Big Data in Portfolio Allocation: *A New Approach to Successful Portfolio Optimization*

IRENE ALDRIDGE

IRENE ALDRIDGE is managing director of research at AbleMarkets in New York, NY, and visiting professor at Cornell University in New York, NY.
irene@ablemarkets.com

According to DeMiguel, Garlappi, and Uppal (2009, p. 1915), the idea of diversifying one's financial portfolio dates back at least to the fourth century AD, when Rabbi Issac bar Aha documented a rule for asset allocation in the Babylonian Talmud (Tractate Baba Mezi'a, folio 42a): "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand."

Modern portfolio theory originated from Markowitz (1952), and the body of work suggested not only diversifying assets and asset classes but also finessing portfolio composition by taking into account mutual co-movement of returns. Investments, the theory goes, should be diversified so that if or when one investment heads south, the others rise or at least counterbalance the total value of the portfolio. Co-movement of returns is often proxied by correlation matrixes. The optimal portfolio weights are computed to be directly proportional to the correlation matrix inverse.

When the number of positions is relatively small and stable, the classic Markowitz framework may work well. For larger portfolios, such as mutual funds and hedge funds with assets valued in the billions of US dollars, diversification suffers with unstable variance–covariance matrixes, costly reallocation requirements, and some illiquid positions. Exchange-traded funds further complicate the situation by providing a low-cost universe of potentially redundant securities that did not exist during Markowitz era, as described by Aldridge and Krawciw (2017). The correlation matrixes become very large. Big data techniques become necessary to intelligently reduce the size of the correlation matrixes, to select the key drivers in portfolios, and to remove redundant securities. Doing so helps portfolio managers improve transaction costs, stability of portfolio weights, and liquidity. With the advent of MiFID II and streamlined, potentially flat transaction fees per financial instrument, the smaller universe of financial instruments traded may be particularly beneficial to institutional investors.

Another benefit of reducing portfolio selection is the shortened history required for a robust performance estimation. As illustrated by DeMiguel, Garlappi, and Uppal (2009), increasing the number of instruments in the portfolio requires a significant increase in the length of historical data. Specifically, DeMiguel, Garlappi, and Uppal (2009) found that a portfolio of 25 assets with monthly reallocation requires a 250-year estimation window (across all positions) to reliably outperform the equally weighted (EW) strategy. This is a difficult requirement to fulfill considering reliable daily records have been kept for less than 70 years. DeMiguel, Garlappi, and Uppal (2009) also showed that the required estimating window scales linearly with the number of assets in the portfolio. Thus, a portfolio with five assets requires only 50 years of monthly data for reliable estimation.

Several techniques have been proposed over the years to mitigate the issues surrounding the Markowitz model. At the core of portfolio management is the following question: Which instruments should be removed and which ones kept? The decision is hardly trivial. Big data techniques do help to pinpoint the keepers in a reasonable time.

Traditional, not–big data solutions to the problem of optimal portfolio allocation fall roughly into two categories: Bayesian and non-Bayesian. Bayesian approaches include statistical, diffuse-priors, shrinkage estimators, and asset-pricing model priors. The diffuse-priors approach was pioneered by Barry (1974) and Bawa, Brown, and Klein (1979). The original shrinkage estimators date back to Jobson, Korkie, and Ratti (1979); Jobson and Korkie (1980); and Jorion (1985, 1986). The original asset-pricing models for establishing a prior were discussed by Pastor (2000) and Pastor and Stambaugh (2000) and, more recently, Brandt et al. (2005). They developed, for example, a simulation-based approach using recursion of approximations to the portfolio policy. Garlappi and Skoulakis (2008) simulated optimal portfolio choices using recursion of approximations to the portfolio value function.

Non-Bayesian non–big data approaches to minimizing estimation errors are similarly numerous. Goldfarb and Iyengar (2003) and Garlappi, Uppal, and Wang (2007) proposed robust portfolio optimization to deal with estimation errors using uncertainty structures and confidence intervals, respectively. MacKinlay and Pastor (2000) restricted the moments of returns by imposing factor dependencies. Best and Grauer (1992); Chan, Karceski, and Lakonishok (1999); and Ledoit and Wolf (2004a, 2004b) proposed methods for reducing the errors in the estimation of variance–covariance matrixes. Frost and Savarino (1988), Chopra and Ziemba (1993), and Jagannathan and Ma (2003) introduced short-selling constraints.

A separate stream of literature considers different portfolio optimization frameworks that depend on the concurrent market regime (i.e., bull versus bear market). For example, Ang and Bekaert (2002) used the Markov regime-switching model to show that regime-switching strategies that rely on macro factors as states outperform static portfolio allocation strategies out of sample.

Optimization problems from other disciplines with similarities to portfolio management and optimal asset allocation have been successfully studied in great detail in the field of big data analytics, and big data has been making inroads in portfolio management. Partovi and Caputo (2004) were the first to apply principal component analysis (PCA) to the portfolio choice problem to decompose principal portfolios uncorrelated by construction. Meucci (2009) followed up on the idea with the creation of maximum entropy portfolios. Garlappi and Skoulakis (2008) applied singular value decomposition (SVD) to solving several portfolio optimization problems in the context of the investor utility maximization. To do so, they deployed SVD to decompose state variables into fundamental drivers and shocks. The highest singular values or eigenvalues portray the drivers, whereas the lowest identify the shocks. Garlappi and Skoulakis (2008) applied the technique to solving the classic portfolio choice problem first proposed by Samuelson (1970) and extended by Hakansson (1971) and, later, Loistl (1976); Pulley (1981, 1983); Kroll, Levy, and Markowitz (1984); and Markowitz (1991), among others. A relatively recent stream of literature applies eigenvalue techniques to covariance matrixes to create eigenportfolios from any set of assets chosen by a researcher or a portfolio manager by some other evaluation criteria (see, for example, Steele (1995), Partovi and Caputo (2004), Avellaneda and Lee (2010), and Boyle (2014)).

The covariance and correlation matrixes, however, have been known to evolve, presenting a challenge to portfolio managers and researchers. Allez and Bouchaud (2012) studied eigenvalue evolution in covariance matrixes and attempted to find a time-based pattern of covariance evolution. They found that the covariance eigenvalues evolve over time, as expected. To deal with the estimation errors in the forward-looking correlation and covariance matrixes, Ledoit and Wolf (2017) proposed shrinking the sample covariance matrix toward a multiple of the identity matrix to push sample eigenvalues toward their mean. They proposed shrinking covariance matrixes by sampling eigenvalues in a nonlinear manner. Fan, Liao, and Mincheva (2013) developed a principal orthogonal complement thresholding method to estimate a high-dimensional covariance matrix with a conditional sparse structure and fast-diverging eigenvalues.

In this article I provide the first study of the big data properties of the inverse of the correlation matrix and show that the inverse is much more informative than the correlation matrix itself, from the big data perspective. Subsequently, the article proposes big data approaches to harness the correlation inverse and to deliver superior out-of-sample returns. The three key advantages of the method proposed in this article are conceptual simplicity, analytically tractable performance improvements, and empirically verified portfolio gains.

BIG DATA OVERVIEW

Many big data techniques, such as spectral decomposition, first appeared in the 18th century when researchers grappled with solutions to differential equations in the context of wave mechanics and vibration physics. Fourier has furthered the field of eigenvalue applications extensively with partial differential equations and other work.

At the heart of many big data models is the idea that the properties of every dataset can be uniquely summarized by a set of values, called *eigenvalues*. An eigenvalue is a total amount of variance in the dataset explained by the common factor. The bigger the eigenvalue, the higher the proportion of the dataset dynamics that eigenvalue captures.

Eigenvalues are obtained via either PCA or SVD. The latter technique is discussed in the following. The eigenvalues and related eigenvectors describe and optimize the composition of the dataset, perhaps best illustrated with an example of an image.

Consider the black-and-white image shown in Exhibit 1. It is a set of data points, *pixels* in computer lingo, whereby each data point describes the color of that point on a 0–255 scale, where 0 corresponds to pure black, 255 to pure white, and all other shades of gray lie in between. This particular image contains 960 rows and 720 columns.

EXHIBIT 1
Original Sample Image

Source: Courtesy Dr. Frank Fabozzi, 2018.

To perform spectral decomposition on the image, I use SVD, a technique originally developed by Beltrami (1873).[1] PCA is a related technique that produces eigenvalues and eigenvectors identical to those produced by SVD when PCA eigenvalues are normalized. Raw, non-normalized, PCA eigenvalues can be negative or positive and do not equal the singular values produced by SVD. For the purposes of the analysis presented here, we assume that all the eigenvalues are normalized, equal to singular values, and we will use the terms *singular values* and *eigenvalues* interchangeably throughout this article because the results presented can be developed using SVD and PCA techniques.

