}_n^\top D^-\mathbf{1}_n} = \frac{1}{n}\mathbf{1}_n.$$

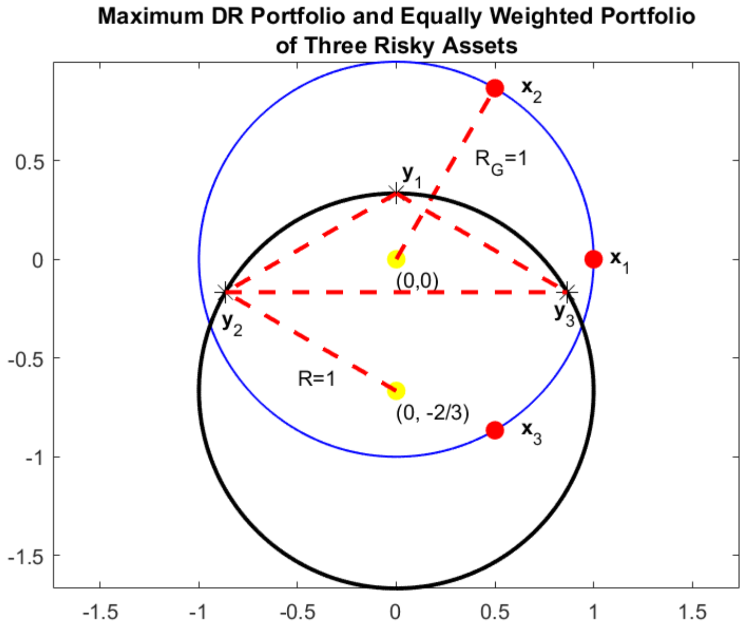


Figure 1. *For the data in Example 3.1, the Gower sphere is centered at the origin with $R_G = 1$ and the three embedding points are $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$. The equally weighted portfolio also lies on a unit sphere with the center $(0, -2/3)$ being away from the origin, and its corresponding embedding points are $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$. The origin is in the convex hull of the embedding points $\mathbf{y}_i$, $i = 1, 2, 3$. However, the equal-weight portfolio is not optimal. The optimal portfolio is $\mathbf{w}^* = (-1, 1, 1)$, which places the center of its unit circle at the origin.*

That is, the equal-weight portfolio is MDRP. ■

The case of equal variance σ^2 and same level of correlations at ρ among n risky assets becomes a direct consequence of Corollary 3.4. This is because for this case $D\mathbf{1}_n = (1 + (n-1)\rho)\sigma^2\mathbf{1}_n$. An example that yields $D\mathbf{1}_n = \lambda\mathbf{1}_n$ for some $\lambda > 0$ but has different level of correlations is the following:

$$V = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

In fact, the condition $D\mathbf{1}_n = \lambda\mathbf{1}_n$ for some $\lambda > 0$ is closely related to regular figures investigated in [15]. Hence, many such examples can be derived from those figures, which results in the equal-weight portfolio being MDRP.

3.2. Maximum volatility portfolio. The maximum volatility portfolio studied in [22] and [20, App. 2] aims to use a riskless asset to increase the diversification return. The main purpose of this subsection is to explain that a formula for the maxVP in [22, 20] may not be well defined and we provide a unified computational formula based on the distance matrix D. Moreover, we identify a necessary and sufficient condition for the riskless asset to (or not to) play a role in maxVP. We also study a condition when the diversification return of maxVP is guaranteed positive.

Suppose there is a riskless asset S_{n+1} available. The covariance matrix $\widetilde{V}$ of the $(n+1)$ risky asset is given by

$$\widetilde{V} = \begin{bmatrix} V & 0 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{\eta} = \begin{bmatrix} \eta \\ 0 \end{bmatrix}, \quad \text{and} \quad \widetilde{\mathbf{w}} = \begin{bmatrix} \mathbf{w} \\ w_{n+1} \end{bmatrix}.$$

The corresponding distance matrix

$$\widetilde{D} := \frac{1}{2}\left(\widetilde{\eta}\mathbf{1}_{n+1}^\top + \mathbf{1}_{n+1}\widetilde{\eta}^\top\right) - \widetilde{V} = \begin{bmatrix} D & \frac{1}{2}\eta \\ \frac{1}{2}\eta^\top & 0 \end{bmatrix}.$$

We use the language of $\widetilde{D}$ to define the maximum volatility portfolio in [22]:

$$\widetilde{\mathbf{w}}^* := \arg\max \ \frac{1}{2}\widetilde{\mathbf{w}}^\top \widetilde{D}\widetilde{\mathbf{w}} \quad \text{s.t.} \quad \mathbf{1}_{n+1}^\top \widetilde{\mathbf{w}} = 1. \tag{3.4}$$

We further note that

$$\begin{aligned} \widetilde{\mathbf{w}}^\top \widetilde{D}\widetilde{\mathbf{w}} &= \mathbf{w}^\top D\mathbf{w} + w_{n+1}\eta^\top \mathbf{w} \\ &= \mathbf{w}^\top D\mathbf{w} + (1 - \mathbf{1}_n^\top \mathbf{w})\eta^\top \mathbf{w} \\ &= \mathbf{w}^\top \left(D - \frac{1}{2}(\mathbf{1}_n\eta^\top + \eta\mathbf{1}_n^\top)\right)\mathbf{w} + \eta^\top \mathbf{w} \\ &= \eta^\top \mathbf{w} - \mathbf{w}^\top V\mathbf{w}. \end{aligned}$$

Therefore, problem (3.4) is equivalent to

$$\max \ \frac{1}{2}\left(\eta^\top \mathbf{w} - \mathbf{w}^\top V\mathbf{w}\right), \quad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} + w_{n+1} = 1,$$

which in turn is an unconstrained optimization problem:

$$\mathbf{w}_\eta := \arg\min_{\mathbf{w}\in\Re^n} \ \mathbf{w}^\top V\mathbf{w} - \eta^\top \mathbf{w}.$$

Its optimal solution is $\mathbf{w}_\eta = (1/2)V^{-1}\eta$. Both [22, 20] used its normalized version for the risky part of their maxVP:

$$\mathbf{w}_{\text{maxvp}} := \frac{V^{-1}\eta}{\mathbf{1}_n^\top V^{-1}\eta}. \tag{3.5}$$

One issue with this formula is that its denominator may be zero. To see this, let us consider $n = 2$ risky assets that have the covariance matrix

$$V = \begin{bmatrix} 2 & 1 \\ 1 & 2/3 \end{bmatrix} \qquad \text{and} \qquad V^{-1} = \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}.$$

It is straightforward to verify that $\mathbf{1}_n^\top V^{-1}\eta = 0$. Hence the formula (3.5) is not well defined for this example although we may calculate the numerator $V^{-1}\eta$. We also note that when V

is singular, maxVP may not even exist. Hence, we assume V is nonsingular in this part. By Lemma 2.3(ii), D is also nonsingular. Another issue with the maxVP (3.5) is that it may lead to negative diversification return. With some simplification, we have

$$\mathrm{DR}(\mathbf{w}_{maxvp}) = \frac{\mathbf{1}_n^\top V^{-1}\eta - 1}{2(\mathbf{1}_n^\top V^{-1}\eta)^2}\eta^\top V^{-1}\eta.$$

If $\mathbf{1}_n^\top V^{-1}\eta < 1$, then $\mathrm{DR}(\mathbf{w}_{maxvp})$ becomes negative. This possibility happened in our numerical experiment.

The following result aims to study when maxVP in (3.5) is well defined and when $\mathrm{DR}(\mathbf{w}_{maxvp})$ is positive. According to Theorem 3.1, we have a unified solution for (3.4):

$$\widetilde{\mathbf{w}}^* = \frac{\widetilde{D}^{-1}\mathbf{1}_{n+1}}{\mathbf{1}_{n+1}^\top \widetilde{D}^{-1}\mathbf{1}_{n+1}} =: \begin{bmatrix} \mathbf{w}_u^* \\ w_f^* \end{bmatrix}, \tag{3.6}$$

where we use $\mathbf{w}_u^*$ to denote the risky part and w_f^* for the riskless part. We will show $\widetilde{D}^{-1}$ exists. maxVP places the $(n+1)$ assets on a sphere and the $\widetilde{\mathbf{w}}^*$-weighted center is the origin of the sphere with radius $\widetilde{R}_G = \sqrt{1/(2\mathbf{1}_{n+1}^\top \widetilde{D}^{-1}\mathbf{1}_{n+1})}$.

Theorem 3.5. *Suppose the covariance matrix V is nonsingular. Then the following hold.*

(i) *It holds that*

$$\eta^\top D^{-1}\eta > 0 \qquad \text{and} \qquad \mathbf{1}_n^\top D^{-1}\eta > 1.$$

(ii) *The matrix $\widetilde{D}$ is nonsingular. Define*

$$\alpha_0 := \frac{\mathbf{1}_n^\top D^{-1}\eta - 2}{\eta^\top D^{-1}\eta} \qquad \text{and} \qquad c_0 := \mathbf{1}_n^\top D^{-1}\mathbf{1}_n - (\eta^\top D^{-1}\eta)\alpha_0^2.$$

We have

$$\widetilde{\mathbf{w}}^* = \begin{bmatrix} \mathbf{w}_u^* \\ w_f^* \end{bmatrix} = \frac{1}{c_0}\begin{bmatrix} D^{-1}\mathbf{1}_n - \alpha_0 D^{-1}\eta \\ 2\alpha_0 \end{bmatrix}.$$

Consequently, $\widetilde{\mathbf{w}}^* = (\mathbf{w}^*, 0)$ if only if $\mathbf{1}_n^\top D^{-1}\eta = 2$. Here $\mathbf{w}^*$ is the MDRP.

