Received 1 June 2018
Revised 25 March 2019
5 May 2019
Accepted 9 May 2019

On sparsity of eigenportfolios to reduce transaction cost

Anqi Xiong
Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark College of Engineering, Newark, New Jersey, USA, and
Ali N. Akansu
Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, New Jersey, USA

Abstract

Purpose – Transaction cost becomes significant when one holds many securities in a large portfolio where capital allocations are frequently rebalanced due to variations in non-stationary statistical characteristics of the asset returns. The purpose of this paper is to employ a sparsing method to sparse the eigenportfolios, so that the transaction cost can be reduced and without any loss of its performance.

Design/methodology/approach – In this paper, the authors have designed pdf-optimized mid-tread Lloyd-Max quantizers based on the distribution of each eigenportfolio, and then employed them to sparse the eigenportfolios, so those small size orders may usually be ignored (sparsed), as the result, the trading costs have been reduced.

Findings – The authors find that the sparsing technique addressed in this paper is methodic, easy to implement for large size portfolios and it offers significant reduction in transaction cost without any loss of performance.

Originality/value – In this paper, the authors investigated the performance the sparsed eigenportfolios of stock returns in S&P500 Index. It is shown that the sparsing method is simple to implement and it provides high levels of sparsity without causing PNL loss. Therefore, transaction cost of managing a large size portfolio is reduced by employing such an efficient sparsity method.

Keywords Transaction cost, Principal component analysis (PCA), Sparse matrix, Eigenportfolio, Karhunen-Loeve transform (KLT), Midtread (zero-zone) pdf-optimized Lloyd-Max quantizer, Eigen decomposition

Paper type Research paper

1. Introduction

In order to reduce volatility, large size portfolios with built-in diversity are commonly used in practice. On the other hand, the portfolio maintenance (re-balancing) becomes more costly when portfolio size is large, e.g. a few hundred asset portfolio. Therefore, calculated small adjustments of some asset positions are judiciously ignored in the implementation during the periodic re-balancing process by employing a method to sparse large size portfolios. A sparsing technique for Markowitz (mean-variance) portfolio (Markowitz, 1959) was proposed in Brodie *et al.* (2008), a penalty (regularization term) which is proportional by employing L1-norm based lasso regression (Tibshirani, 1996). In this paper, we propose a method to design sparse eigenportfolios and present its merit by using market data.

Eigendecomposition, also called spectral decomposition, principal component analysis or Karhunen-Loeve Transform (KLT), is the factorization of a diagonalizable matrix in terms of its eigenvalues and eigenvectors. It has been a popular method for multivariate data analysis and dimension reduction problems where commonly the matrix of interest is a correlation or

Journal of Capital Markets Studies
Vol. 3 No. 1, 2019
pp. 82-90
Emerald Publishing Limited
2514-4774
DOI 10.1108/JCMS-06-2018-0024

covariance matrix of a random vector process (Hotelling, 1933; Karhunen, 1947; Loeve, 1955; Wilkinson, 1965; Akansu and Haddad, 1992). The proper interpretation of eigenvectors (principal components) and eigenvalues (weight of principal components) for the given covariance (or correlation) matrix and application is a major aspect of eigenanalysis. In finance, this analysis method is employed to design a set of eigenportfolios for a group of stocks in a basket where eigenportfolio returns are perfectly pairwise decorrelated. Moreover, they are statistically independent with the assumption of jointly Gaussian stock returns. Small components of an eigenvector increase the transaction cost in generation and maintenance (rebalancing) of the relevant eigenportfolio due to the tracking of statistical variations in time (Akansu and Torun, 2015; Torun *et al.*, 2011). The use of eigenvector components (loading coefficients) in such an application makes sparsity (cardinality reduction of a vector space) an important consideration. In contrast, the uneven distribution of signal energy among eigenvectors' spectra (as reflected in eigenvalues or eigencoefficient variances) leads to dimension reduction that is inherent in subspace methods. This characteristic is utilized in image and video compression through transform coding techniques (Jayant and Noll, 1984; Akansu and Haddad, 1992).

The dimension reduction and sparsity of basis functions (eigenvectors) are important properties of orthonormal transforms to emphasize frequency and time domain specifics of a signal vector. These built-in time and frequency domain properties of signals as emphasized in a subspace are independently utilized in most applications. The simultaneous interpretation of Eigen coefficients and eigenvector components in subspace methods leads to the time-frequency representation of signal energy. We use it in this paper.