In SVD, a matrix X is decomposed into three matrixes: U, S, *and* V

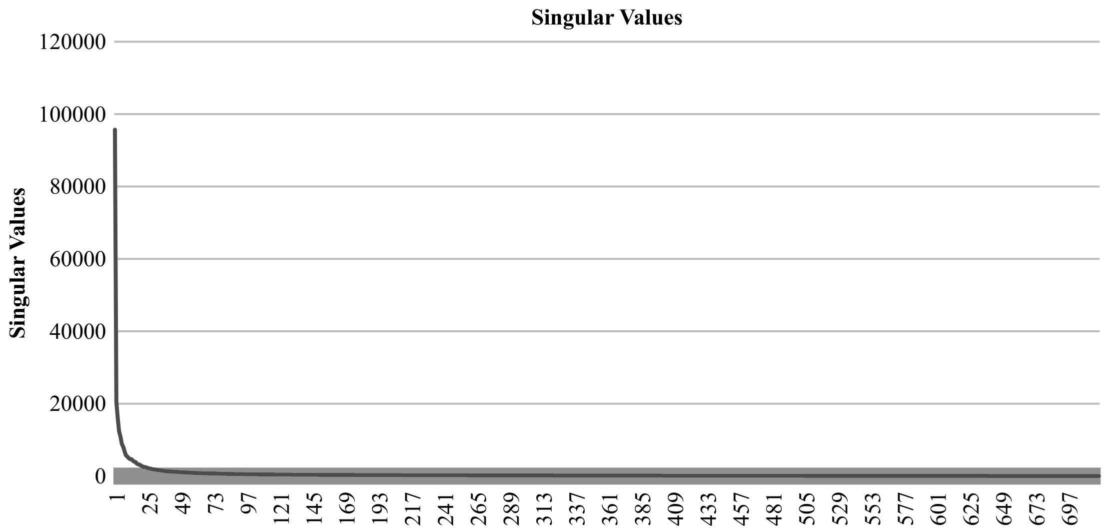
$$X = USV' \tag{1}$$

where X is the original $n \times m$ matrix; S is an $m \times m$ diagonal matrix of singular values or eigenvalues sorted from the highest to the lowest on the diagonal; V' is a transpose of the $m \times m$ matrix of so-called singular vectors, sorted according to the sorting of S; and U is an $n \times n$ user matrix containing characteristics of rows vis-a-vis singular values.

SVD delivers singular values sorted from largest to smallest. The plot of the singular values corresponding to the image in Exhibit 1 is shown in Exhibit 2. The plot of singular values is known as a *scree plot* because it resembles a real-life scree, a rocky mountain slope.

[1] For a detailed history of SVD, please see Stewart (1993).

EXHIBIT 2
Scree Plot Corresponding to the Image in Exhibit 1



A scree plot is a simple line segment plot that shows the fraction of total variance in the data as explained or represented by each singular value (eigenvalue). The singular values are ordered and are assigned a number label by decreasing order of contribution to total variance.

To reduce the dimensionality of a dataset, we select k singular values. If we were to use the most significant of the singular values, typically containing macroinformation common to the dataset, we would select the first k values. However, in applications involving idiosyncratic data details, we may be interested in the last k values (e.g., when we need to evaluate the noise in the system). A rule of thumb dictates breaking the eigenvalues into sets before the elbow and after the elbow sets in the scree plot.

What is the perfect number of singular values to keep in the image of Exhibit 1? An experiment presented in the seven panels in Exhibit 3 shows the evolution of the data with varying number of eigenvalues included. The eigenvalues and the corresponding eigenvectors composed of linear combinations of the original data create new dimensions of data. As the seven panels in the Exhibit 3 show, as few as 10 eigenvalues allow a human eye to identify the content of the image, effectively reducing dimensionality of the image from 720 columns to 10.

However, the guesswork is not at all needed because the optimal method of discarding the eigenvectors associated with the smallest eigenvalues has already been developed (see, for example, Carrasco, Florens, and Renault 2007). The method is known as the *spectral cutoff method*. Carrasco and Noumon (2011) further proposed a data-driven method to select the optimal number of principal components to be kept in the spectral cutoff method.

To create the reduced dataset, we restrict the number of columns in the S and V matrixes to k by selecting k first elements, determined by the spectral cutoff method. The resulting matrix $X_{reduced}$ has dimensions n rows and k columns, where

$$X_{reduced,nxk} = U_{nxk}\ S_{kxk}\ V_{kxk}^{T} \tag{2}$$

TRADITIONAL PORTFOLIO OPTIMIZATION AND BIG DATA APPLICATIONS

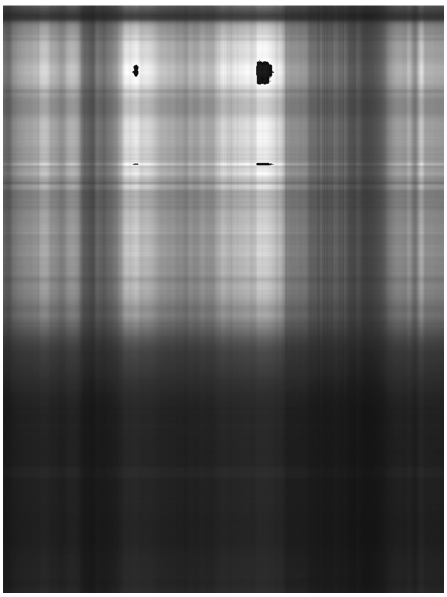
Markowitz-style portfolio optimization is often known as *mean–variance optimization* (MVO) because it seeks to increase mean returns while simultaneously decreasing variance in portfolios. Denoting the beginning prices of each asset i X_i, $i = 0, 1, \ldots, n$, we can express the investment portfolio as

$$w_0 X_0 + w_1 X_1 + \cdots + w_n X_n \tag{3}$$

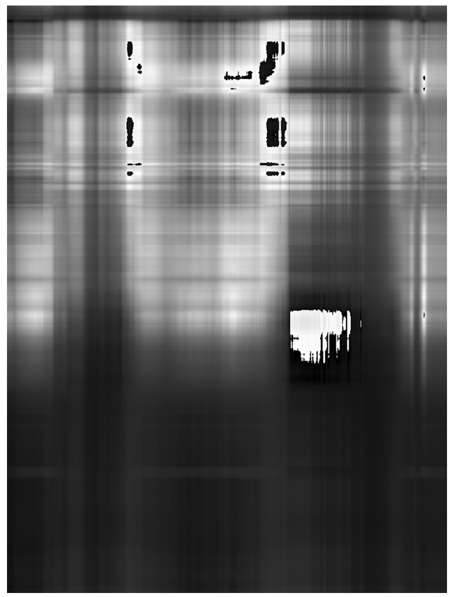
where w_i, $i = 0, 1, \ldots, n$ are portfolio weights: the proportion of the total portfolio wealth that is invested in

EXHIBIT 3
Reconstruction of the Image of Exhibit 1

Panel A: Reconstruction with Just the First Eigenvalue



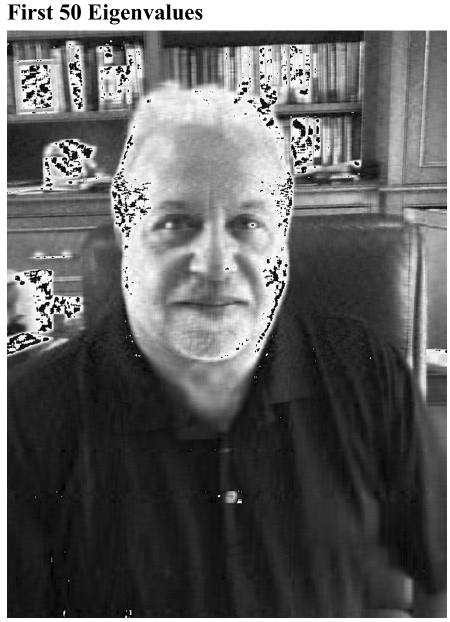
Panel B: Reconstruction with the First Two Eigenvalues



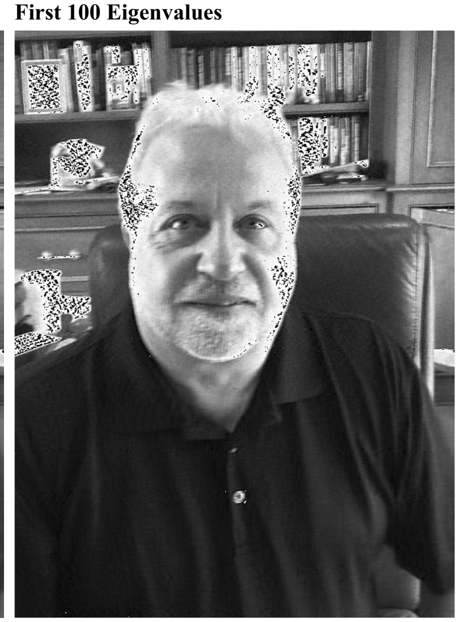
Panel C: Reconstruction with the First Five Eigenvalues (the outlines of the figure are beginning to appear)

Panel D: Reconstruction with the First 10 Eigenvalues

Panel E: Reconstruction with the First 50 Eigenvalues



Panel F: Reconstruction with the First 100 Eigenvalues



the asset i. The sum of the weights of the portfolio assets is then equal to 1, and $w_0 + w_1 + \cdots + w_n = 1$. The asset with $i = 0$ is often assumed to be the prevailing risk-free rate, denoted r_0.