(iii) *The risky part $\mathbf{w}_u^*$ is a null portfolio (i.e., $\mathbf{1}_n^T\mathbf{w}_u^* = 0$) if and only if*

$$\mathbf{1}_n^\top D^{-1}\eta = 1 + \sqrt{1 + (\mathbf{1}_n^\top D^{-1}\mathbf{1}_n)(\eta^\top D^{-1}\eta)}.$$

Moreover, if $\mathbf{1}_n^\top D^{-1}\eta \le 2$, then $w_f^* \le 0$, $\mathbf{1}_n^\top \mathbf{w}_u^* \ge 1$, and maxVP is well defined and satisfies

$$\mathbf{w}_{maxvp} = \frac{\mathbf{w}_u^*}{\mathbf{1}_n^\top \mathbf{w}_u^*} \quad \text{and} \quad \mathrm{DR}(\mathbf{w}_{maxvp}) \ge \frac{1}{(\mathbf{1}_n^\top \mathbf{w}_u^*)^2}\mathrm{DR}(\mathbf{w}^*).$$

Proof. (i) We consider the following problem:

$$\max \frac{1}{2}\mathbf{w}^\top D\mathbf{w} \qquad \text{s.t.} \quad \eta^\top \mathbf{w} = 1. \tag{3.7}$$

Note that the objective function under the constraint becomes $\frac{1}{2}\mathbf{1}_n^\top \mathbf{w} - \frac{1}{2}\mathbf{w}^\top V\mathbf{w}$, which is strongly concave because V is nonsingular. Therefore, problem (3.7) has a unique optimal solution, denoted by $\mathbf{w}_\eta$, which satisfies the optimality condition: $D\mathbf{w}_\eta = \lambda_0\eta$ for some λ_0. Since D is elementwise nonnegative, the optimal objective is positive. Hence, $\lambda_0 = \mathbf{w}_\eta^\top D\mathbf{w}_\eta > 0$. It follows that $\eta^\top D^{-1}\eta = 1/\lambda_0 > 0$.

From the definition

$$D = \frac{1}{2}\left(\eta\mathbf{1}_n^\top + \mathbf{1}_n\eta^\top\right) - V$$

we get

$$D^{-1} = D^{-1}DD^{-1} = \frac{1}{2}\left(D^{-1}\eta\mathbf{1}_n^\top D^{-1} + D^{-1}\mathbf{1}\eta^\top D^{-1}\right) - D^{-1}VD^{-1}.$$

Multiplying $\eta^\top$ and η on both sides of the above identity, we have

$$\eta^\top D^{-1}\eta = (\eta^\top D^{-1}\eta)(\eta^\top D^{-1}\mathbf{1}_n) - \eta^\top D^{-1}VD^{-1}\eta,$$

which is equivalent to

$$(\eta^\top D^{-1}\mathbf{1}_n - 1)(\eta^\top D^{-1}\eta) = \eta^\top D^{-1}VD^{-1}\eta.$$

Since V is positive definite, the right-hand side is positive. We have proved that $\eta^\top D^{-1}\eta > 0$. Therefore, $\mathbf{1}_n^\top D^{-1}\eta > 1$.

(ii) Since D is nonsingular and $\eta^\top D^{-1}\eta \neq 0$, the Schur-complement of D in the matrix $\widetilde{D}$ is $-\eta^\top D^{-1}\eta \neq 0$. Hence, $\widetilde{D}$ is nonsingular and its inverse is given by

$$\widetilde{D}^{-1} = \frac{1}{\eta^\top D^{-1}\eta}\begin{bmatrix} (\eta^\top D^{-1}\eta)D^{-1} - D^{-1}\eta\eta^\top D^{-1}, & 2D^{-1}\eta \\ 2\eta^\top D^{-1}, & -4 \end{bmatrix}.$$

We calculate

$$\widetilde{D}^{-1}\mathbf{1}_{n+1} = \frac{1}{\eta^\top D^{-1}\eta}\begin{bmatrix} (\eta^\top D^{-1}\eta)D^{-1}\mathbf{1}_n + (2 - \mathbf{1}_n^\top D^{-1}\eta)D^{-1}\eta \\ 2(\mathbf{1}_n^\top D^{-1}\eta - 2) \end{bmatrix}. \tag{3.8}$$

Noticing $c_0 = \mathbf{1}_{n+1}^\top \widetilde{D}^{-1}\mathbf{1}_{n+1}$ and applying the normalization formula (3.6), we arrive at the stated solution $\widetilde{\mathbf{w}}^*$. It is obvious that $\alpha_0 = 0$ if and only if $\mathbf{1}_n^\top D^{-1}\eta = 2$ and the maxVP reduces to $(\mathbf{w}^*, 0)$.

(iii) The condition $\mathbf{1}_n^\top \mathbf{w}_u^* = 0$ implies

$$\mathbf{1}_n^\top D^{-1}\mathbf{1}_n - \alpha_0\mathbf{1}_n^\top D^{-1}\eta = 0.$$

Substitute the definition of α_0 into the equation and solve for $\mathbf{1}_n^\top D^{-1}\mathbf{1}_n$ to get the first claim in (iii). If $\mathbf{1}_n^\top D^{-1}\eta \le 2$, then the definition of α_0 ensures $\alpha_0 \le 0$. With the fact $c_0 > 0$, we have $w_f^* \le 0$. Consequently, $\mathbf{1}_n^\top \mathbf{w}_u^* = 1 - w_f^* \ge 1$. The fact that the optimal objective value of (3.4) is positive yields

$$0 < \frac{1}{2}(\widetilde{\mathbf{w}}^*)^\top \widetilde{D}\widetilde{\mathbf{w}}^* = \eta^\top \mathbf{w}_u^* - \frac{1}{2}(\widetilde{\mathbf{w}}^*)^\top V \widetilde{\mathbf{w}}^*,$$

which implies $\eta^\top \mathbf{w}_u^* > 0$. Since $(\mathbf{w}^*, 0)$ is feasible to problem (3.4), we have

$$\begin{aligned}
\mathrm{DR}(\mathbf{w}^*) &\le \frac{1}{2}(\widetilde{\mathbf{w}}^*)^\top \widetilde{D}\widetilde{\mathbf{w}}^* \\
&= \frac{1}{2}(\mathbf{w}_u^*)^\top D\mathbf{w}_u^* + w_f^*(\eta^\top \mathbf{w}_u^*) \\
&\le \frac{1}{2}(\mathbf{w}_u^*)^\top D\mathbf{w}_u^* = (\mathbf{1}_n^\top \mathbf{w}_u^*)^2 \mathrm{DR}(\mathbf{w}_{\mathrm{maxvp}}) \qquad (\text{using } \eta^\top \mathbf{w}_u^* > 0 \text{ and } w_f^* \le 0).
\end{aligned}$$

This establishes the lower bound for $\mathrm{DR}(\mathbf{w}_{\mathrm{maxvp}})$. ■

We make a brief comment on the condition $\mathbf{1}_n^\top D^{-1}\eta = 2$ for the case in which all risky assets share the same variance. That is, $\eta = \sigma^2 \mathbf{1}_n$ for some value σ^2. In this case, we have

$$\mathbf{1}_n^\top D^{-1}\mathbf{1}_n = \frac{2}{\sigma^2} \quad \text{and} \quad R_G^2 = \frac{1}{2\mathbf{1}_n^\top D^{-1}\mathbf{1}_n} = \frac{\sigma^2}{4}.$$

Therefore, $R_G = \frac{\sigma}{2}$. Since all embedding points of MDRP sit on the sphere of the radius R_G with the center at the origin, the Euclidean distance between any two points must be not bigger than the diameter of the sphere, leading to

$$\sqrt{D_{ij}} \le \sigma \qquad \forall\ i \ne j. \tag{3.9}$$

On the other hand, we have from the construction $D = (\eta\mathbf{1}^\top + \mathbf{1}\eta^\top)/2 - V$ that

$$D_{ij} = \sigma^2 - \sigma_{ij} = \sigma^2(1 - \rho_{ij}),$$

where ρ_{ij} is the correlation between assets S_i and S_j. Therefore, the condition (3.9) must require $\rho_{ij} \ge 0$ for all i, j. To put it another way, if there exist a pair of assets that have negative correlation $\rho_{ij} < 0$, then the condition $\mathbf{1}_n^\top D^{-1}\eta = 2$ cannot hold and the riskless asset must play an active role in maxVP.

3.3. Ridge-regularized MDRP. One of the major issues with MDRP is that it may contain negative weights on some assets. Various techniques have been proposed to convert MDRP to a long-only portfolio. One technique is to take only the positive weights in MDRP, and this strategy may lead to portfolio that has significantly fewer assets in it. In this section, we propose a regularized MDR which is guaranteed to be a long-only portfolio when the regularization parameter is set above the certain threshold. The regularized portfolio defined in (1.4) is copied below:

$$\mathbf{w}_\rho := \arg\max \frac{1}{2}\mathbf{w}^\top D\mathbf{w} - \frac{\rho}{2}\|\mathbf{w}\|^2, \qquad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1, \tag{3.10}$$

where $\rho \geq 0$ is a regularization parameter and $\|\mathbf{w}\|$ is the Euclidean norm of $\mathbf{w}$. We note that the regularization term in (3.10) is known as the Herfindahl concentration index in economics and finance (see, e.g., [19, sect. 1.2.3]).

A key property that we will use is the PF eigen-pair of D. Note that D is EDM and hence is a nonnegative matrix. We summarize some of the results about D below.