This work is a continuation of the subspace sparsing framework proposed in Yilmaz and Akansu (2015). It is based on the rate-distortion theory and employs zero-zone (mid-tread) pdf-optimized (Lloyd-Max) quantizer created for the histogram of an eigenvector or eigenmatrix and the desired level of sparsity in the subspace (Max, 1960; Lloyd, 1982; Akansu and Haddad, 1992). We focus on sparsing the eigenportfolios of stocks in S&P500 index by using this method and their resulting performance.

The method to sparse subspace is presented in Section 2. Section 3 focuses on sparsity of eigenportfolios. The quantization of eigenvector components or eigenmatrix elements of such eigensubspace is described. The impact of sparsity on PNL performance of the eigenportfolios is displayed. The conclusions are presented in the following section of the paper.

2. Sparsity in subspace methods

The energy compaction that is achieved through the unevenness of transform coefficient variances, and their pairwise correlations are the performance metrics derived from the rate-distortion theory to evaluate orthogonal sets (orthogonal subspace methods). The energy compaction measure emphasizes the spectral (frequency domain) features of a subspace representation and the foundation of transform coding, that is the industry standard for image and video compression standards (Jayant and Noll, 1984; Akansu and Haddad, 1992). In contrast, sparse representation aims to replace insignificant components of the basis vectors, cardinality reduction, that define an orthogonal subspace. Hence, it highlights the signal domain (time domain) characteristics of subspace representation. We investigate both the time and the frequency domain sparsities of subspace methods where both the explained variance and sparsity are quantified.

Herein, we revisit the mathematical definitions of orthogonal subspace representation. Let x be an $N \times 1$ input signal vector:

$$\mathrm{x} = [x_0, x_1, \ldots, x_{N-1}]^{\mathrm{T}}. \tag{1}$$

θ be an $N \times 1$ coefficient vector:

$$\theta = [\theta_0, \theta_1, \ldots, \theta_{N-1}]^{\mathrm{T}}. \tag{2}$$

And, let Φ be an $N \times N$ orthogonal transform matrix:

$$\Phi = \left[\phi_0^{\mathrm{T}}, \phi_1^{\mathrm{T}}, \ldots, \phi_{N-1}^{\mathrm{T}}\right], \tag{3}$$

where ϕ_k is an $N \times 1$ vector, and Φ matrix satisfies the orthogonality properties:

$$\Phi\Phi^{-1} = \Phi\Phi^{\mathrm{T}} = \mathrm{I}, \tag{4}$$

I is the identity matrix of size $N \times N$. Then, we define the forward transform (projection) of signal vector x onto orthogonal subspace:

$$\theta = \Phi \mathrm{x}, \tag{5}$$

and its inverse transform (signal representation by basis vectors) as:

$$\mathrm{x} = (\Phi)^{-1}\theta = \Phi^{T}\theta, \tag{6}$$

to reconstruct the original signal vector x.

Most of the sparse representation techniques reported in the literature are based on various subspace optimization methods where sparsity is imposed in the design. More recently, the quantization of basis vector components of a subspace was experimented as a more efficient alternative to the existing sparsing approaches. The following subsections focus on the quantization of orthogonal Eigen subspace vectors and matrices that are used as the capital allocation coefficients of the resulting eigenportfolios (Akansu and Haddad, 1992; Yilmaz and Akansu, 2015).

2.1 Quantization of subspace

We focus on the quantization of forward transform matrix (vectors) in this section. Then, its quantized (sparsed) version rather than the original matrix is employed for signal representation. The motivation for such vector sparsing (cardinality reduction) is to replace insignificant vector components with zero. Hence, one may reduce computational and implementation cost of orthogonal set based subspace applications spanning from image compression to eigenportfolios.