Denoting risk aversion as γ, we now express the traditional MVO as follows:

$$max_{w,w_0}(w_0 r_0 + w'\mu - \gamma w'\Sigma w) \qquad s.t.\ w_0 + w'1 = 1 \quad (4)$$

where Σ represents the variance–covariance matrix of the returns of the n assets under consideration.

Subtracting the risk-free rate, the maximization problem can be rewritten as follows:

$$max_{w,w_0}(w'(\mu - r_0 1) - \gamma w'\Sigma w) \qquad s.t.\ w_0 + w'1 = 1 \quad (5)$$

Equation 2 then leads to the following optimal solution:

$$w = \frac{1}{2\gamma}\sum^{-1}(\mu - r_0 1) \qquad (6)$$

Although the vector of returns is typically assumed to be the long-running average of returns on assets under consideration (see, for example, Jegadeesh and Titman (1993)), the covariance matrix presents several challenges to researchers and practitioners. Specifically, the covariance matrix can in turn be decomposed into variance and correlation matrixes, although variances tend to be sticky and reasonably predictable by techniques such as generalized autoregressive conditional heteroskedasticity[2] and correlations of asset returns are notoriously volatile.[3] It is the properties of correlation matrixes that induce two key problems portfolio managers encounter when implementing MVO:

1. Possibly extreme positions in selected assets (i.e., a large proportion of the portfolio) resulting in liquidity constraints and violating the economic equilibrium of the portfolio allocation. To solve the issue, Black and Litterman (1993) and others proposed a blended solution between economic equilibrium and MVO.
2. Possibly extreme changes in portfolio weights from one investment period to the next, resulting in large transaction costs. Bertsimas and Lo (1998), Liu (2004), Muthuraman and Kumar (2006), Lynch and Tan (2008), and Mei, DeMiguel, and Nogales (2016), for example, propose penalizing the MVO function with transaction costs as the remedy to the problem. However, such methods often tend to be opaque in practice.

Big data techniques, such as spectral decomposition, have appealed to researchers for their data size reduction and stabilization properties but have produced variable results to date. Several techniques have been developed and popularized over the years, all deploying big data on the correlation matrix or, worse, on the covariance matrix itself, instead of tackling the root of the portfolio management woes: the correlation matrix inverse.

Reduction of the covariance matrix can be considered erroneous for the following reasons: The volatility properties have been well studied and can be successfully modeled independently of the correlation framework. As a result, including variances in the optimization bag together with the correlations prevents the researchers from finessing the optimization with the independent volatility properties.

The prevalent techniques for the stand-alone correlation optimization suffer from an even bigger flaw. The classic foundation technique, known as PCA, is at the heart of most current optimization frameworks for the correlation matrix. The technique decomposes the correlation matrix into its eigenvalue-related principal components and then shrinks the correlation matrix by setting the eigenvalue tail to zeros. The technique follows the principles of big data optimization discussed in the previous section.

Two immediate issues arise. First, the largest eigenvalue of the correlation matrix has long been known to be a market portfolio, whereas the eigenvalue tail corresponds to the idiosyncratic properties of the assets under consideration. Retaining the dominant market portfolio while discarding the idiosyncratic pieces goes completely against the spirit of the classical Markowitz optimization, which seeks instead to diversify away from the market. Second, setting eigenvalues to zero prior to matrix inversion renders matrixes singular and, therefore, noninvertible. In other words, reducing the spectral dimensionality of the correlation matrixes and subsequent inversion blow up the outcome. To overcome the issue, researchers often use *whitening*—replacing set-to-zero eigenvalues with white noise $N(0, 1)$ to allow matrix invertibility. The process introduces noise into the system, resulting in classic "garbage in, garbage out" situations well known in engineering disciplines.

Most models, such as shrinkage operators and Bayesian optimization frameworks, use the described faulty PCA as their underlying core, producing suboptimal

[2] See Engle (1982), Bollerslev (1986), and Andersen et al. (2006).

[3] See Davis and Mikosch (1998), Gourieroux (1997), and Cont (2001).

results. The same argument applies to recently popular eigenportfolios and other techniques that apply spectral decomposition or PCA to correlation or covariance matrixes, instead of correlation matrix inverses.

BIG DATA WITH THE INVERSE OF THE CORRELATION MATRIX: A NOVEL APPROACH

In contrast to the established techniques tackling the correlation matrix, big data application to the inverse of the correlation matrix appears to be more promising and robust. The eigenvectors of an invertible matrix are also the eigenvectors of the matrix's inverse. To show this, consider an invertible matrix A. Matrix A is invertible if and only if its determinant is not zero (Lipschutz 1991, p. 45), which in turn implies that matrix A columns are linearly independent, further implying that its eigenvalue λ is not zero. Suppose that matrix A has eigenvectors $\boldsymbol{v}$. By definition of eigenvectors, $A\boldsymbol{v} = \lambda\boldsymbol{v}$. Multiplying by A^{-1} from the left, we obtain

$$v = A^{-1}\lambda\boldsymbol{v} \tag{7}$$

$$A^{-1}v = (1/\lambda)\boldsymbol{v} \tag{8}$$

Another solution is to exploit the fact that singular values of a matrix may be found and the dimensions reduced after the inversion with equal success and without sacrificing data precision.

$$(AB)^{-1} = B^{-1}A^{-1} \tag{9}$$

More generally,

$$\left(\prod_{k=0}^{N} A_k\right)^{-1} = \prod_{k=0}^{N} A_{N-k}^{-1} \tag{10}$$

So, for SVD

$$A(p) = U(p)S(p)V'(p) \tag{11}$$

the inverse becomes

$$(A(p))^{-1} = (V')^{-1}S^{-1}U^{-1} \tag{12}$$

SVD of the inverse of the correlation matrix is, therefore, much more precise because no data are lost as a result of the poorly specified input to the inversion process that occurs with whitening methodology. Accordingly, the SVD in the case of the matrix inversion can be performed as follows: The spectral decomposition can be performed after the matrix inversion without sacrificing results.

If SVD decomposes a correlation matrix C into $C = USV^T$, then the inverse of the matrix C can be written as $C^{-1} = (V^T)^{-1}S^{-1}U^{-1}$, where S^{-1} is the inverse of the diagonal matrix S

$$S = \begin{matrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{matrix}$$

$$S^{-1} = \begin{matrix} 1/\lambda_n & 0 & 0 & \cdots & 0 \\ 0 & 1/\lambda_{n-1} & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & 1/\lambda_2 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1/\lambda_1 \end{matrix}$$

Inverting the correlation matrix first and then spectrally decomposing it to retrieve eigenvalues $\{\lambda\}$ therefore allows researchers to retain much more precision. Instead of replacing the irrelevant eigenvalues with noise to allow inversion, the proposed process is to replace the eigenvalues directly with 0 postinversion.

Which eigenvalues should you keep or discard? This, once again, is a nontrivial question. Spectral decomposition of the original, noninverted correlation matrix results in principal components or portfolios sorted according to their universality vis-a-vis all assets considered. Thus, the largest component often represents the global macro portfolio factor driving most of the performance and typically reflecting the broad market movement. Several of the following eigenvalues deliver portfolios that induce synchronized fluctuations of groups of stocks; these can be, for example, factors driving industries. The remaining small components are idiosyncratic in nature. Spectral decomposition of the inverted correlation matrix produces eigenvalues sorted in the opposite order: from smallest to the largest.

Numerous big data techniques have been developed to help us understand the information content of the matrix under consideration—in our case, the inverse of the correlation matrix. Here, we develop and prove a conjecture that the top eigenvalue information content of the inverse of the correlation matrix always exceeds that of the correlation matrix itself. As a result, the big data analysis pertaining to the optimal portfolio allocation should be carried out on the correlation matrix inverse, not on the correlation matrix as is done at present. The invert-then-optimize methodology proposed in this article, and diametrically opposite to established methodologies, not only delivers superior results but also delivers explicit tractable solutions to the most-cited woes of existing portfolio optimization methodologies: correlation instability and extreme portfolio weights.

In short, the proposed methodology is to retain the largest eigenvalues in the inverse of the correlation matrix. These eigenvalues correspond to the smallest eigenvalues of the original correlation matrix, the values discarded in traditional analyses. We show that these values, long known to contain idiosyncratic properties of assets, are indeed key to successful portfolio optimization.

CORRELATION MATRIXES VERSUS INVERSES: STABILITY AND SENSITIVITY TO PERTURBATIONS

Given that the main problems associated with large-scale portfolio optimization revolve around the instability of the resulting portfolio weights, the objective of the decomposition should be to preserve the most stable components and to discard the least stable ones. Much of the traditional literature interprets this as retaining the top eigenvalues of the correlation matrix and discarding the smallest values. However, this does not make sense given that the final portfolio weights are proportional to the inverse of the correlation matrix instead. As this section shows, the inverse of the correlation matrix is necessarily less stable than the correlation matrix itself; to stabilize portfolios, one needs to stabilize the inverse of the correlation matrix, not the correlation matrix itself.