Lemma 3.6. *Let G be any (nonzero) $n \times n$ EDM. The following results hold.*

(i) [6, eq. (1028)] *G has just one positive eigenvalue, denoted as λ_1, and it has $(n-1)$ nonpositive eigenvalues. The largest eigenvalue is known as the PF eigenvalue and satisfies the following bound:*

$$\min_i \sum_{j=1}^n G_{ij} \leq \lambda_1 \leq \max_i \sum_{j=1}^n G_{ij}.$$

(ii) [24] *Assume further that G is irreducible (i.e., there exists an integer $k > 0$ such that $G^k > 0$ positive componentwise with G^k being the multiplication of G with itself k times). Let $\mathbf{p}_1$ be an eigenvector of G corresponding to its largest eigenvalue λ_1. The pair $(\lambda_1, \mathbf{p}_1)$ is known as the PF eigen-pair and $\mathbf{p}_1$ is strictly positive.*

By using the constraint $\mathbf{1}_n^\top \mathbf{w} = 1$, problem (3.10) is equivalent to

$$\begin{aligned} \mathbf{w}_\rho = \arg\max \quad & \tfrac{1}{2}\mathbf{w}^\top D\mathbf{w} - \tfrac{\rho}{2}\|\mathbf{w}\|^2 + \tfrac{\rho}{2} \\ = \arg\max \quad & \tfrac{1}{2}\mathbf{w}^\top D\mathbf{w} - \tfrac{\rho}{2}\|\mathbf{w}\|^2 + \tfrac{\rho}{2}\mathbf{w}^\top \mathbf{1}_n \mathbf{1}_n^\top \mathbf{w} \\ & = \tfrac{1}{2}\mathbf{w}^\top \Big(D + \rho \underbrace{(\mathbf{1}_n\mathbf{1}_n^\top - I)}_{=:D_0} \Big) \mathbf{w} \\ \text{s.t.} \quad & \mathbf{1}_n^\top \mathbf{w} = 1. \end{aligned} \tag{3.11}$$

By using the fact $J^2 = J$ and $J\mathbf{1}_n = 0$, we have

$$-JD_0J = \rho J \succeq 0 \quad \text{and} \quad \operatorname{diag}(D_0) = 0.$$

It follows from Lemma 2.1 that D_0 is Euclidean. Hence, the matrix $(D + \rho D_0)$ is Euclidean. By Theorem 3.1, we get

$$\mathbf{w}_\rho = \frac{(D+\rho D_0)^{-}\mathbf{1}_n}{\mathbf{1}_n^\top (D+\rho D_0)^{-}\mathbf{1}_n} \tag{3.12}$$

and the n risky assets can be embedded on a sphere centered at the origin with the radius $\sqrt{1/(2(\mathbf{1}_n^\top (D+\rho D_0)^{-}\mathbf{1}_n))}$. We would like to know when $\mathbf{w}_\rho$ results in long-only portfolios. We denote $D_\rho := D - \rho I$. We have the following results.

Theorem 3.7. *Suppose D is irreducible. Let $(\lambda_1, \mathbf{p}_1)$ be the PF eigen-pair. Let $p_{\min}$ be the smallest element in $\mathbf{p}_1$. The following hold.*

(i) *If $\rho = \lambda_1$, then*

$$\mathbf{w}_\rho = \frac{1}{\mathbf{1}_n^\top \mathbf{p}_1}\mathbf{p}_1 > 0.$$

(ii) *If ρ satisfies*

$$\frac{1}{\lambda_1} < \frac{1}{\rho} < \left(1 + \frac{p_{\min}}{(n-1)\sqrt{n}}\right) \frac{1}{\lambda_1},$$

then $\mathbf{w}_\rho > 0$.

(iii) *If ρ satisfies*

$$\rho \geq \left(1 + (n-1)\sqrt{n}\right) \max_i \sum_{j=1}^n D_{ij},$$

then $\mathbf{w}_\rho > 0$.

Proof. Since D is irreducible, the PF eigenvector is strictly positive by Lemma 3.6. Consequently $p_{\min} > 0$. Since the ridge regularization is used, the objective function in (1.4) is strictly concave, and its unique optimal solution satisfies the optimality condition:

$$D_\rho \mathbf{w} = D\mathbf{w} - \rho \mathbf{w} = \lambda \mathbf{1}_n \qquad \text{and} \qquad \mathbf{1}_n^\top \mathbf{w} = 1.$$

If $\rho = \lambda_1$ (the largest eigenvalue of D), then $D\mathbf{p}_1 = \rho \mathbf{p}_1$. The weight $\mathbf{w}_\rho$ in (i) satisfies the optimality condition with the Lagrange multiplier $\lambda = 0$. This proves (i).

We now prove (ii). We first note that D has one positive eigenvalue λ_1 and $(n-1)$ non-positive eigenvalues $\lambda_i \leq 0$, $i = 2, \ldots, n$. Suppose D has the following eigenvalue-eigenvector decomposition:

$$D = \lambda_1 \mathbf{p}_1 \mathbf{p}_1^\top + \sum_{i=2}^n \lambda_i \mathbf{p}_i \mathbf{p}_i^\top,$$

where $\mathbf{p}_i$ are the orthonormal eigenvectors corresponding to λ_i, $i = 1, \ldots, n$. Therefore,

$$D_\rho = (\lambda_1 - \rho) \mathbf{p}_1 \mathbf{p}_1^\top + \sum_{i=2}^n (\lambda_i - \rho) \mathbf{p}_i \mathbf{p}_i^\top.$$

Since $\rho \neq \lambda_1$ and $\lambda_i - \rho \leq -\rho < 0$ for $i = 2, \ldots, n$, D_ρ is invertible and

$$D_\rho^{-1} = \frac{1}{\lambda_1 - \rho} \mathbf{p}_1 \mathbf{p}_1^\top + \sum_{i=2}^n \frac{1}{\lambda_i - \rho} \mathbf{p}_i \mathbf{p}_i^\top. \tag{3.13}$$

It follows that

$$D_\rho^{-1} \mathbf{1}_n = \frac{\mathbf{1}_n^\top \mathbf{p}_1}{\lambda_1 - \rho} \mathbf{p}_1 + \sum_{i=2}^n \frac{\mathbf{1}_n^\top \mathbf{p}_i}{\lambda_i - \rho} \mathbf{p}_i.$$

The solution $\mathbf{w}_\rho$ is given by

$$\mathbf{w}_\rho = \frac{D_\rho^{-1} \mathbf{1}_n}{\mathbf{1}_n^\top D_\rho^{-1} \mathbf{1}_n}.$$

Using the facts $\mathbf{p}_1 > 0$, $\lambda_1 \geq \rho$, and $\lambda_i \leq 0$, $i = 2, \ldots, n$, we now establish a lower bound for each component of $D_\rho^{-1}\mathbf{1}_n$: for each $j = 1, \ldots, n$,

$$\left(D_\rho^{-1}\mathbf{1}_n\right)_j \geq \frac{\mathbf{1}_n^\top \mathbf{p}_1}{\lambda_1 - \rho} p_{\min} - \sum_{i=2}^n \frac{|\mathbf{1}_n^\top \mathbf{p}_i|}{\rho}$$
$$\geq \frac{1}{\lambda_1 - \rho} p_{\min} - \sum_{i=2}^n \frac{\sqrt{n}}{\rho} = \frac{1}{\lambda_1 - \rho} p_{\min} - \frac{(n-1)\sqrt{n}}{\rho}.$$

The second inequality used the facts $\mathbf{1}_n^\top \mathbf{p}_1 \geq \|\mathbf{p}_1\|_2 = 1$ and $|\mathbf{1}_n^\top \mathbf{p}_i\| \leq \sqrt{n}$. It is easy to see that $\mathbf{w}_\rho > 0$ under the stated condition.

(iii) Since D is assumed to be irreducible, Lemma 3.6(i) implies that the largest eigenvalue λ_1 of D satisfies $\lambda_1 \leq \max_i \sum_{j=1}^n D_{ij}$. By the choice of ρ in (iii), we have $\rho > \lambda_1$. Using (3.13) and the fact

$$\frac{1}{\rho - \lambda_i} - \frac{1}{\rho} = \frac{\lambda_i}{\rho(\rho - \lambda_i)}, \quad i = 1, \ldots, n,$$

we get

$$-D_\rho^{-1} = \frac{1}{\rho} I + \left(-D_\rho^{-1} - \frac{1}{\rho} I\right) = \frac{1}{\rho} I + \sum_{i=1}^n \left(\frac{1}{\rho - \lambda_i} - \frac{1}{\rho}\right) \mathbf{p}_i \mathbf{p}_i^\top$$
$$= \frac{1}{\rho} I + \sum_{i=1}^n \frac{\lambda_i}{\rho(\rho - \lambda_i)} \mathbf{p}_i \mathbf{p}_i^\top,$$

where the second equality used $I = \sum_{i=1}^n \mathbf{p}_i \mathbf{p}_i^\top$. Because λ_1 is the only positive eigenvalue of D and the trace of D is zero, we must have $|\lambda_i| \leq \lambda_1$ for all $i = 2, \ldots, n$. Then, for each index $j = 1, \ldots, n$,

$$\left(-D_\rho^{-1}\mathbf{1}_n\right)_j \geq \frac{1}{\rho} + \frac{\lambda_1 \mathbf{1}_n^\top \mathbf{p}_1}{\rho(\rho - \lambda_1)} (\mathbf{p}_1)_j - \sum_{i=2}^n \frac{|\lambda_i|}{\rho(\rho - \lambda_i)} |\mathbf{1}_n^\top \mathbf{p}_i|$$
$$> \frac{1}{\rho} - \sum_{i=2}^n \frac{|\lambda_1|}{\rho(\rho - \lambda_i)} \sqrt{n} = \frac{1}{\rho} - \frac{(n-1)\sqrt{n}\lambda_1}{\rho(\rho - \lambda_1)}.$$