We quantize orthogonal set (transform matrix) by using a quantizer $Q_{\Phi}\{\cdot\}$ (Yilmaz and Akansu, 2015):

$$\hat{\Phi} = Q_{\Phi}\{\Phi\}, \tag{7}$$

such that the error matrix of the quantized set is written as:

$$\tilde{\Phi} = \Phi - \hat{\Phi}. \tag{8}$$

It is noted that this quantization compromises the orthogonality property of the set. The levels of non-orthogonality and the sparsity are coupled in this problem. Then the transform coefficients for the quantized set are calculated as:

$$\hat{\theta}_{FT} = \hat{\Phi}\mathrm{x}. \tag{9}$$

Therefore, the quantization error embedded in transform coefficients due to the use of $\hat{\Phi}$ rather than the (original) matrix Φ is shown as:

$$\tilde{\theta}_{FT} = \theta - \hat{\theta}_{FT} = \Phi\mathrm{x} - \hat{\Phi}\mathrm{x} = \left(\Phi - \hat{\Phi}\right)\mathrm{x} = \tilde{\Phi}\mathrm{x}. \tag{10}$$

Then, the reconstructed signal by using the original inverse transform matrix $\Phi^{-1} = \Phi^{\mathrm{T}}$ is obtained as:

$$\hat{\mathrm{x}}_{FT} = \Phi^{T}\hat{\theta}_{FT}. \tag{11}$$

The reconstruction error is given in the expression:

$$\tilde{\mathrm{x}}_{FT} = \mathrm{x} - \hat{\mathrm{x}}_{FT} = \Phi^{\mathrm{T}}\theta - \Phi^{\mathrm{T}}\hat{\theta}_{FT} = \Phi^{\mathrm{T}}\left(\theta - \hat{\theta}_{FT}\right) = \Phi^{\mathrm{T}}\tilde{\theta}_{FT}. \tag{12}$$

The MSE of coefficients is equal to the MSE of reconstructed signal due to the orthogonality of the original set as:

$$E\{\tilde{\mathrm{x}}_{FT}^{\mathrm{T}}\tilde{\mathrm{x}}_{FT}\} = E\left\{\left(\Phi^{\mathrm{T}}\tilde{\theta}_{FT}\right)^{\mathrm{T}}\left(\Phi^{\mathrm{T}}\tilde{\theta}_{FT}\right)\right\} = E\left\{\tilde{\theta}_{FT}^{\mathrm{T}}\underbrace{\Phi\Phi^{T}}_{I}\tilde{\theta}_{FT}\right\} = E\left\{\tilde{\theta}_{FT}^{\mathrm{T}}\tilde{\theta}_{FT}\right\}. \tag{13}$$

In summary:

$$E\{\tilde{\mathrm{x}}_{FT}^{\mathrm{T}}\tilde{\mathrm{x}}_{FT}\} = E\left\{\tilde{\theta}_{FT}^{\mathrm{T}}\tilde{\theta}_{FT}\right\} = E\left\{\mathrm{x}^{T}\tilde{\Phi}^{\mathrm{T}}\tilde{\Phi}\mathrm{x}\right\}. \tag{14}$$

One can easily exchange the roles of the forward and the inverse transform matrices for the applications like eigenportfolio design where the sparsity of the representation (inverse) set is desired.

2.2 pdf-optimized quantizer

We use Lloyd-Max quantizer that minimizes the quantization error in the mean square sense. The pdf optimized quantizer is calculated iteratively as described in Max (1960) and Lloyd (1982). The random input x has a pdf $f(x)$ with zero-mean and unit variance, the end values of N-level quantizer are x_k and x_{k+1}, and y_k represents all numbers fall into the kth interval (bin) $[x_k, x_{k+1}]$, where $k = 1, 2, ..., N$, and $x_1 = -\infty$ and $x_{N+1} = \infty$. The quantization error in MSE for such quantizer is calculated as:

$$\sigma_q^2 = \sum_{k=1}^{N}\int_{x_k}^{x_{k+1}}(x - y_k)^2 f(x)dx. \tag{15}$$

The Lloyd-Max quantizer design algorithm updates the intervals $[x_k, x_{k+1}]$ and y_k iteratively, by satisfying two conditions (Max, 1960; Lloyd, 1982):

$$\begin{aligned}
&\frac{\partial\sigma_q^2}{\partial x_k} = 0; k = 2, 3, ..., N,\\
&\frac{\partial\sigma_q^2}{\partial y_k} = 0; k = 1, 2, 3, ..., N,
\end{aligned} \tag{16}$$

which also implies that:

$$\begin{aligned}
x_k &= \frac{1}{2}(y_k + y_{k-1}); k = 2, 3, ..., N,\\
y_k &= \frac{\int_{x_k}^{x_{k+1}} xf(x)dx}{\int_{x_k}^{x_{k+1}} f(x)dx}; k = 1, 2, 3, ..., N.
\end{aligned} \tag{17}$$

Note that the noise variances of all bins are the same in a pdf-optimized quantizer.