A vast stream of literature focusing on the stability of matrixes and their sensitivity to perturbations dates back to Gershgorin (1931). Gershgorin circles allow us to identify the span of possible values for eigenvalues in our system. The Gershgorin circles define the radii around each a_{ii} in a matrix A, within which lies eigenvalue i

$$\left|\lambda_i - a_{ii}\right| = \sum_{i \neq j}\left|a_{ij}\right| \tag{13}$$

The tighter the Gershgorin circle around i, the more stable the eigenvalue i to small perturbations in the matrix under consideration. Correspondingly, the larger the Gershgorin circle around i, the less stable the ith eigenvalue and the more sensitive the matrix is to even the smallest changes in the underlying data.

Gershgorin circles form a convenient visual representation of the sensitivity of data to small perturbations. As an example, consider just five equities (A, AA, AAL, AAMC, and AAN) over a three-week period ending October 27, 2017, with the summary statistics shown in Exhibit 4. Exhibit 5 shows the normalized eigenvalues of the correlation matrix, the respective Gershgorin radii of the correlation matrix, the eigenvalues of the inverse of the correlation matrix, and the Gershgorin radii of the inverse of the correlation matrix.

The two panels in Exhibit 6 represent the resulting Gershgorin circles visually. As this exhibit shows, the Gershgorin circles of the inverse are much larger, indicating that the inverse of the matrix is much more unstable than the sample correlation matrix itself.

Similar empirical results can be obtained with the Bauer–Fike theorem (Bauer and Fike, 1960) and other methods, such as the Robinson and Wathen (1992) method. The Bauer–Fike theorem proposes comparing operator vector norms of eigenvectors. The vector norms serve as upper bounds for perturbations for respective eigenvectors. Exhibit 7 shows the upper bounds for matrix perturbations for vanilla correlation and correlation inverse matrixes for data in Exhibit 1. As shown in Exhibit 7, the bounds on the inverse of the correlation matrix are considerably higher than that on the correlation matrix itself, implying once again that the inverse of the correlation matrix is much less stable than the correlation matrix.

Similar results can be obtained using key relation between matrixes, ordered eigenvalues {□}, and matrix inverses derived by Robinson and Wathen (1992)

Exhibit 4

An Illustration of Gershgorin Circles on Sample Correlation Matrixes

	Mean	St Dev	Corr A	AA	AAL	AAMC	AAN
A	0.001384	0.008118	1.000	0.391	–0.024	0.315	–0.150
AA	0.003440	0.019856	0.391	1.000	0.344	0.365	–0.194
AAL	–0.00064	0.017241	–0.024	0.344	1.000	0.123	–0.125
AAMC	–0.00400	0.023821	0.315	0.365	0.123	1.000	0.693
AAN	–0.00390	0.014468	–0.150	–0.194	–0.125	0.693	1.000

$$\frac{1}{\lambda_1}+\frac{(\lambda_1-1)^2}{\lambda_1(\lambda_1-s_{ii})}\le (A^{-1})_{ii}\le \frac{1}{\lambda_n}-\frac{(\lambda_n-1)^2}{\lambda_n(-\lambda_n+s_{ii})},$$

where $s_{ii}=\sum_k a_{ik}{}^2$.

A formal theoretical conclusion showing the higher instability of the correlation inverse is as follows: The largest eigenvalue of the inverse of the correlation matrix is always larger than the largest eigenvalue of the correlation matrix itself. The proof of this theoretical conclusion is provided in the online supplement.

The obtained results are independent of the underlying distribution of returns. Indeed, the result accommodate Gaussian, leptokurtic, and other distributions with equal effect, making the strategy robust to a variety of financial return models. Furthermore, the result of the theoretical conclusion extends far beyond financial data and is applicable to any datasets, whether advertising, healthcare, or genomics.

SENSITIVITY OF CORRELATION MATRIXES VERSUS THEIR INVERSES: SIMULATION

To ascertain the validity of our conjecture, we perform 10,000 experiments of the following nature:

1. We create a random symmetric 100×100 matrix $\{A_{ij}\}$ simulating the real-life correlation structure: All the values on the diagonal are set to 1.0, and all other values for $i \neq j$ range in the interval [−1.0, 1.0], with entries $a_{ij} = a_{ji} = \forall i, j$.
2. We compute and document the eigenvalues of the correlation matrix and its inverse.

As the results presented in Exhibit 8 illustrate, the top eigenvalue of the inverse is considerably higher than the top eigenvalue of the correlation matrix itself. As the simulation results show, it is the inverse of the correlation matrix that is the unstable component of the portfolio optimization puzzle. Because the portfolio weights are directly proportional to the inverse of the correlation

Exhibit 5

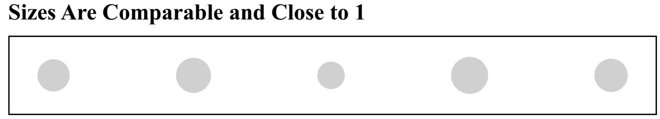
Comparative Dispersion of Eigenvalues via Gershgorin Radii for the Correlation Matrix of Exhibit 4 and the Inverse

Normalized Eigenvalues of the Correlation Matrix	Gershgorin Radii of the Correlation Matrix	Normalized Eigenvalues of the Inverse of the Correlation Matrix	Gershgorin Radii of the Inverse of the Correlation Matrix
0.09290	0.880	0.54085	3.41071
0.46752	1.294	0.63756	4.07172
1.02214	0.616	0.97834	1.76692
1.56849	1.496	2.13897	8.49492
1.84896	1.162	10.7639	8.09458

Exhibit 6

Gershgorin Circles of the Sample Correlation Matrix of Exhibit 4 and the Inverse of the Matrix, Graphical Representation

Panel A: Circles of the Correlation Matrix All Centered on 1; Sizes Are Comparable and Close to 1



Panel B: Circles of the Inverse of the Correlation Matrix (Centers: 1.69597, 2.14316, 1.23938, 5.32429, and 4.65683; Radii: 3.41071, 4.07172, 1.76692, 8.49492, and 8.09458)

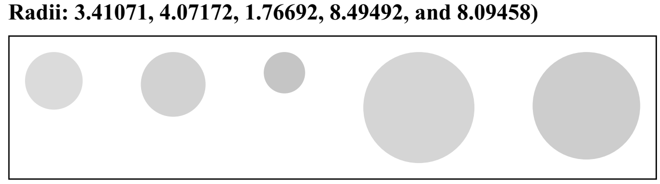


EXHIBIT 7

Bauer–Fike Norms for Eigenvectors of Correlations and Inverse Correlations of Data in Exhibit 4

	||V|| Corr	||V–1|| Corr	||V|| Corr Inverse	||V–1|| Corr Inverse
1	2.10	2.11	2.10	2.11
2	**1.31**	**1.24**	**1.45**	**1.60**
∞	2.11	2.10	2.11	2.10

Note: Bolded value show much higher inverse dispersion.

EXHIBIT 8

Summary Statistics for Eigenvalues of 10,000 Simulated Correlation Matrixes and Their Inverses

	Top 1	Bottom 1	Top 1 Inverse	Bottom 1 Inverse
Mean	11.60537822	0.05523877	**312.46715013**	0.08622607
St Dev	0.30493585	0.04221966	**8,508.6027317**	0.00225207
Skew	0.30964001	1.00654776	**50.50362873**	–0.14982966
Kurt	0.17120805	0.88971901	**2,776.6559658**	0.02559518
Max	12.83546300	0.25704200	500,000.00000	0.09435499
99%	12.39773146	0.18056210	1,545.2583525	0.09108325
95%	12.12671910	0.13857045	262.28504037	0.08983642
90%	12.00535900	0.11442080	121.92152061	0.08909438
75%	11.80207150	0.08023350	46.19737847	0.08777734
50%	11.59161850	0.04672950	21.39976314	0.08626923
25%	11.39246225	0.02164625	12.46362244	0.08473089
10%	11.22405270	0.00820200	8.73966974	0.08329613
5%	11.13134355	0.00381265	7.21654767	0.08246254
1%	10.97896710	0.00064715	5.53826083	0.08065992
Min	10.59827400	0.00000200	3.89041480	0.07790915

Note: Bolded value show much higher inverse dispersion.

matrix, stabilization and other optimization of the inverse of the correlation matrix—not the correlation matrix itself—are critical for successful portfolio allocation.

OUT-OF-SAMPLE APPLICATIONS TO FINANCIAL DATA

I next test the theory (the importance of the optimization of the correlation matrix inverse) on the historical financial data. I performed two experiments:

1. Comparison of the core portfolio management techniques on the S&P 500 data for the 20-year period from 1998 through 2017, with monthly reallocation
2. Comparison of the portfolio management techniques on 1,000 portfolios with 50 or more stocks each, the constituents of which were randomly drawn from the S&P 500 from 1998 through 2017, with monthly reallocation

Both experiments show that regardless of portfolio composition, the correlation inverse optimization proposed in this article significantly outperforms the other core portfolio allocation strategies.