The strict inequality above used the following facts: $\rho > \lambda_1$ and $\mathbf{p}_1$ is positive. Hence, $(-D_\rho^{-1}\mathbf{1}_n)_j > 0$ if

$$\rho > (1 + (n-1)\sqrt{n})\lambda_1.$$

The stated condition in (iii) is sufficient for the above inequality and hence

$$\mathbf{w}_\rho = \frac{-D_\rho^{-1}\mathbf{1}_n}{-\mathbf{1}_n^\top D_\rho^{-1}\mathbf{1}_n} > 0$$

by using the fact that the quantity $(-\mathbf{1}_n^\top D_\rho^{-1}\mathbf{1}_n)$ is necessarily positive under the condition on ρ. ∎

Theorem 3.7 states three scenarios that guarantee long-only portfolios. The first scenario (i) is to set the regularization parameter at the level of λ_1. This would generate a portfolio defined by the PF eigenvector. We call it the PF portfolio. Therefore, the PF portfolio, like MDRP, also has a Gower spherical representation. The second scenario (ii) is to set ρ below the level of λ_1, but close to it so that the ridge regularization pulls the optimal portfolio toward the PF portfolio. The third scenario (iii) is to set ρ well above λ_1 so that the regularization pulls the optimal portfolio toward the equal-weight portfolio. To see why it is the case, we note that the regularized problem (3.10) is actually equivalent to

$$\max \frac{1}{2}\mathbf{w}^\top D\mathbf{w} - \frac{\rho}{2}\|\mathbf{w}-\mathbf{w}_{1/n}\|^2 \quad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1, \tag{3.14}$$

using the fact $\mathbf{1}_n^\top \mathbf{w}_{1/n} = 1$. When ρ is sufficiently large, the regularization term dominates the diversification return term. Consequently, the optimal portfolio is pulled toward the equal-weight portfolio $\mathbf{w}_{1/n}$. The bounds on ρ in (ii) and (iii) are the worst-case analysis. It is important to note that it is quite possible that the values outside of those estimated intervals also lead to long-only portfolios. This possibility is illustrated by our numerical examples.

3.4. Long-only MDRP. This subsection shows that the spherical representation of MDRP also extends to the ℓMDRP (1.3). However, the difference from the previous subsections is that its representation involves an unknown quantity and needs to be computed by an iterative algorithm.

Since the feasible region of (1.3) is bounded, the ℓMDRP $\mathbf{w}_L^*$ is well defined and satisfies the following KKT conditions for some $\boldsymbol{g}_* \in \Re^n$ and $\lambda_* \in \Re$:

$$\begin{cases} D\mathbf{w}_L^* + \boldsymbol{g}_* = \lambda_* \mathbf{1}_n, \\ \mathbf{1}_n^\top \mathbf{w}_L^* = 1, \end{cases} \qquad \mathbf{w}_L^* \ge 0, \quad \boldsymbol{g}_* \ge 0, \quad (\boldsymbol{g}_* \circ \mathbf{w}_L^*) = 0. \tag{3.15}$$

Here $(\boldsymbol{g}_* \circ \mathbf{w}_L^*)$ is the vector that contains the Hadamard (componentwise) product of the two vectors. For any given $\boldsymbol{g} \in \Re^n$, define

$$V_{\boldsymbol{g}} := V + 2\mathrm{Diag}(\boldsymbol{g}), \quad \eta_{\boldsymbol{g}} := \mathrm{diag}(V_{\boldsymbol{g}}), \quad D_{\boldsymbol{g}} := \frac{1}{2}\left(\eta_{\boldsymbol{g}}\mathbf{1}_n^\top + \mathbf{1}\eta_{\boldsymbol{g}}^\top\right) - V_{\boldsymbol{g}}.$$

We have the following characterization of $\mathbf{w}_L^*$.

Proposition 3.8. *Let $(\mathbf{w}_L^*, \boldsymbol{g}_*, \lambda_*)$ be the KKT point defined by (3.15). Then the matrix $D_{\boldsymbol{g}_*}$ is EDM and*

$$\mathbf{w}_L^* = \arg\max \frac{1}{2}\mathbf{w}^\top D_{\boldsymbol{g}}\mathbf{w}, \quad \text{s.t.} \quad \mathbf{1}_n^\top \mathbf{w} = 1. \tag{3.16}$$

Proof. Since $\boldsymbol{g}_* \ge 0$, the matrix $V_{\boldsymbol{g}_*} = V + 2\mathrm{Diag}(\boldsymbol{g}_*)$ is positive semidefinite and hence is a legitimate covariance matrix. It follows from Lemma 2.3(ii) that $D_{\boldsymbol{g}_*}$ is EDM and

$$\begin{aligned} D_{\boldsymbol{g}_*} &= \frac{1}{2}\left((\eta + 2\boldsymbol{g}_*)\mathbf{1}_n^\top + \mathbf{1}_n(\eta + 2\boldsymbol{g}_*)^\top\right) - V - 2\mathrm{Diag}(\boldsymbol{g}_*) \\ &= D + \left(\boldsymbol{g}_*\mathbf{1}_n^\top + \mathbf{1}_n\boldsymbol{g}_*^\top\right) - 2\mathrm{Diag}(\boldsymbol{g}_*). \end{aligned}$$

It is easy to verify that

$$D_{\boldsymbol{g}_*}\mathbf{w}_L^* = D\mathbf{w}_L^* + \boldsymbol{g}_* = \lambda_* \mathbf{1}_n \quad \text{and} \quad \mathbf{1}_n^\top \mathbf{w}_L^* = 1,$$

where we used the properties in (3.15). We proved that $(\mathbf{w}_L^*, \lambda_*)$ satisfies the KKT conditions for the problem (3.16) and hence $\mathbf{w}_L^*$ is its optimal solution. ■

It follows from Theorem 3.1 that

$$\mathbf{w}_L^* = \frac{D_{\boldsymbol{g}_*}^- \mathbf{1}_n}{\mathbf{1}_n^\top D_{\boldsymbol{g}_*}^- \mathbf{1}_n}$$

and ℓMDRP $\mathbf{w}_L^*$ corresponds to a hyperspherical representation. However, $\mathbf{w}_L^*$ depends on the unknown Lagrange multiplier vector $\boldsymbol{g}_*$. This is in contrast to the previous portfolios (MDRP, maxVP, and rMDRP), which all enjoy a closed-form formula that does not involve any unknown quantity.

4. Application: Measuring distance between a new asset and MDRP.

Suppose we have a new asset S_{n+1} available. Our interest is not to compute the new MDRP of this asset together with the existing assets S_i, $i = 1, \ldots, n$, but to understand how far it is from the current MDRP. The geometric sphere representation of MDRP suggests that we may measure the distance between the new asset and the sphere, and it leads to the following computational approach.

Let $\sigma_{i,n+1}$ denote the covariance between S_{n+1} and S_i, $i = 1, \ldots, n$. Let $\widehat{V} := (\sigma_{ij})_{i,j=1}^{n+1}$ be the covariance matrix of those $(n+1)$ assets. According to Theorem 3.1(i), the matrix $\widehat{D} := (\widehat{\eta}\mathbf{1}_{n+1}^\top + \mathbf{1}_{n+1}\widehat{\eta}^\top)/2 - \widehat{V}$ is a Euclidean distance matrix, where $\widehat{\eta} := \mathrm{diag}(\widehat{V})$. Therefore, the last column of $\widehat{D}$ consists of the (squared) Euclidean distances between assets S_i and S_{n+1}:

$$\widehat{d}_{i,n+1}^2 := \frac{1}{2}(\sigma_i^2 + \sigma_{n+1}^2) - \sigma_{i,n+1}, \quad i = 1, \ldots, n.$$

Let MDRP$_n$ be the MDRP of the n risky assets and Sphere(R_G) denote the embedding sphere of MDRP$_n$. Let the n embedding points be $\mathbf{x}_i$, $i = 1, \ldots, n$. The new asset S_{n+1} is embedded onto Sphere(R_G) and we denote the embedding point as $\mathbf{x}_{n+1}$. We then calculate the pairwise (squared) Euclidean distance between $\mathbf{x}_{n+1}$ and $\mathbf{x}_i$:

$$d_{i,n+1}^2 = \|\mathbf{x}_i - \mathbf{x}_{n+1}\|^2, \quad i = 1, \ldots, n.$$

We define the embedding error by

$$\ell(x_{n+1}) := \sum_{i=1}^n \left(d_{i,n+1}^2 - \widehat{d}_{i,n+1}^2 \right)^2.$$

If $\ell(x_{n+1}) = 0$, then S_{n+1} can be exactly embedded onto Sphere(R_G). Adding S_{n+1} to the existing assets universe and calculating the new MDRP$_{n+1}$ would lead to the same portfolio given by MDRP$_n$, i.e., the weight on S_{n+1} will be zero. Naturally, we define the distance between S_{n+1} and MDRP$_n$ to be the least error loss:

$$d^2(S_{n+1}, \mathrm{MDRP}_n) := \min_{\mathbf{x}_{n+1} \in \mathrm{Sphere}(R_G)} \ell(\mathbf{x}_{n+1}). \tag{4.1}$$

This problem is very similar to the problem of adding a new point to an existing vector diagram [11] or landmark MDS [7]. Both considered the case where the existing points $\mathbf{x}_i$ are geometrically centered, i.e., $\sum_{i=1}^n \mathbf{x}_i = 0$, which is not satisfied here. Moreover, we require the embedding to be on a fixed hypersphere rather than to the space spanned by $\{\mathbf{x}_i\}$. We follow the approach by Gower [11] to derive a formula for (4.1). We show that the problem (4.1) can be solved in two steps:

S.1 Solve the problem without restricting $\mathbf{x}_{n+1}$ to the sphere. That is, we solve the problem

$$\mathbf{x}_{n+1} = \arg\min \ \ell(\mathbf{x}) \qquad \text{s.t.} \quad \mathbf{x} \in \mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}, \tag{4.2}$$

where $\mathrm{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is the space spanned by $\mathbf{x}_i$, $i = 1, \ldots, n$.