The rate-distortion theory based sparsity method was detailed in Yilmaz and Akansu (2015), and it is used to generate sparse eigenportfolios in this study. We employ the midtread (zero-zone) quantizer type to quantize each basis function (components of each vector) or the entire basis set of a transform to achieve a sparse representation. It is noted that only the center bin (zero-zone) of the mid-tread quantizer around zero is used in sparsity applications. The size of this zero-zone is adjusted to achieve the desired level of sparsity.

3. Sparsity of eigenportfolios

The calculated eigenvectors of empirical correlation matrix may have components (capital allocation coefficients) with small values. The maintenance of portfolios with large number of assets becomes burdensome and costly. It is a common practice to avoid investing in assets of a portfolio with small capital allocations in the overall investment (Brodie *et al.*, 2008; Akansu *et al.*, 2016). A design framework to sparse portfolios was proposed in Yilmaz and Akansu (2015), and we utilize it in this study.

The eigenportfolios are generated based on eigenvectors of empirical correlation matrix obtained from past returns of a pre-selected basket of assets for a given time window. We used the returns of stocks in S&P500 index for the period of December 1, 2015 to December 1, 2017 for various sparsity levels in order to validate the merit of the sparsity method.

There are 492 tickers of S&P500 index that continue to exist during that time period. Therefore, we used their end of day (EOD) simple returns, $r(n)$, to calculate the empirical correlation matrix at time n, $R_E(n)$, as follows (Akansu and Torun, 2012, 2015):

$$\mathbf{r}(n) = [r_k(n)]; k = 1, 2, ..., 492, \tag{18}$$

and, the resulting empirical correlation matrix $R_E(n)$ is calculated as:

$$\mathbf{R}_E(n) = \left[E\{\mathbf{r}(n)\mathbf{r}^{\mathrm{T}}(n)\}\right] = [R_{k,l}(n)]$$
$$= \begin{bmatrix} R_{1,1}(n) & R_{1,2}(n) & \dots & R_{1,492}(n) \\ R_{2,1}(n) & R_{2,2}(n) & \dots & R_{2,492}(n) \\ \vdots & \vdots & \ddots & \vdots \\ R_{492,1}(n) & R_{492,2}(n) & \dots & R_{492,492}(n) \end{bmatrix}, \tag{19}$$

where its elements are obtained as:

$$R_{k,l}(n) = E\{r_k(n)r_l(n)\} = \frac{1}{W}\sum_{m=0}^{W-1} r_k(n-m)r_l(n-m), \tag{20}$$

and, $R_{k,l}(n)$ represents the pairwise empirical correlation of the kth and the lth stocks at time n. We used the time window of $W = 60$ days in this study. Note that the returns are normalized to zero mean and unit variance, and $R_E(n)$ is real, symmetric and positive definite matrix. Then, the eigenmatrix A_{KLT} of $R_E(n)$ that satisfies the eigenmatrix decomposition property as given below is obtained (Akansu and Haddad, 1992; Akansu and Torun, 2012; Yilmaz and Akansu, 2015):

$$R_E(n) = A^{\mathrm{T}}_{KLT}(n)\Lambda(n)A_{KLT}(n) = \sum_{k=1}^{K} = \lambda_k(n)\phi_k(n)\phi_k^{\mathrm{T}}(n), \tag{21}$$

where $\{\lambda_k(n), \phi_k(n)\}$ are eigenvalue-eigenvector (eigenportfolio) pairs.

Now, we focus on the elements of A_{KLT} with small values. The histogram of the elements of an eigenmatrix A_{KLT} is displayed in Figure 1.

Then, design a mid-tread (zero-zone) pdf-optimized quantizer for this histogram to replace the small valued elements of A_{KLT} by zero. We adjust the zero-zone of the quantizer by simultaneously adding the neighboring pairs of intervals on both sides to achieve the desired sparsity level. We only use the zero-zone of the quantizer for this application.

Each eigenvector may have different histogram, and one can design a separate quantizer for every one of them in particular for portfolios with large number of assets.