Out-of-Sample Application to the S&P 500

The test uses daily closing price data for the S&P 500 constituents for the 20-year period spanning 1998–2017 and obtained from Yahoo!. We assume monthly portfolio reallocation and test the following strategies on the S&P 500 data: EW, vanilla MVO, PCA with the top eigenvalues retained, and PCA_Inverse with the bottom eigenvalues of the inverse taken into account and the bottom eigenvalues discarded.

To compute strategy performance, the lognormal daily returns from the price data are first determined

$$r_t = log(P_t) - log(P_{t-1}) \tag{14}$$

Next the monthly correlation matrixes using the returns falling into each calendar month in the 1998–2017 span are computed. Each correlation matrix then serves as an input to the strategy evaluation over the following month. For example, the correlation matrix computed on January 30, 1998 serves as the input for portfolio selection for February 1998.

Monthly performance of the strategies is next measured using the strategy weights computed on the last day of the previous month using the daily returns for the previous month. For analytical tractability, the risk aversion coefficient is chosen to be 1; however, it can be easily scaled up or down because the portfolio weights of the MVO, PCA of MVO, and PCA_Inverse of MVO strategies are directly proportional to the risk aversion coefficient. The performance evaluation applies the weights to the returns observed on the last trading day of the following month vis-a-vis the price levels observed on the last day of the portfolio creation month. Thus, the performance of portfolios created on January 30, 1998 is tested by returns observed from the closing price on January 30, 1998 to the closing price observed on February 27, 1998.

The four panels in Exhibit 9 document the performance of the monthly reallocation of the strategies. As the exhibit shows, the PCA_Inverse strategy outperforms the other strategies when the number of selected eigenvalues is small, such as the top one eigenvalue selected in the PCA_Inverse strategy shown in Panel A (with outliers) and Panel B (outliers removed for clarity) of Exhibit 9. As the number of retained eigenvalues increases, the PCA_Inverse strategy loses its power and eventually yields to the EW strategy.

Exhibit 10 shows the Sharpe ratios from the obtained strategies. As the exhibit shows, the PCA_Inverse strategy consistently outperforms other portfolio management strategies, particularly when the outliers, such as extreme one-time returns, are discarded from the data. Exhibit 11 presents average monthly returns for each strategy computed over the 1998–2017 period. As shown in the exhibit once again, the PCA_Inverse strategy delivers superior results when a concentrated number of eigenvalues is deployed to create an optimal portfolio allocation.

The results of the analysis so far show that just the top eigenvalue of the inverse of the correlation matrix contains enough portfolio information to outperform the other strategies. Just how many instruments does such a strategy contain? Exhibits 12 and 13 help answer this question. The number of positions with the absolute value greater than or equal to 2% of the total portfolio value varied throughout the 20-year period; the number of stocks was significantly smaller than that of other strategies, pointing to a smart diversification portfolio selection of the PCA_Inverse strategy.

Bootstrapping the S&P 500: Technique Comparison on Randomly Selected Subportfolios over 1998–2017 Period

To anticipate the objections of researchers and portfolio managers dealing with assets other than the prim and proper S&P 500 and to showcase the strength and capability of the correlation inverse optimization proposed in this article, the following tests were conducted:

1. On January 1, 1998, we randomly select 50 or more names from the S&P 500. There are no restrictions on the name selections or their quantity, other than the randomly chosen portfolio must include at least 50 names. As noted earlier in this article, portfolios of fewer than 50 names are considered suitable for vanilla MVO and may not be as interesting for our purposes.
2. The four core portfolio management strategies with monthly reallocation on the portfolio randomly chosen in Step 1 were then run: (a) EW; (b) MVO; (c) spectral decomposition and optimization via PCA of the asset correlation matrix (PCA), retaining the top eigenvector only; and (d) the methodology proposed in this article, spectral decomposition and optimization of the inverse of the asset correlation matrix (PCA_Inverse), again, retaining only the top eigenvector, this time of the inverse.

The portfolio compositions do not change from 1998 through 2017. The portfolio weights are computed on the last trading day of each month. The EW weights do not change, unless the originally chosen stock is no longer trading. For MVO, PCA, and PCA_Inverse, the correlation matrixes used to set portfolio weights for the following month are computed on the last day of each trading month using daily log returns based on closing prices for the past month. Thus, the correlation matrix used to compute the weights for March 2005 is determined on the last trading day of February 2005 using all the closing daily returns for February 2005, including the first and the last trading days.

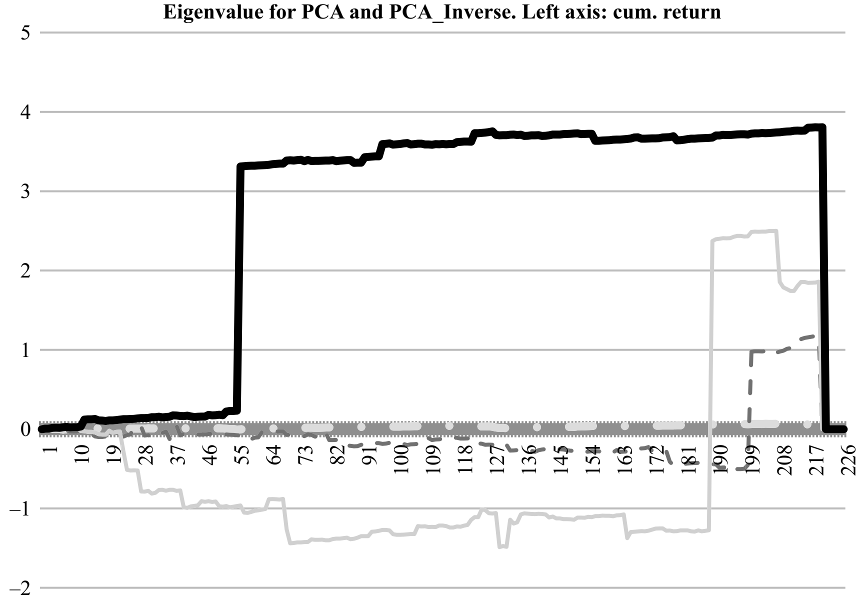
The traditional PCA approach to the correlation matrix is analogous to the eigenportfolio selection. As our analysis shows, the methodology on the correlation inverse PCA (PCA_Inverse) proposed in this article is far superior to the plain eigenportfolio construction. Panel A in Exhibit 14 shows the cumulative returns of the four core strategies averaged by month across 30 random draws of 50 or more securities comprising the S&P 500. Panel B shows standard deviations of the 30 independent repetitions by month from 1998 through 2017, illuminating outliers. As the two panels of Exhibit 14 show, even with severe outliers, the proposed methodology significantly outperforms other methods, regardless of portfolio construction. Panel C of Exhibit 14 shows the cumulative returns of PCA_Inverse over the 20-year period from 1998 through 2017.

Exhibit 9

Portfolio Strategy Performance Comparison, S&P 500, 1998–2017

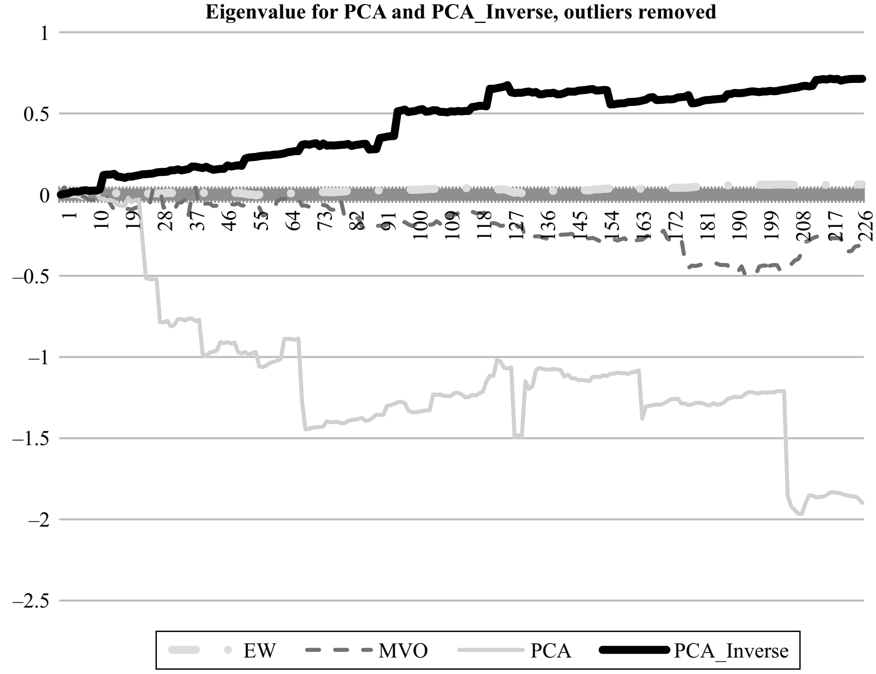
Panel A: S&P 500 Strategy, Monthly Reallocations: Keep the Top One Eigenvalue in the Correlation Matrix (bottom one eigenvalue in the correlation matrix inverse)

Portfolio Strategy Performance Comparison 1998–2017: Top 1 Eigenvalue for PCA and PCA_Inverse. Left axis: cum. return



Panel B: S&P 500 Strategy, Monthly Reallocations: Keep the Top One Eigenvalue in the Correlation Matrix (bottom one eigenvalue in the correlation matrix inverse), Outliers Removed