S.2 The optimal solution of (4.1) is given by

$$\mathbf{x}^*_{n+1} = \begin{cases} R_G \frac{\mathbf{x}_{n+1}}{\|\mathbf{x}_{n+1}\|} & \text{if } \mathbf{x}_{n+1} \neq 0, \\ \text{any point } \mathbf{x} \text{ on Sphere}(R_G) & \text{otherwise.} \end{cases} \tag{4.3}$$

Let us collect what we have known for the embedding points $\mathbf{x}_i$, $i = 1, \ldots, n$. They can be computed from the decomposition of the matrix B in (2.3) with $\mathbf{s} = D^- \mathbf{1}_n / (\mathbf{1}_n^\top D^- \mathbf{1}_n)$. Let k be the rank of B and k is known as the embedding dimension. Suppose B has the eigenvalue-eigenvector decomposition

$$B = -\frac{1}{2} J_{\mathbf{s}}^\top D J_{\mathbf{s}} = [\mathbf{u}_1, \ldots, \mathbf{u}_k] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_k^\top \end{bmatrix},$$

where $\lambda_1 \geq \cdots \geq \lambda_k > 0$ are the positive eigenvalues of B and $\mathbf{u}_i$ are the orthonormal eigenvectors. The embedding points $\mathbf{x}_i$ can be obtained by

$$X := [\mathbf{x}_1, \cdots, \mathbf{x}_n] = \mathrm{Diag}\Big(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_k}\Big) \begin{bmatrix} \mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_k^\top \end{bmatrix}. \tag{4.4}$$

In particular, we have

$$X\mathbf{s} = 0, \quad XX^\top = \mathrm{Diag}\Big(\lambda_1, \cdots, \lambda_k\Big), \quad \|\mathbf{x}_i\| = R_G, \ i = 1, \ldots, n.$$

Proposition 4.1. *Let $\boldsymbol{\delta}$ be the column vector with its ith element being $(\widehat{d}^2_{i,n+1} - R_G^2)$. Then the optimal solution of (4.2) is given by*

$$\begin{aligned} \mathbf{x}_{n+1} &= \frac{1}{2}\Big(XX^\top\Big)^{-1} X(I - \mathbf{1}_n \mathbf{s}^\top)\boldsymbol{\delta} \\ &= \frac{1}{2}\mathrm{Diag}\Big(1/\sqrt{\lambda_1}, \ldots, 1/\sqrt{\lambda_k}\Big) \begin{bmatrix} \mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_k^\top \end{bmatrix} (\mathbf{1}_n \mathbf{s}^\top - I)\boldsymbol{\delta}. \end{aligned}$$

Moreover, $\mathbf{x}^*_{n+1}$ *defined in* (4.3) *solves the problem* (4.1) *and*

$$\mathbf{s}^\top \boldsymbol{\delta} \geq \|\mathbf{x}_{n+1}\|^2 \quad \textit{and} \quad d^2(S_{n+1}, \mathrm{MDRP}_n) = (\mathbf{s}^\top \boldsymbol{\delta} - \|\mathbf{x}_{n+1}\|^2) + (R_G - \|\mathbf{x}_{n+1}\|)^2.$$

Proof. Since $\widehat{d}^2_{i,n+1}$ are squared Euclidean distances between S_{n+1} and S_i, $i=1,\ldots,n$ and the embedding points $\mathbf{x}_i$, $i=1,\ldots,n$ are in $\Re^k$, there must exist a representation of S_{n+1} in $\Re^{k+1}$ denoted as $\widehat{\mathbf{x}}_{n+1}$:

$$\widehat{\mathbf{x}}_{n+1} = \begin{bmatrix} \mathbf{x}_{n+1} \\ x_0 \end{bmatrix},$$

where $\mathbf{x}_{n+1} \in \Re^k$ and $x_0 \in \Re$ satisfies

$$\|\widehat{\mathbf{x}}_{n+1} - \widehat{\mathbf{x}}_i\|^2 = \widehat{d}^2_{i,n+1}, \quad \widehat{\mathbf{x}}_i := \begin{bmatrix} \mathbf{x}_i \\ 0 \end{bmatrix} \in \Re^{k+1}, \quad i=1,\ldots,n.$$

Expanding leads to

$$\widehat{d}^2_{i,n+1} = \|\widehat{\mathbf{x}}_{n+1}\|^2 + \|\mathbf{x}_i\|^2 - 2\mathbf{x}_i^\top \mathbf{x}_{n+1}, \quad i=1,\ldots,n. \tag{4.5}$$

Multiplying each equation above by s_i and adding them together yields

$$\sum_{i=1}^n s_i \widehat{d}^2_{i,n+1} = \|\widehat{\mathbf{x}}_{n+1}\|^2 + \sum_{i=1}^n s_i \|\mathbf{x}_i\|^2 - 2\Big(X\mathbf{s}\Big)^\top \mathbf{x}_{n+1} = \|\widehat{\mathbf{x}}_{n+1}\|^2 + \sum_{i=1}^n s_i R_G^2,$$

where we used the facts $\mathbf{1}_n^\top \mathbf{s} = 1$ and $X\mathbf{s} = 0$. Hence,

$$\|\widehat{\mathbf{x}}_{n+1}\|^2 = \sum_{i=1}^n s_i (\widehat{d}^2_{i,n+1} - R_G^2) = \mathbf{s}^\top \boldsymbol{\delta}. \tag{4.6}$$

Substituting back to (4.5) we get

$$\mathbf{x}_i^\top \mathbf{x}_{n+1} = \frac{1}{2}\mathbf{s}^\top \boldsymbol{\delta} - \frac{1}{2}\delta_i, \quad i=1,\ldots,n,$$

whose vector form is

$$X^\top \mathbf{x}_{n+1} = \frac{1}{2}\Big(\mathbf{1}_n \mathbf{s}^\top - I\Big)\boldsymbol{\delta}.$$

Multiplying the above equation by X on both sides we obtain

$$\Big(XX^\top\Big)\mathbf{x}_{n+1} = \frac{1}{2}X\Big(\mathbf{1}_n \mathbf{s}^\top - I\Big)\boldsymbol{\delta}.$$

Noticing the nonsingularity of $(XX^\top)$ and the formula (4.4), we arrive at the claimed characterization for $\mathbf{x}_{n+1}$. Moreover, the x_0-part in $\widehat{\mathbf{x}}_{n+1}$ can be computed from (4.6):

$$\|\mathbf{x}_{n+1}\|^2 + x_0^2 = \mathbf{s}^\top \boldsymbol{\delta} \implies x_0^2 = \mathbf{s}^\top \boldsymbol{\delta} - \|\mathbf{x}_{n+1}\|^2 \geq 0.$$

Let $\mathbf{h}^\top := (0, x_0) \in \Re^{k+1}$. It is easy to see that $\mathbf{h}$ is orthogonal to the space $\mathrm{Span}\{\mathbf{x}_i\}$ and $\mathbf{x}_{n+1}$ is the projection of $\widehat{\mathbf{x}}_{n+1}$ to $\mathrm{Span}\{\mathbf{x}_i\}$. The nearest point $\mathbf{x}^*_{n+1}$ of $\mathbf{x}_{n+1}$ to the $\mathrm{Sphere}(R_G)$ is defined by (4.3). Let $\mathbf{x}$ be any other point on $\mathrm{Sphere}(R_G)$, and

$$\widehat{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}, \qquad \widehat{\mathbf{x}}^*_{n+1} = \begin{bmatrix} \mathbf{x}^*_{n+1} \\ 0 \end{bmatrix}.$$

By the Pythagorean theorem, we have

$$\begin{aligned} \|\widehat{\mathbf{x}}_{n+1} - \widehat{\mathbf{x}}\|^2 = x_0^2 + \|\mathbf{x}_{n+1} - \mathbf{x}\|^2 &\geq x_0^2 + \|\mathbf{x}_{n+1} - \mathbf{x}^*_{n+1}\|^2 \\ &= \mathbf{s}^\top \boldsymbol{\delta} - \|\mathbf{x}_{n+1}\|^2 + (R_G - \|\mathbf{x}_{n+1}\|)^2. \end{aligned}$$

This gives the formula for $d^2(S_{n+1}, \mathrm{MDRP}_n)$. ■

We note that the landmark MDS approach [7] would lead to a different formula for $\mathbf{x}_{n+1}$. But it can be proved that the formula is equivalent to the one we obtained here. We omit the details. If the new asset S_{n+1} is not to make a contribution to the existing MDRP_n, we must have

$$\|\mathbf{x}_{n+1}\| = R_G \quad \text{and} \quad \mathbf{s}^\top \boldsymbol{\delta} = R_G^2,$$

so that $d^2(S_{n+1}, \mathrm{MDRP}_n) = 0$. In particular, if the new asset is chosen to be any existing asset S_i, then $d^2(S_i, \mathrm{MDRP}_n) = 0$, which implies $\mathbf{s}^\top \boldsymbol{\delta} = R_G^2$. Simplifying it leads to

$$\sum_{i=1}^n s_i D_{ij} = 2R_G^2, \quad i = 1, \ldots, n.$$

Those identities can be proved directly from the distance relations among the embedding points $\{\mathbf{x}_i\}$, but they are natural consequences of our characterization in Proposition 4.1.

5. Numerical illustration.

There exist extensive numerical experiments in [3, 20] that have demonstrated that diversification return driven portfolios can perform well on risk-adjusted returns in certain circumstances. Hence, it is not the purpose of this part to enhance those conclusions. Instead, through comparison with some benchmark portfolios, we draw a few key observations when using MDRP related portfolios.