We also investigate the explained variance (eigenvalue) of the eigenvectors $\phi_k(n)$ in order to identify significant ones and evaluate their eigenportfolio performance. The kth eigenvalue at time n, $\lambda_k(n)$, is equivalent to the variance of the kth transform coefficients, σ_k^2, and calculated as $\{\lambda_k(n) = \phi_k^T(n) R_E(n) \phi_k(n)\} \forall k$ (Akansu and Haddad, 1992). Figure 2 illustrates the cumulative explained variance as a function of the number of eigenvalues for

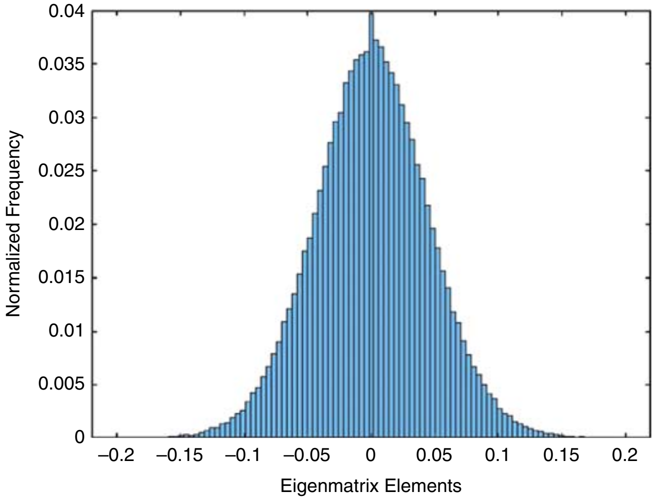

Figure 1. Normalized histogram of eigenmatrix elements for empirical correlation matrix of end of day (EOD) returns for 492 stocks in S&P 500 index with $W = 60$-day measurement window starting on December 1, 2015, ending on February 26, 2016, and widow shifts for another 252 days, starting from February 26, 2016, ending on February 24, 2017

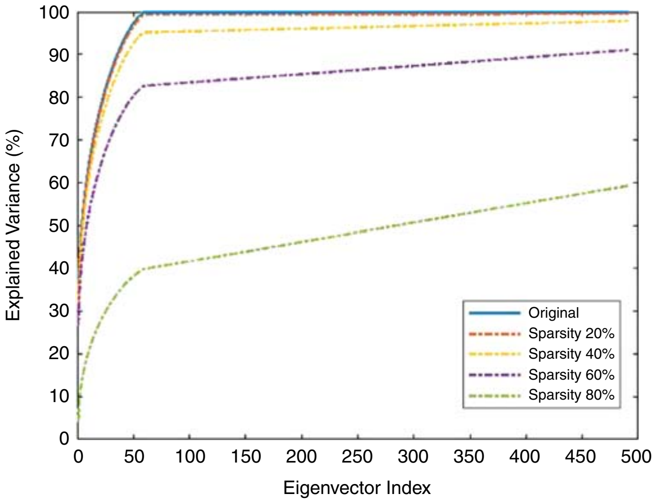

Figure 2. Cumulative explained variances as a function of the number of eigenvectors included in representation for different levels of sparsity

various sparsity levels. Similarly, the explained variances (eigenvalues) of the sparsed eigenvectors at time n, $\{\hat{\phi}_k(n)\}$, are calculated as $\{\hat{\lambda}_k(n) = \hat{\phi}_k^{T}(n)R_E(n)\hat{\phi}_k(n)\}\forall k$.

Note that first 59 out of 492 eigenvalues (eigenvectors or principle components) explained 99 percent of the total variance (of the random vector process) in this experiment. We designed a specific pdf-optimized quantizer for each one of the first 25 eigenvectors that explain 80 percent of the total variance.

The relationship between the sparsity and the resulting variance loss for these 25 eigenvectors are displayed in Figure 3. The variance loss of the kth eigenvector due to the sparsity is defined as:

$$E\{VL_k(n)\} = \frac{1}{N}\sum_{n=0}^{N-1} VL_k(n), \tag{22}$$

where:

$$\left\{VL_k(n) = \left(1-\frac{\hat{\lambda}_k(n)}{\lambda_k(n)}\right) \times 100\right\}\forall k.$$

We assume \$1 normalized investment in each eigenportfolio with long and short positions, in general, and no transaction cost is considered in these experiments. Then, the Profit and Loss (PNL) curve is calculated (Yilmaz and Akansu, 2016). It is observed from Figure 4 and Table I that the sparse eigenportfolios with significant reduction in transaction cost perform similar to the original eigenportfolios for the stocks of S&P500 index stocks.