Portfolio Strategies Cumulative Return Comparison 1998–2017: Top 1 Eigenvalue for PCA and PCA_Inverse, outliers removed

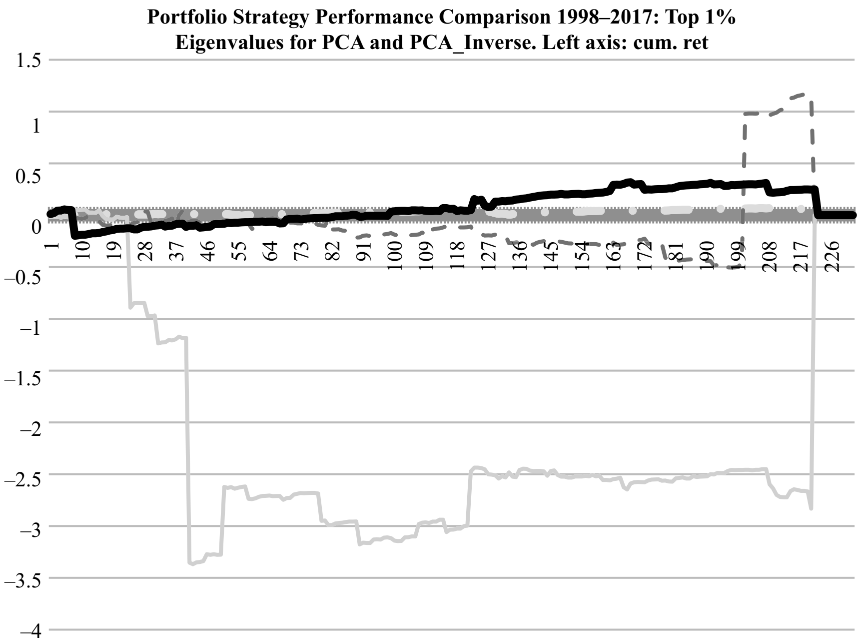


(continued)

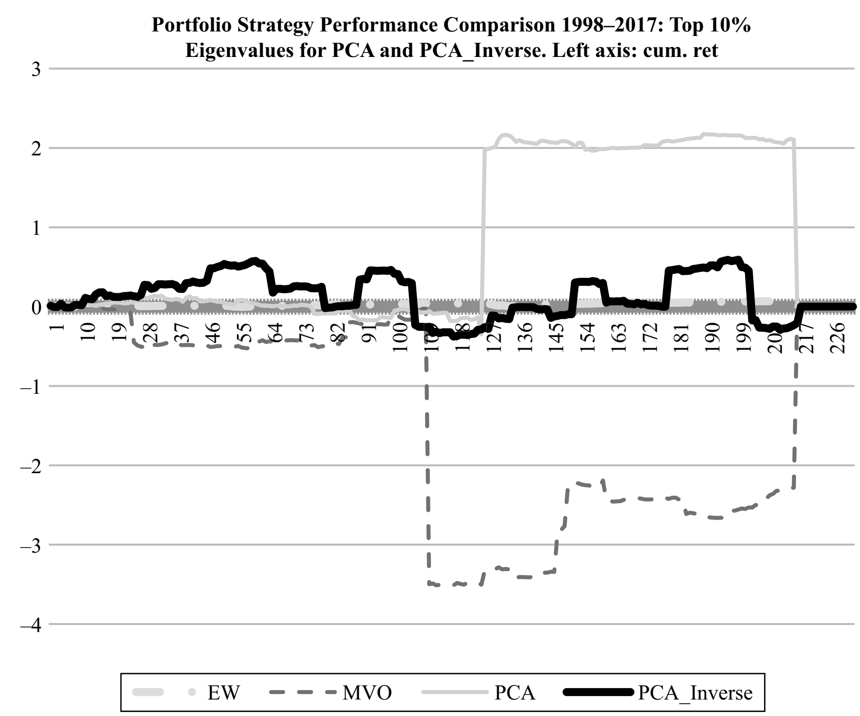
Exhibit 9 *(continued)*

Portfolio Strategy Performance Comparison, S&P 500, 1998–2017

Panel C: S&P 500 Strategy, Monthly Reallocations: Keep the Top 1% of Eigenvalues in the Correlation Matrix (bottom 1% of eigenvalues in the correlation matrix inverse)



Panel D: S&P 500 Gross Cumulative Annualized Returns of EW, Standard MVO, Inverse Correlation Largest Eigenvalue Decile (inverse largest), and Inverse Correlation Smallest Eigenvalue Decile (inverse smallest) Portfolios



Notes: Gross cumulative annualized returns of EW, standard MVO, inverse correlation largest eigenvalue deciles (inverse largest), and inverse correlation smallest eigenvalue decile (inverse smallest) portfolios.

EXHIBIT 10

Sharpe Ratios on Strategy Performance, S&P 500, 1998–2017, Monthly Reallocation

	EW	MVO	PCA	PCA_Inverse
10% Eigenvalues	0.4398529652	0.1660338977	0.2175831572	–0.05572290167
1% Eigenvalues	0.4398529652	0.1660338977	–0.2620669154	0.1854018356
1 Eigenvalue, with outliers	0.4398529652	0.1660338977	0.1121639833	0.2838331832
1 Eigenvalue, outliers removed	0.4398529652	–0.1695154284	–0.3799479568	0.6117174285

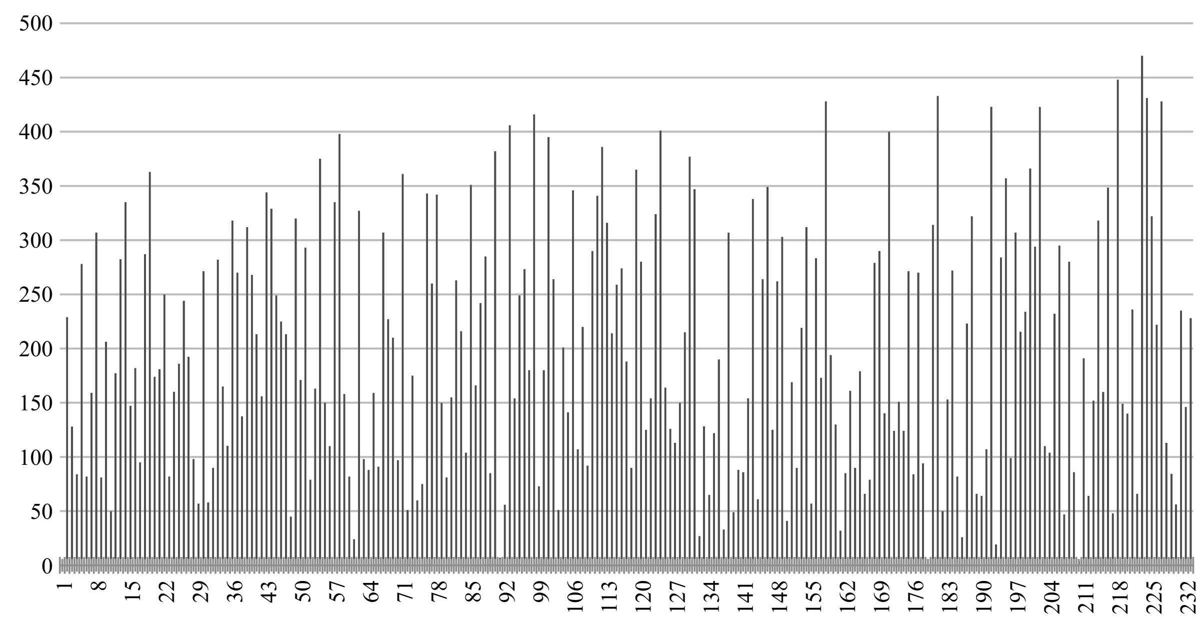
EXHIBIT 11

Average Monthly Returns per Strategy, S&P 500, 1998–2017, Monthly Reallocation

	EW	MVO	PCA	PCA_Inverse
10% Eigenvalues	0.0002994329004	0.004832313043	0.01287782073	–0.001001351801
1% Eigenvalues	0.0002994329004	0.004832313043	–0.01278558696	0.001204575758
1 Eigenvalue, with outliers	0.0002994329004	0.004832313043	0.007869265217	0.01649431169
1 Eigenvalue, outliers removed	0.0002994329004	–0.001353022026	–0.008363651982	0.003141938326

EXHIBIT 12

Number of Securities Selected Each Month from the S&P 500 by PCA_Inverse Method, 1998–2017



Notes: Using only the top eigenvalue of the inverse. It is common for the algorithm to deliver a single-digit number of names under this portfolio construction.