5.1. Comparison of MDRP related portfolios.

(a) Data sets. We choose two real yet small data sets for our numerical illustration for two reasons. One is that the results reported can be easily reproduced, and the second reason is that they bring out contrasting behavior of MDRP related portfolios. The first data set consists of 30 stocks having appeared in the German DAX Index (GDAXI) and the second data set is from [21]. We describe the two data sets below.

Example 5.1. (DAX30 stocks) This data set consists of 30 stocks that have appeared in the DAX30 Index (DAX30) and was used in [16, p. 336]. The ticker symbols for those stocks are ADS.DE, ALV.DE, BAS.DE, BAYN.DE, BEI.DE, BMW.DE, CBK.DE, CON.DE, DAI.DE, DB1.DE, DBK.DE, DPW.DE, DTE.DE, EOAN.DE, FME.DE, FRE.DE, HEI.DE, HEN3.DE, IFX.DE, LHA.DE, LIN.DE, LXS.DE, MRK.DE, MUV2.DE, RWE.DE, SAP.DE, SDF.DE, SIE.DE, TKA.DE, VOW3.DE. The data period is from January 3, 2017 to December

Table 1

Eight agricultural commodities [21, Tab. 3]. The asset codes are defined as follows. Corn (CC), Live Cattle (CLC), Lean Hogs (CLH), Soybeans (CS), Wheat (CW), Cotton (NCT), Coffee (NKC), Sugar (NSB). Data period: January 1979 to March 2008.

Code	Return	σ	Correlation matrix (%)							
CC	4.5%	21.4%	100	2.7	4.2	61.8	51.6	13.9	4.6	9.3
CLC	17.2%	14.8%		100	31.0	4.5	3.5	2.5	0.8	3.7
CLH	14.4%	22.6%			100	7.0	5.9	5.0	−0.7	3.1
CS	10,5%	21.8%				100	42.8	16.2	6.3	10.4
CW	5.1%	23.7%					100	10.9	5.6	7.9
NCT	3.6%	23.2%						100	3.4	7.3
NKC	4.2%	36.5%							100	6.6
NSB	5.0%	43.8%								100

31, 2021. The mean and the covariance matrix of the daily returns were annualized in this experiment.[1]

Example 5.2 (agricultural assets). This data set consists of 8 light agricultural commodities taken from [21, Tab. 3] and is given in Table 1. It shows a large heterogeneity in volatilities and similarities of correlation coefficients around low levels (0%–10%).

(b) Portfolios. We recall from (3.1) that whenever maxVP is well defined, the MDRP can be represented as an affine combination of the minimum variance portfolio and maxVP:

$$\mathbf{w}^* = \alpha \mathbf{w}_{\mathrm{mvp}} + (1-\alpha)\mathbf{w}_{\mathrm{maxvp}} \quad \text{with} \quad \alpha = \mathbf{1}_n^\top V^{-1}\eta/2.$$

This motivates us to consider hybrid portfolios of two known portfolios $\mathbf{w}_1$ and $\mathbf{w}_2$:

$$\mathbf{w}_h := \alpha \mathbf{w}_1 + (1-\alpha)\mathbf{w}_2, \quad \alpha \in \Re.$$

If $\alpha \in [0,1]$, then $\mathbf{w}_h$ is a convex combination of $\mathbf{w}_1$ and $\mathbf{w}_2$. This class of portfolios was also studied in [23]. We refer to $\mathbf{w}_h$ as $\mathbf{w}_1$-$\mathbf{w}_2$ portfolios. We calculate DR at $\mathbf{w}_h$:

$$\mathrm{DR}(\mathbf{w}_h) = \mathrm{DR}(\mathbf{w}_2) + \alpha \mathbf{w}_2^\top D(\mathbf{w}_1 - \mathbf{w}_2) + \alpha^2 \mathrm{DR}(\mathbf{w}_1 - \mathbf{w}_2).$$

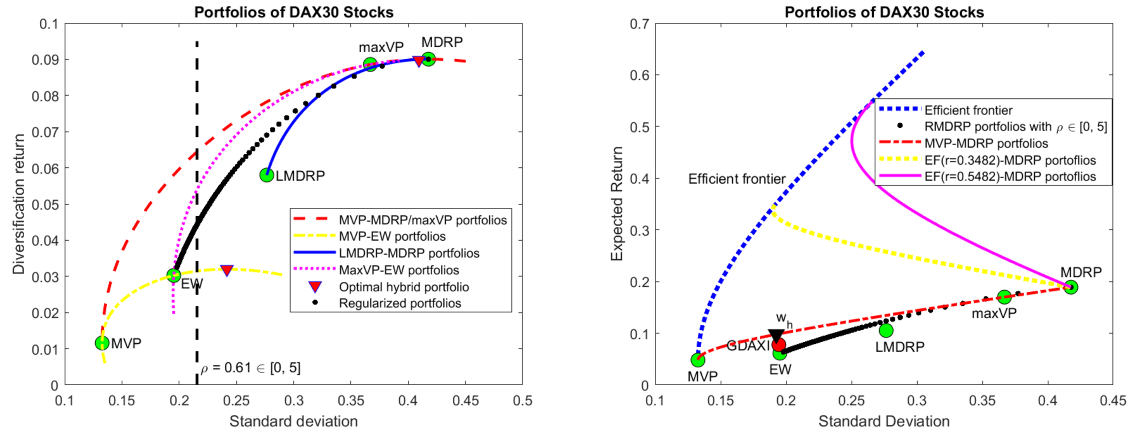
We like to compute the largest $\mathrm{DR}(\mathbf{w}_h)$:

$$\alpha^* := \arg\max \mathrm{DR}(\mathbf{w}_h), \qquad \text{s.t.} \quad \alpha \in \Re. \tag{5.1}$$

Since D is negative semidefinite on $\mathbf{1}_n^\perp$, $\mathrm{DR}(\mathbf{w}_1 - \mathbf{w}_2) \le 0$. Under the assumption that D is nonsingular, the problem (5.1) is strictly concave (i.e., $\mathrm{DR}(\mathbf{w}_1 - \mathbf{w}_2) < 0$. The optimal solution is given by

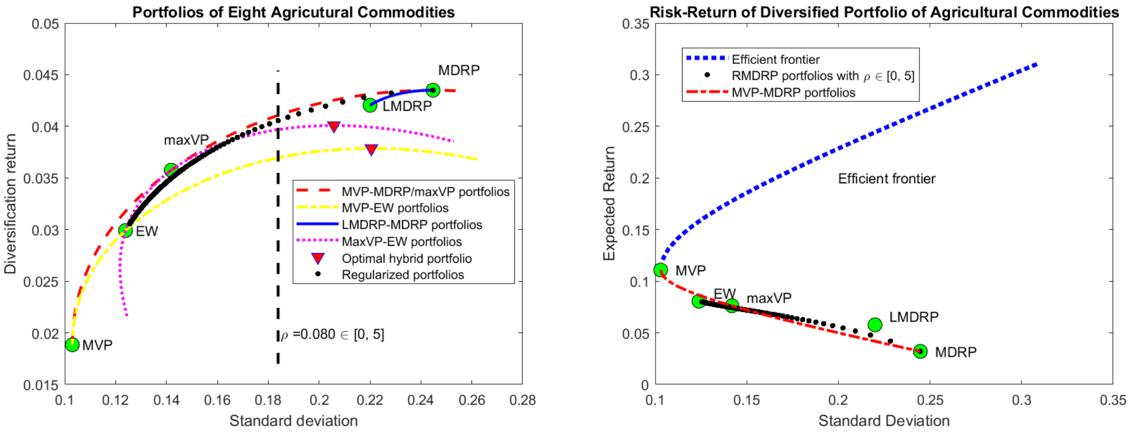
$$\alpha^* = -\frac{\mathbf{w}_2^\top D(\mathbf{w}_1 - \mathbf{w}_2)}{(\mathbf{w}_1 - \mathbf{w}_2)^\top D(\mathbf{w}_1 - \mathbf{w}_2)} \quad \text{and} \quad \mathbf{w}_h^* := \alpha^* \mathbf{w}_1 + (1-\alpha^*)\mathbf{w}_2.$$

[1] Suppose the observations are from time t_0 to time T. Let n denote the number of returns in this period and let N denote the number of calendar days between t_0 and T. The annualized time step is $\delta = N/(365n)$. Let $\widehat{\mu}$ and $\widehat{V}$ be the sample mean and covariance matrix of the returns. Then the annualized mean and covariance matrix are, respectively, given by $\mu = \widehat{\mu}/\delta$ and $V = \widehat{V}/\delta$. The value of δ for DAX30 data set is 0.003952.



(a) (b)

Figure 2. *Comparison of portfolios of DAX30 assets.*



(a) (b)

Figure 3. *Comparison of portfolios of eight agricultural commodities.*

The best portfolio $\mathbf{w}_h^*$ is called the optimal hybrid portfolio. There are many combinations. But we choose the following combinations that are enough to bring out our key observations: MVP-maxVP, MVP-MDRP, MVP-EW, and ℓMDRP-MDRP. We note that the MVP-maxVP and MVP-MDRP form the same set of hybrid portfolios because MDRP is an affine combination of MVP and maxVP in (3.1). We also study the behavior of ridge-regularized portfolios rMDRP when $\rho > 0$ varies. The diversification returns of those portfolios against the corresponding standard deviations are plotted in Figure 2 (for the DAX30 assets) and Figure 3 (for the 8 agricultural commodities), along with the more familiar efficient-frontier graphs. Contrasting features were brought out by the two examples.

(c) Key observations. We have the following observations.