4. Conclusions

Portfolio managers and investors desire to have smaller number of positions to open and rebalance. Therefore, one needs to develop a methodology to define a threshold where an investment allocation is deemed insignificant. This problem becomes important for very large size portfolios, (e.g. Russell 2000, VTI, VGTSX) and it can be formulated under the rate-distortion theory and a solution by using the mid-tread (zero-zone) pdf-optimized quantizer to sparse orthogonal subspaces was proposed in the literature. Note that the optimal quantizers are tuned for different portfolios with desired sparsity levels. Usually, the execution

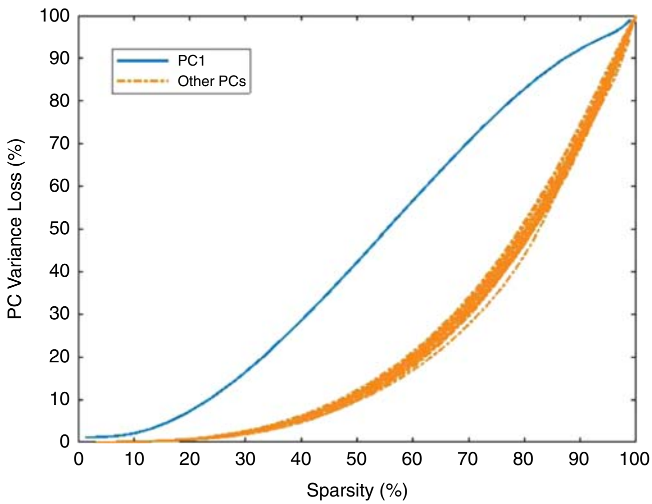

Figure 3.
Variance loss of first 25 eigenvectors (principle components) as a function of sparsity level

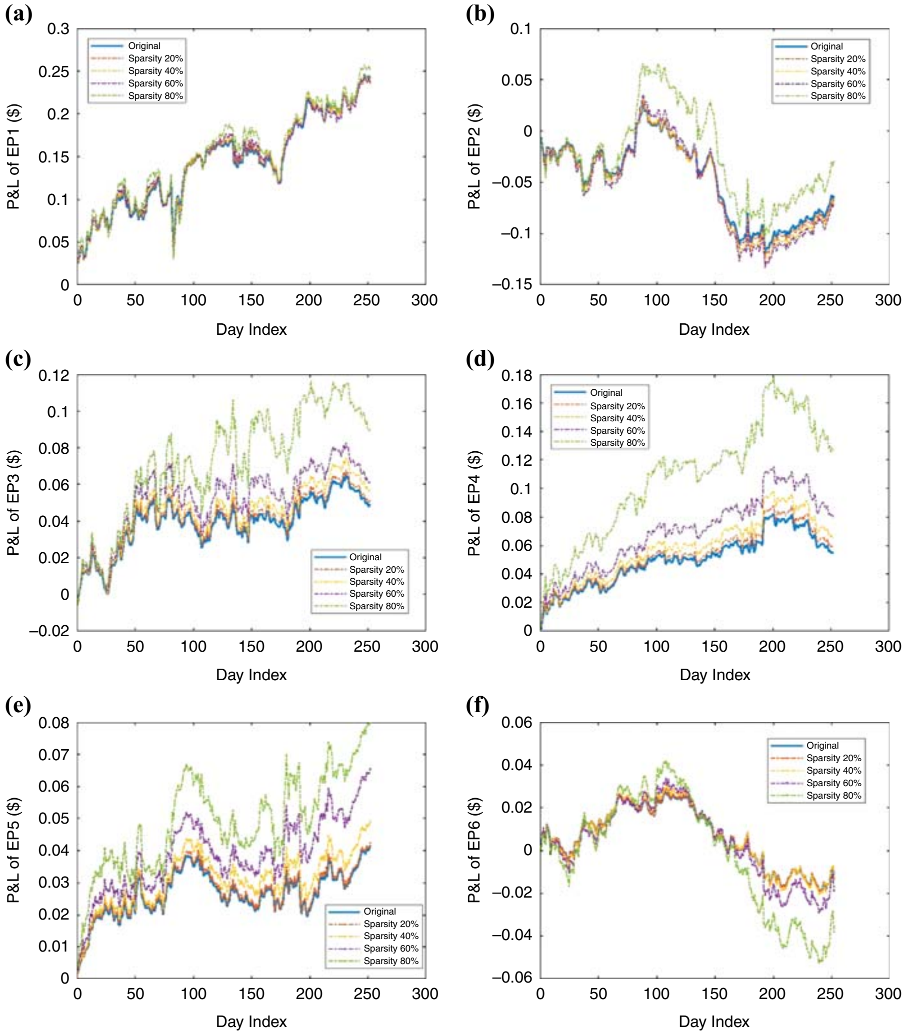

Figure 4.
PNL curves of the original and sparsed eigenportfolios for EOD returns of stocks in S&P500 index between February 26, 2016 and February 24, 2017