CONCLUSIONS

In this article, I demonstrate that the inverse of the correlation matrix is inherently more sensitive to perturbations than the correlation matrix itself, affecting the Markowitz portfolio allocation strategies. To harness the power of big data analytics to capitalize on this information content, I propose a big data refinement to portfolio selection: applying spectral decomposition to the inverse of the correlation matrix, instead of to the correlation matrix. The proposed methodology is tested on the S&P 500 Index and random subportfolios of the

EXHIBIT 13

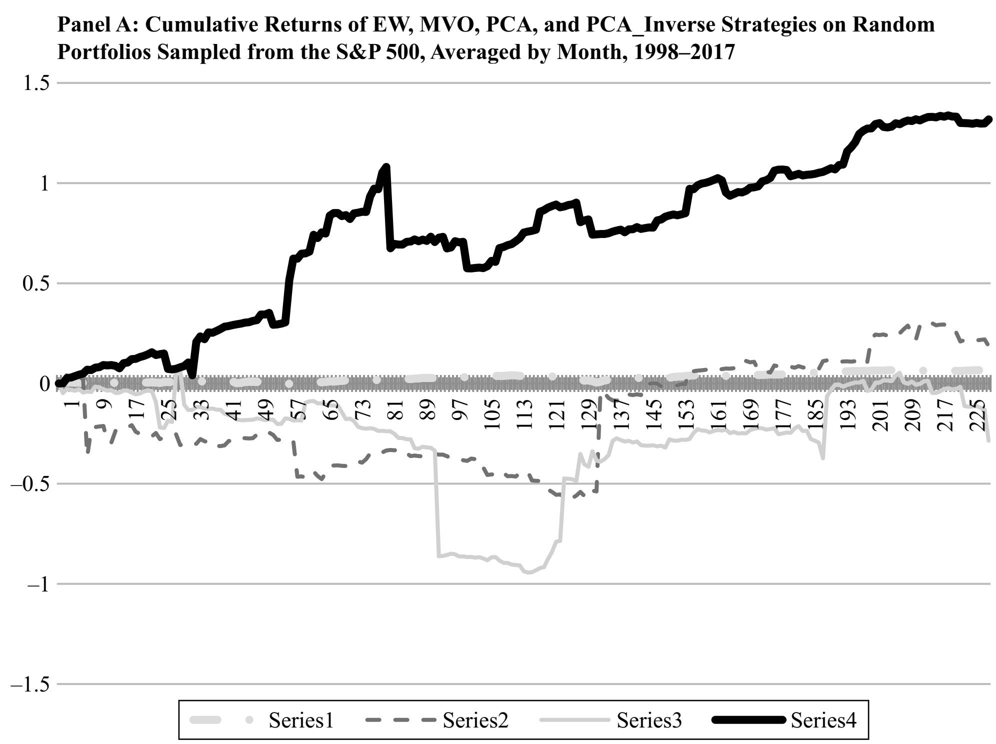
Mean and Standard Deviation (in parentheses) for the Number of Equities from the S&P 500 with Absolute Values of Weights Exceeding 1% or 2% of the Entire Portfolio Selected Monthly by Vanilla MVO, PCA, and PCA_Inverse Methods for Different Eigenvalue Cutoffs

		MVO	PCA	PCA_Inverse
Top 1 Eigenvalue	|weight| > 1%	375.49 (82.16)	375.15 (82.94)	192.44 (117.07)
	|weight| > 2%	343.57 (88.33)	343.40 (89.77)	112.64 (116.94)
Top 1% of Eigenvalues	|weight| > 1%	375.49 (82.16)	376.49 (84.30)	190.90 (115.95)
	|weight| > 2%	343.57 (88.33)	346.27 (91.65)	109.98 (115.37)
Top 10% of Eigenvalues	|weight| > 1%	375.69 (84.82)	377.83 (85.41)	372.13 (82.73)
	|weight| > 2%	341.54 (93.03)	343.97 (93.17)	335.50 (89.24)

Note: Data: 1998–2017, monthly portfolio rebalancing.

EXHIBIT 14

Performance of Portfolio Randomly Selected from the S&P 500 Constituents, 1998–2017

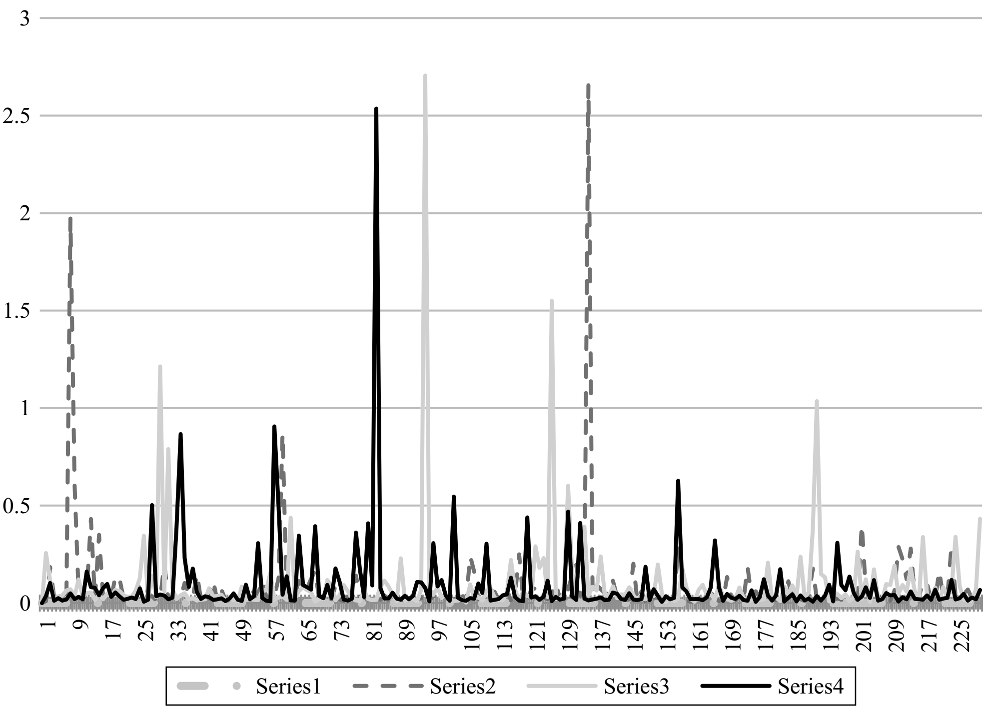


(continued)

EXHIBIT 14 *(continued)*

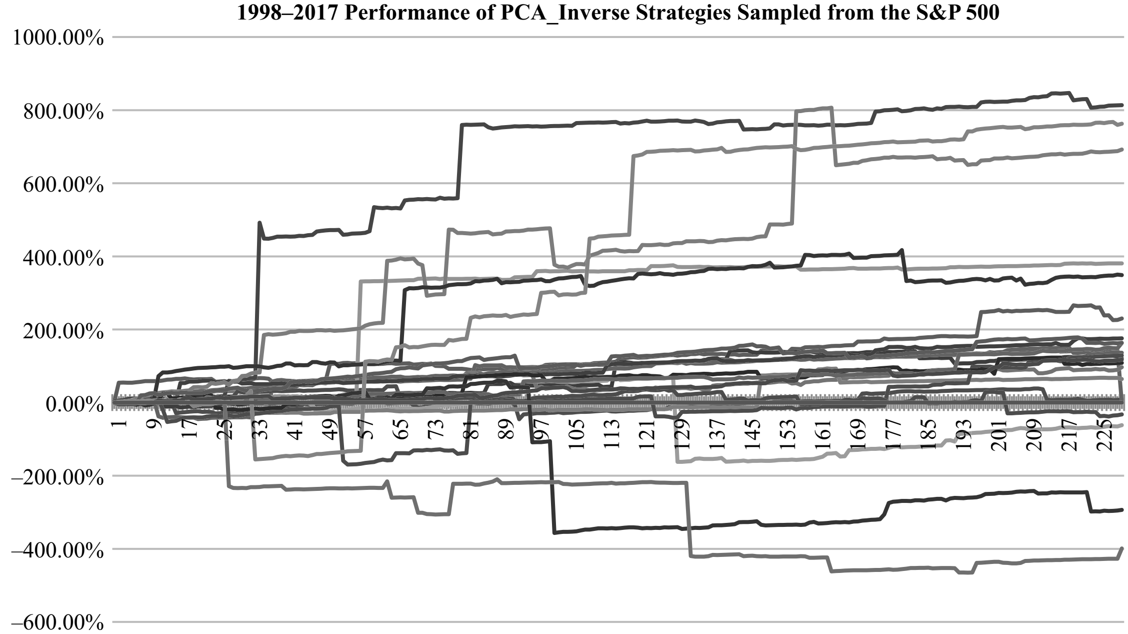
Performance of Portfolio Randomly Selected from the S&P 500 Constituents, 1998–2017

Panel B: Standard Deviation of Monthly Returns of EW, MVO, PCA, and PCA_Inverse Strategies on Random Portfolios Sampled from the S&P 500, by Month, 1998–2017



Panel C: Cumulative Return Paths of the 30 Portfolios Randomly Drawn from the S&P 500, 1998–2017

1998–2017 Performance of PCA_Inverse Strategies Sampled from the S&P 500



S&P 500 from 1998 through 2017. Out of sample, the methodology consistently outperforms other common methods, such as EW portfolio allocation, plain MVO, and previously suggested big data portfolio optimization methodologies.