1. The first striking feature is that in both cases the hybrid MVP-MDRP portfolios dominate others in the sense that the hybrid MVP-MDRP yields the highest diversification return given a level of standard deviation. It is more like an efficient frontier in the space of standard deviation and the diversification return. However, it is not easy to theoretically prove this is the case.
2. The ridge-regularized portfolios when $\rho > 0$ varies closely follow the EW-MDRP hybrid portfolios. This is consistent with the theory of rMDRP in the sense that at one end rMDRP is MDRP and at the other end it approximates equal-weight portfolio EW; see the comments following the problem (3.14). Therefore, rMDRP provides important portfolios that balance the EW portfolio and MDRP. However, rMDRP is sensitive to the choice of ρ. In both data sets, small ρ (0.61 in Figure 2(a) and 0.0800 in Figure 3(a)) can lead to long-only portfolios. In particular, for the first data set, when $\rho \geq 0.61$, it generates portfolios that are close to the EW portfolio, leaving not much room to search for other long-only portfolios. We also note that the PF portfolio is also one of the regularized portfolios.
3. Optimal hybrid portfolios also show interesting features. We depicted two such portfolios (optimal MVP-EW and maxVP-EW) for both cases. The other optimal hybrid portfolio is MDRP. For DAX30, the optimal hybrid maxVP-EW is very close to MDRP, and the optimal hybrid MVP-EW is close to EW. However, for the agricultural commodity dataset, they are very different from their generators (EW, maxVP, MVP). More numerical experiments would be needed to quantify what benefit those new portfolios would bring out in terms of risk-adjusted returns.
4. Comparing with the efficient portfolios in Figures 2(b) and 3(b), the portfolios related to the diversification return are far from being efficient. Would this suggest that those portfolios should be less preferred in practice? In Figure 2(b), we also plotted the risk and return of the GDAXI, which can be regarded as a market portfolio. It can be seen that GDAXI is also far from being efficient but is very close to the hybrid MVP-MDRP portfolio line. In fact, the dominating portfolio, denoted as $\mathbf{w}_h$ on the line over GDAXI, can be calculated. Its return and risk (standard deviation) as well as those from other portfolios are reported in Table 2. Under the same risk, the portfolio $\mathbf{w}_h$ would make 2% more return than GDAXI. We also note that the return per unit risk of the efficient frontier is probably too high for DAX30 data and does not reflect the true market performance.

The observation above seems to suggest that diversification return related portfolios and the efficient portfolios are at opposite ends of the spectrum of certain kinds of portfolios. This motivates us to investigate the intrinsic relationship among them. Recall from Corollary 3.2 that any portfolio $\mathbf{w}$ can be embedded on a sphere of the size of the Gower sphere, but

Table 2
Risk and return of particular portfolios on DAX30 stocks.

	MDRP	maxVP	LMDRP	$\mathbf{w}_h$	GDAXI	EW	MVP
Return (%)	18.91	16.99	10.53	9.78	7.79	6.14	4.82
Risk (σ)	0.4177	0.3669	0.2765	0.1926	0.1926	0.1953	0.1327

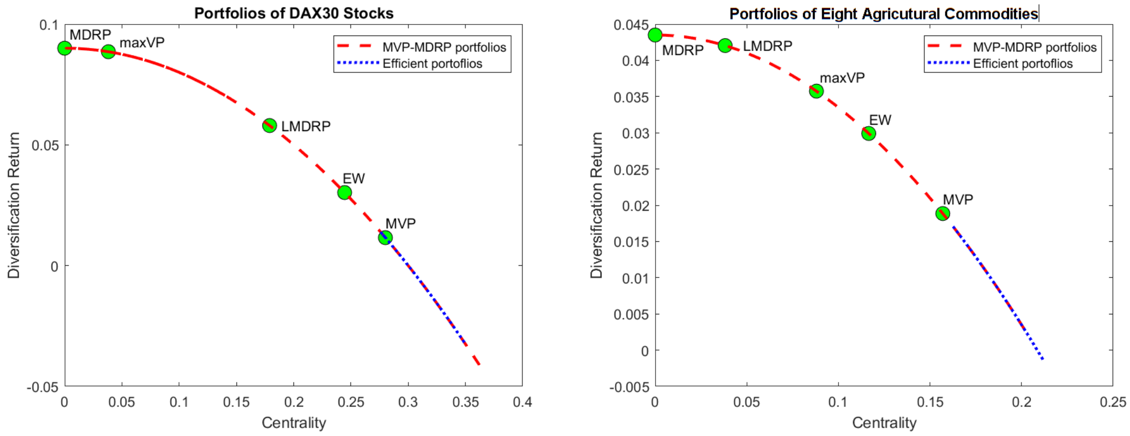


Figure 4. *Centrality vs. diversification return.*

with its center being away from the origin. The origin is the center of the Gower sphere representing MDRP. The distance $c(\mathbf{w})$ of the center to the origin is defined as the centrality of the portfolio in Definition 3.3. We plot the centrality of a portfolio against its diversification return $q(\mathbf{w})$ in Figure 4 for both data sets. We highlight two observations. One is that the centrality $c(\mathbf{w})$ and the diversification return $q(\mathbf{w})$ seem to form a smooth concave curve. Whether it is a parabola (similar to the efficient frontier in the risk-return plane) would need a mathematical proof. The second observation is that the curve is decreasing with MDRP at one end and the efficient portfolios at the other end. This echoes our previous observation that diversification-return-based portfolios are opposite to efficient portfolios. It is an interesting question whether efficient portfolios can be analyzed on this curve and this would provide a new perspective of the efficient frontier from the viewpoint of diversification return.

5.2. Asset distance and DR contribution. The main purpose of this part is to demonstrate that the distance measure developed in section 4 is a good indication of how an individual asset would make a DR contribution when added to an existing portfolio. The data sets above are too small for this purpose. Hence, we choose to use a bigger data set FF100,[2] accessed on November 9, 2022. The dataset contains daily average-value weighted returns of 100 portfolios (assets) from 1926 to 2022. The detail of the assets is also described in the companion webpage.[3] The 100 assets were formed in the following way. The original assets were first divided into 10 categories according to size (market equity, ME). They were also divided into 10 categories according to the ratio of book equity to market equity (BE/ME). Combining the categories of ME and BE/ME in pairs, there are 100 cases in total. This is the total number of newly formed assets in the dataset. This data set was recently studied in [33]. We consider two subdatasets: (a) the averaged monthly returns from November 1978 to June of 2020; in this dataset, after removing the assets with missing values, there are 96 assets left ($N = 96$). (b) the annualized daily returns from January 3, 2017 to December 31, 2020; there are 100 assets in this dataset ($N = 100$). The figures (a) and (b) in the figures

[2] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/100_Portfolios_10x10_Daily_CSV.zip.
[3] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_100_port_sz.html.

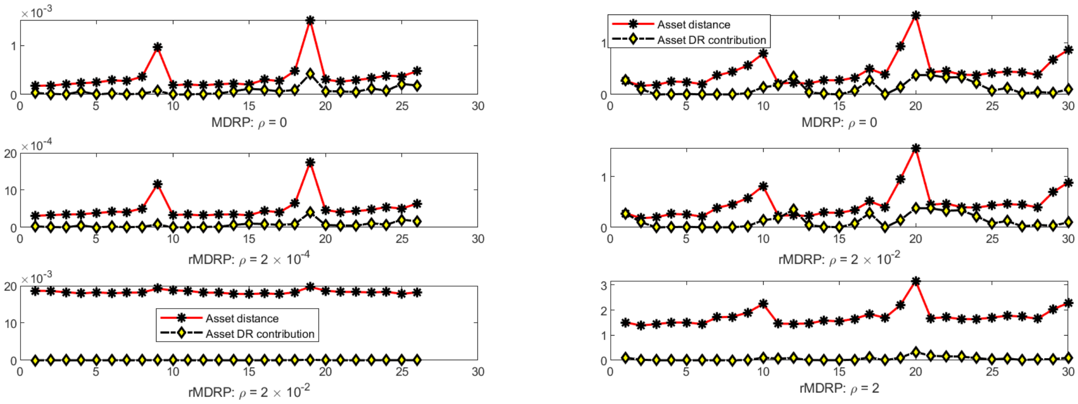
reported below, respectively, correspond to the datasets (a) and (b).

We tested two types of optimal portfolios: MDRP and rMDRP (with $\rho > 0$ varying). We use the first 70 stocks ($n = 70$) to form an optimal portfolio, which is denoted as Port_n. For the remaining assets, we calculate the distance d_i of each asset S_i to Port_n by $d_i = d^2(S_i, \text{Port}_n)$ using (4.1). The distances are well defined because both MDRP and rMDRP have spherical representations. We next calculate the asset's diversification return contribution by adding this new asset to Port_n and recalculate the DR of the new optimal portfolio, denoted by $\text{Port}_{n,i}$. The asset's DR contribution is defined to be

$$\text{DR}_i := \text{DR of Port}_{n,i} - \text{DR of Port}_n, \qquad i = n+1, \ldots, N.$$

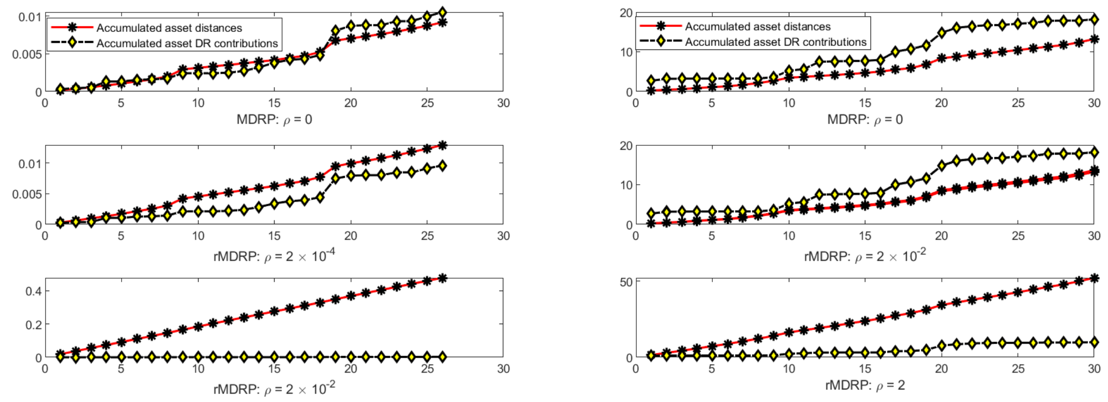
We plot DR_i against d_i for those remaining assets in Figure 5. It can be clearly seen for both data sets that d_i has a nice positive correlation with the corresponding DR contribution. In particular, whenever the distances are significantly far away from Port_n, the corresponding assets contributed more in terms of DR (see the top two panels of Figure 5). A potential use of the distance measure is to detect significantly different assets from the existing ones and those assets have potential to have a significant influence on portfolio construction. The bottom panel in Figure 5 also demonstrates one important point in portfolio construction, that the impact of the newly available asset on Port_n does not just depend on its distance to Port_n (as discussed above), but also depends on the property of Port_n. In the bottom panel case, the portfolio is close to EW portfolio and it seems that the DR of EW portfolio is very stable and is not conducive to change.