Table I.
Annualized Sharpe Ratios of each eigenportfolios at different level of sparsity, using S&P 500 EOD returns, starting on February 26, 2016, ending on February 24, 2017

Sharpe ratio	EP 1	EP 2	EP 3	EP 4	EP 5	EP 6
Original	1.88	−0.77	0.93	1.3	1.29	−0.42
20% Sparsity	1.81	−0.78	0.94	1.27	1.29	−0.41
40% Sparsity	1.77	−0.78	0.96	1.36	1.4	−0.38
60% Sparsity	1.66	−0.74	0.88	1.44	1.57	−0.46
80% Sparsity	1.64	−0.37	0.93	1.48	1.48	−0.78

related concerns, i.e. hard to find ticker, lot size and trading cost, are known in advance, and implemented in the adjustment of the zero-zone pdf-optimized quantizer, accordingly. In this paper, we investigated the performance of such quantizers to sparse eigenportfolios of stock returns in S&P500 Index. It is shown that the method is simple to implement and it provides high levels of sparsity without causing PNL loss. Therefore, transaction cost of maintaining a large size portfolio is reduced by employing such an efficient sparsity method.

References

Akansu, A.N. and Haddad, R.A. (1992), *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*, Academic Press, New York, NY.

Akansu, A.N. and Torun, M.U. (2012), "Toeplitz approximation to empirical correlation matrix of asset returns: a signal processing perspective", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6 No. 4, pp. 319-326.

Akansu, A.N. and Torun, M.U. (2015), *A Primer for Financial Engineering: Financial Signal Processing and Electronic Trading*, Academic Press, New York, NY.

Akansu, A.N., Kulkarni, S.R. and Malioutov, D.M. (2016), *Financial Signal Processing and Machine Learning*, Wiley-IEEE Press, available at: https://doi.org/10.1002/9781118745540

Brodie, J., Daubechies, I., De Mol, C., Giannone, D. and Loris, I. (2008), "Sparse and stable Markowitz portfolios", Working Paper Series, European Central Bank, No. 936, pp. 1-23, available at: http://ssrn.com/abstract_id=1258442

Hotelling, H. (1933), "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, Vol. 24 No. 7, pp. 498-520, available at: http://dx.doi.org/10.1037/h0070888

Jayant, N.S. and Noll, P. (1984), *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice-Hall Professional Technical Reference, Englewood Cliffs, NJ.

Karhunen, K. (1947), "Uber lineare Methoden in der Wahrscheinlichkeitsrechnung", *Annales Academiae Scientiarum Fennicae Series A. I. Mathematica – Physica*, Vol. 37, pp. 1-79.

Lloyd, S. (1982), "Least squares quantization in PCM", *IEEE Transactions on Information Theory*, Vol. 28 No. 2, pp. 129-137.

Loeve, M. (1955), *Probability Theory*, D Van Nostrand, Princeton, NJ.

Markowitz, H.M. (1959), *Portfolio Selection: Efficient Diversification of Investments*, Wiley, New York, NY.

Max, J. (1960), "Quantizing for minimum distortion", *IRE Transactions on Information Theory*, Vol. 6 No. 1, pp. 7-12.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society. Series B*, Vol. 58 No. 1, pp. 267-288.

Torun, M.U., Akansu, A.N. and Avellaneda, M. (2011), "Portfolio risk in multiple frequencies", *IEEE Signal Processing Magazine, Special Issue on Signal Processing for Financial Applications*, Vol. 28 No. 5, pp. 61-71.

Wilkinson, J.H. (1965), *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford.

Yilmaz, O. and Akansu, A.N. (2015), "Quantization of Eigen subspace for sparse representation", *IEEE Transactions on Signal Processing*, Vol. 63 No. 14, pp. 3616-3625.

Yilmaz, O. and Akansu, A.N. (2016), "Performance analysis of eigenportfolios for AR(1) process", *IEEE Annual Conference on Information Science and Systems (CISS), Princeton, NJ, March 21-23.*

Corresponding author
Anqi Xiong can be contacted at: ax2@njit.edu