ACKNOWLEDGMENTS

I am grateful to Marco Avellaneda, Frank Fabozzi, Sasha Stoikov, and Victoria Averbukh; participants of the 2018 Sixth Annual Big Data Finance Conference; and my students at Cornell University for suggestions and advice.

REFERENCES

Aldridge, I., and S. Krawciw. *Real-Time Risk: What Investors Should Know about Fintech, High-Frequency Trading and Flash Crashes*. Hoboken, NJ: Wiley, 2017.

Allez, R., and J. P. Bouchaud. 2012. "Eigenvector Dynamics: General Theory and Some Applications." *Physical Review E* 86: 46202.

Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold. "Volatility and Correlation Forecasting." In *Handbook of Economic Forecasting*, edited by G. Elliot, C.W.J. Granger, and A. Timmermann, pp. 778–878. Amsterdam: North-Holland; 2006.

Ang, A., and G. Bekaert. 2002. "International Asset Allocation with Regime Shifts." *Review of Financial Studies* 15 (4): 1137–1187.

Avellaneda, M., and J. H. Lee. 2010. "Statistical Arbitrage in the US Equities Market." *Quantitative Finance* 10 (7): 761–782.

Barry C. 1974. "Portfolio Analysis under Uncertain Means, Variances, and Covariances." *The Journal of Finance* 29 (2): 515–522.

Bauer, F. L., and C. T. Fike. 1960. "Norms and Exclusion Theorems." *Numerische Matematik* 2: 137–141.

Bawa, V. S., S. J. Brown, and R. W. Klein. *Estimation Risk and Optimal Portfolio Choice*. Amsterdam: North-Holland; 1979.

Beltrami, E. 1873. "Sulle funzioni bilineari." *Giornale di Matematiche ad Uso degli Studenti Delle Universita* 11: 98–106.

Bertsimas, D., and A. W. Lo. 1998. "Optimal Control of Execution Costs." *Journal of Financial Markets* 1 (1): 1–50.

Best, M. J., and R. R. Grauer. 1992. "Positively Weighted Minimum-Variance Portfolios and the Structure of Asset Expected Returns." *Journal of Financial and Quantitative Analysis* 27 (4): 513–537.

Black, F., and R. Litterman. 1993. "Global Portfolio Optimization." *Financial Analysts Journal* 48: 28–43.

Bollerslev, T. 1986. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31 (3): 307–327.

Boyle, P. 2014. "Positive Weights on the Efficient Frontier." *North American Actuarial Journal* 18 (4): 462–477.

Brandt, M. W., A. Goyal, P. Santa-Clara, and J. R. Stroud. 2005. "A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning about Return Predictability." *Review of Financial Studies* 18: 831–873.

Carrasco, M., J. P. Florens, and E. Renault. 2007. "Linear Inverse Problems in Structural Econometrics Estimation Based on Spectral Decomposition and Regularization." *Handbook of Econometrics* 6: 5633–5751.

Carrasco, M., and N. Noumon. 2011. "Optimal Portfolio Selection Using Regularization." Discussion paper.

Chan, L. K. C., J. Karceski, and J. Lakonishok, 1999. "Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model." *Review of Financial Studies* 12 (5): 937–974.

Chopra, V., and W. Ziemba. 1993. "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice." *The Journal of Portfolio Management* 19: 6–12.

Cont, R. 2001. "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues." *Quantitative Finance Volume* 1: 223–236.

Davis, R. A., and T. Mikosch. 1998. "Limit Theory for the Sample ACF of Stationary Process with Heavy Tails with Applications to ARCH." *The Annals of Statistics* 26: 2049–2080.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009. "Optimal versus Naive Diversification: How Inefficient Is the $1/N$ Portfolio Strategy?" *Review of Financial Studies* 22: 1915–1953.

Engle, R. F. 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50 (4): 987–1008.

Fan, J., Y. Liao, and M. Mincheva. 2013. "Large Covariance Estimation by Thresholding Principal Orthogonal Complements." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 75 (4): 603–680.

Frost, P. A., and J. E. Savarino. 1988. "For Better Performance: Constrain Portfolio Weights." *The Journal of Portfolio Management* 15 (1): 29–34.

Garlappi, L., and G. Skoulakis. 2008. "A State Variable Decomposition Approach for Solving Portfolio Choice Problems." *Review of Economic Studies* 23: 3346–3400.

Garlappi, L., R. Uppal, and T. Wang. 2007. "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach." *The Review of Financial Studies* 20 (1): 41–81.

Gerschgorin, S. 1931. "Über die Abgrenzung der Eigenwerte einer Matrix." Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk 7, 749–754.

Goldfarb, D., and G. Iyengar. 2003. "Robust Portfolio Selection Problems." *Mathematics of Operations Research* 28 (1): 1–38.

Gourieroux, C. *ARCH Models and Financial Applications*. Berlin: Springer, 1997.

Hakansson, N. H. 1971. "Capital Growth and the Mean–Variance Approach to Portfolio Selection." *Journal of Financial and Quantitative Analysis* 6: 517–557.

Jagannathan, R., and T. Ma. 2003. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *The Journal of Finance* 58 (4): 1651–1683.

Jegadeesh, N., and S. Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *The Journal of Finance* 48 (1): 65–91.

Jobson, J. D., and B. Korkie. 1980. "Estimation for Markowitz Efficient Portfolios." *Journal of the American Statistical Association* 75: 544–554.

Jobson, J. D., R. Korkie, and V. Ratti. 1979. "Improved Estimation for Markowitz Portfolios Using James-Stein Type Estimators." *Proceedings of the American Statistical Association* 41: 279–292.

Jorion, P. 1985. "International Portfolio Diversification with Estimation Risk." *The Journal of Business* 58 (3): 259–278.

——. 1986. "Bayes-Stein Estimation for Portfolio Analysis." *The Journal of Financial and Quantitative Analysis* 21 (3): 279–292.

Kroll, Y., H. Levy, and H. M. Markowitz. 1984. "Mean–Variance versus Direct Utility Maximization." *The Journal of Finance* 39: 47–75.

Ledoit, O., and M. Wolf. 2004a. "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices." *Journal of Multivariate Analysis* 88 (2): 365–411.

——. 2004b. "Honey, I Shrunk the Sample Covariance Matrix." *The Journal of Portfolio Management* 30 (4): 110–119.

——. 2017. "Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks." *The Review of Financial Studies* 30 (12): 4349–4388.

Lipschutz, S. *Schaum's Outline: Linear Algebra*. New York: McGraw-Hill; 1991.

Liu, H. 2004. "Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets." *The Journal of Finance* 59 (1): 289–338.

Loistl, O. 1976. "The Erroneous Approximation of Expected Utility by Means of a Taylor's Series Expansion: Analytic and Computational Results." *The American Economic Review* 66: 904–910.

Lynch, A. W., and S. Tan. 2008. "Multiple Risky Assets, Transaction Costs and Return Predictability: Allocation Rules and Implications for US Investors." *Journal of Financial and Quantitative Analysis* 45 (4): 1015.

MacKinlay, A. C., and L. Pastor. 2000. "Asset Pricing Models: Implications For Expected Returns and Portfolio Selection." *Review of Financial Studies* 13 (4): 883–916.

Markowitz, H. M. 1952. "Portfolio Selection." *The Journal of Finance*, 7 (1): 77–91.

—— 1991. "Foundations of Portfolio Theory." *The Journal of Finance* 46: 469–477.

Mei, X., V. DeMiguel, and F. Nogales. 2016. "Multiperiod Portfolio Optimization with Multiple Risky Assets and General Transaction Costs." *Journal of Banking and Finance* 69: 108–120.

Meucci, A. "Managing Diversification." *Risk,* May 1, 2009.

Muthuraman, K., and S. Kumar. 2006. "Multidimensional Portfolio Optimization with Proportional Transaction Costs." *Mathematical Finance* 16 (2): 301–335.

Partovi, M. H., and M. Caputo. 2004. "Principal Portfolios: Recasting the Efficient Frontier." *Economics Bulletin* 7 (3): 1–10.

Pastor, L. 2000. "Portfolio Selection and Asset Pricing Models." *The Journal of Finance* 55 (1): 179–223.

Pastor, L., and R. Stambaugh, 2000. "Comparing Asset Pricing Models: An Investment Perspective." *Journal of Financial Economics* 56 (3): 335–381.

Pulley, L. B. 1981. "A General Mean–Variance Approximation to Expected Utility for Short Holding Periods." *Journal of Financial and Quantitative Analysis* 16: 361–373.

——. 1983. "Mean–Variance Approximations to Expected Logarithmic Utility." *Operations Research* 31: 685–696.

Robinson, P. D., and A. J. Wathen. 1992. "Variational Bounds on the Entries of the Inverse of a Matrix." *IMA Journal of Numerical Analysis* 12: 463–486.

Samuelson, P. A., 1970. "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Moments." *Review of Economic Studies* 37: 537–542.

Steele, A. 1995. "On the Eigen Structure of the Mean Variance Efficient Set." *Journal of Business Finance & Accounting* 22 (2): 245–255.

Stewart, G. W. 1993. "On the Early History of the Singular Value Decomposition." *SIAM Review* 35 (4): 551–566.