We further investigate whether the positive correlation between d_i and DR_i passes on when we consider their accumulated version. For k newly available assets denoted by K, we define their distance to Port_n and the collective DR contribution from K, respectively, by



(a) (b)

Figure 5. *Comparison of individual asset DR contribution and the corresponding distance for FF 100 data. Figure (5a) is based on monthly return data from November of 1978 to June of 2020. After dropping those assets that have missing values, there are 96 assets in this experiment. Figure (5b) is based on annualized daily return data from January 3, 2017 to December 31, 2020. There are 100 assets in these experiments.*



(a) (b)

Figure 6. *Comparison of accumulated DR contributions and the corresponding accumulated distances for FF* 100 *data. There exists a strong correlation between accumulated distances and the corresponding DR contributions* (*the top two panels*)*. The correlation in the bottom panel is less obvious as the underlying portfolios are close to the EW portfolio and are less sensitive to changes.*

$$D_K := \sum_{i \in K} d_i \quad \text{and} \quad \mathrm{DR}_K := \mathrm{DR\ of\ Port}_{n,K} - \mathrm{DR\ of\ Port}_n,$$

where $\mathrm{Port}_{n,K}$ is the portfolio of $(n+k)$ assets. We choose K to be $\{n+1\}$, $\{n+1, n+2\}$, and up to $\{n+1, n+2, \cdots, N\}$ (i.e., K contains one new asset, two new assets, and up to $N-n$ new assets). We plot the accumulated D_i against DR_K in Figure 6, where DR_K in Figure 6(a) was scaled 10 times to match the increase in D_i. This linear scaling is just for the convenience of visualization and does not present any issue in our key observations. As seen clearly from the top two panels of the plot, the positive correlation still prevails and it once again demonstrates that the proposed distance measure is a good indication how newly available assets would influence the DR contributions. The bottom panel enhances our observation that an EW-like portfolio is not conducive to DR changes as the DR contribution line stays flat.

6. Conclusion. Despite being a simple convex optimization problem in terms of the covariance matrix, the MDRP problem has a deep geometric interpretation in terms of Euclidean embedding. We derived such a geometric representation via the Rao's quadratic entropy. The resulting spherical representation of MDRP also extends to other portfolios including the long-only MDRP, the ridge-regularized MDRP, and the maximum volatility portfolios. We studied the weakness and strength of those portfolios and cautioned the use of the maximum volatility as it may result in negative diversification return. Those results rely on the fact that the distance matrix D_V is Euclidean. However, there are a few empirical diversification measures whose corresponding D is not Euclidean; see [3] for such instances. One important example is the distance matrix D given

$$D = \mathrm{diag}(\sqrt{\eta}) \Big(\mathbf{1}_n \mathbf{1}_n^\top - C \Big) \mathrm{diag}(\sqrt{\eta}),$$

where C is the Pearson correlation matrix of n assets. This matrix may not be Euclidean and hence the results obtained in this paper cannot be applied to this case. The optimization problem is not convex any more. It remains to be seen how to tackle such choices. In the numerical part, we studied hybrid portfolios consisting of two known portfolios. It is observed that the MVP-MDRP hybrid portfolios seem to dominate other portfolios in terms of the diversification return given the level of the standard deviation. It calls for theoretical justification when this observation is valid. We intend to tackle those questions in our future research.

Acknowledgement. We would like to thank the two referees for the stimulating questions that have led to the development of the concept of portfolio centrality.

REFERENCES

[1] A. Y. Alfakih, *Euclidean Distance Matrices and Their Applications in Rigidity Theory*, Springer, New York, 2018.
[2] D. G. Booth and E. F. Fama, *Diversification returns and asset contributions*, Financ. Anal. J., 48 (1992), pp. 26–32.
[3] B. Carmichael, G. B. Koumou, and K. Moran, *Rao's quadratic entropy and maximum diversification indexation*, Quant. Finance, 18 (2018), pp. 1017–1031.
[4] D. R. Chambers and J. S. Zdanowicz, *The limitations of diversification return*, J. Portfolio Manage., 40 (2014), pp. 65–76.
[5] K. Cuthbertson, S. Hayley, N. Motson, and D. Nitzsche, *Diversification Returns, Rebalancing Returns and Volatility Pumping*, https://dx.doi.org/10.2139/ssrn.2311240 (2015).
[6] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*, https://meboo.convexoptimization.com/Meboo.html (2010).
[7] V. De Silva and J. B. Tenenbaum, *Sparse Multidimensional Scaling Using Landmark Points*, Technical report, Stanford University, 2004.
[8] V. DeMiguel, L. Garlappi, F. J. Nogales, and R. Uppal, *A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms*, Manage. Sci., 55 (2009), pp. 798–812.
[9] C. Ding and H.-D. Qi, *Convex optimization learning of faithful Euclidean distance representations in nonlinear dimensionality reduction*, Math. Program., 164 (2017), pp. 341–381.
[10] C. B. Erb and C. R. Harvey, *The strategic and tactical value of commodity futures*, Financ. Anal. J., 62 (2006), pp. 69–97.
[11] J. C. Gower, *Adding a point to vector diagrams in multivariate analysis*, Biometrika, 55 (1968), pp. 582–585.
[12] J. C. Gower, *Euclidean distance geometry*, Math. Sci., 1 (1982), pp. 1–14.
[13] J. C. Gower, *Properties of Euclidean and non-Euclidean distance matrices*, Linear Algebra Appl., 67 (1985), pp. 81–97.
[14] J. T. Greene and D. A. Rakowski, *The Sources of Portfolio Returns: Underlying Stock Returns and the Excess Growth Rate*, https://dx.doi.org/10.2139/ssrn.1802591 (2011).
[15] T. Hayden and P. Tarazaga, *Distance matrices and regular figures*, Linear Algebra Appl., 195 (1993), pp. 9–16.
[16] Y. Hilpisch, *Python for Finance: Analyze Big Financial Data*, O'Reilly Media, Sebastopol, CA, 2014.
[17] M. Ho, Z. Sun, and J. Xin, *Weighted elastic net penalized mean-variance portfolio design and computation*, SIAM J. Financial Math., 6 (2015), pp. 1220–1244.
[18] S. Lesyk and A. Dougan, *The Role of Diversification Return in Factor Portfolios*, FTC Russell, 2021.
[19] F.-S. Lhabitant, *Portfolio Diversification*, Elsevier, New York, 2017.
[20] J.-M. Maeso and L. Martellini, *Maximizing an equity portfolio excess growth rate: A new form of smart beta strategy?*, Quant. Finance, 20 (2020), pp. 1185–1197.

[21] S. Maillard, T. Roncalli, and J. Teïletche, *The properties of equally weighted risk contribution portfolios*, J. Portfolio Manage., 36 (2010), pp. 60–70.
[22] D. Mantilla-Garcia, *Maximizing the Volatility Return: A Risk-Based Stratgey for Homegeneous Groups of Assets*, https://dx.doi.org/10.2139/ssrn.2740052 (2016).
[23] H. Mausser and O. Romanko, *Computing equal risk contribution portfolios*, IBM J. Res. Dev., 58 (2014), pp. 5:1–5:12.
[24] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.
[25] F. Meyer-Bullerdiek, *Rebalancing and diversification return-evidence from the German stock market*, J. Finance Invest. Anal., 6 (2017), pp. 1–28.
[26] H.-D. Qi, *A semismooth Newton method for the nearest Euclidean distance matrix problem*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 67–93.
[27] E. Qian, *Diversification return and leveraged portfolios*, J. Portfolio Manage., 38 (2012), pp. 14–25.
[28] C. Rao and T. Nayak, *Cross entropy, dissimilarity measures, and characterizations of quadratic entropy*, IEEE Trans. Inform. Theory, 31 (1985), pp. 589–593.
[29] C. R. Rao, *Diversity: Its measurement, decomposition, apportionment and analysis*, Sankhyā Indian J. Statist. Ser. A, 44 (1982), pp. 1–22.
[30] C. Ricotta and L. Szeidl, *Towards a unifying approach to diversity measures: Bridging the gap between the Shannon entropy and Rao's quadratic index*, Theoret. Popul. Biol., 70 (2006), pp. 237–243.
[31] I. Schoenberg, *Remarks to Maurice Frechet's article "Sur la definition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de Hilbert*," Ann. of Math., 36 (1935), pp. 724–732.
[32] S. Willenbrock, *Diversification return, portfolio rebalancing, and the commodity return puzzle*, Financ. Anal. J., 67 (2011), pp. 42–49.
[33] H. Zhao, L. Kong, and H.-D. Qi, *Optimal portfolio selections via $\ell_{1,2}$-norm regularization*, Comput. Optim. Appl., 80 (2021), pp. 853–881.