FILTERATION OF MULTICOMPONENT SEISMIC WAVEFIELD DATA USING FREQUENCY SVD

Aws Al-Qaisi, W.L. Woo, and S.S. Dlay

School of Electrical, Electronics and Computer Engineering
Newcastle University, Newcastle upon Tyne
England, United Kingdom
E-mail: {aws.al-qaisi, w.l.woo, s.s.dlay}@ncl.ac.uk

ABSTRACT

This paper proposes a new statistical approach based on frequency singular value decomposition (SVD) to enhance the SNR of the noisy multicomponent seismic wavefield. Our filtering algorithm consists of three main steps: Firstly, the frequency transformed multicomponent seismic wavefield data is rearranged into one long vector containing information on all frequencies and all component interactions. Secondly, the reduced dimensional spectral covariance matrix of the long vector data is estimated by means of singular value decomposition. Finally, the separation of the primary seismic waves from the noise is achieved by projecting the dominant eigenvector that has the highest eigenvalue of the reduced dimensional covariance matrix onto the long data vector. The experimental results have shown that the proposed algorithm outperforms the conventional separation technique in terms of accuracy and complexity.

1. INTRODUCTION

In seismic exploration, a seismic wavelet is sent to the earth layers and seismic wavefield is recorded by linear arrays of multi-component antenna due to the impedance mismatches between different geological layers [1]. The recorded geophysical data is often contaminated by multiple interference and random noise. The main objective of seismic data processing is to enhance the signal to noise ratio, as well as identify different wave fields to obtain a better image of the earth reservoir [2]. In the case of single component sensors arrays, many techniques have been developed. Firstly, Radon transform can be used to enhance reflection events in a seismic wave-field [3]. Moreover, F-K transform technique transforms the seismic wavefield data into the frequency–wave number domain where the plane waves can be easily identified [4]. However, these two techniques as well as τ -p transform technique [5] in multicomponent case have poor performance in case of short arrays or in the presence of non-plane waves. The basic Singular Value Decomposiion (SVD) approach to remove noise from a seismic wavefield is given in [6]. As multicomponent sensor array technology has been dramatically used in seismic exploration. In particular filtering techniques are required for a multicomponent seismic wavefield data profile. For example, singular value decomposition of quaternion matrices [7] and SVD compound with partial Independent Component Analysis [8] are capable of separating such multicomponent seismic wavefield from contaminated noise. Nevertheless, these techniques require a pre-processing step such as wave alignment. When filtering is performed in the frequency domain, the method is called spectral-matrix filtering [9]. However, it is computationally expensive to diagonalize the whole spectral matrix**.** The proposed filtration process which is derived from MC-WBSMF algorithm [10] [11], reduces the dimension of the multicomponent spectral covariance matrix and estimates it by means of the SVD technique. As a result the algorithm is substantially less complex, so massive computational time is saved and no wave alignment is needed as second order statistics based algorithm does not take the channel phase changing into account. In this algorithm the seismic wavefield data subspace is separated from the noise by first finding the signal a subspace which is defined by the eigenvectors that related to the dominant eigenvalues of the reduced dimensional spectral matrix, and then projecting the dominant eigenvector onto the long vector data that contains information on all frequencies and all components interactions of the multicomponent seismic wave-field. This paper is arranged as follows: In section 2, multicomponent seismic model will be presented in detail. Section 3 presents the mathematical analysis of proposed algorithm. The experimental simulations will be shown in Section 4.

2. MODEL FORMULATION

Consider a uniform array shown in Figure. 1 that consists of K_x Omni-directional sensors recording the propagation of P waves as well as ground roll with $\mathrm{P} < \mathrm{K}_x$. Multicomponent seismic data that is collected during K_t samples on the d^{th} component $(d = 1,\ldots,\mathrm{K}_d)$ of the i^{th} sensor $(i = 1,\ldots\ldots,\mathrm{K}_x)$ with additive noise can be written mathematically as [8].

$$y_{di}(m) = \sum_{j=1}^{\mathrm{P}} a_{dij} x_j(m - m_{ij}) + b_{di}(m) \qquad (1)$$

Where $x_j(m)$ represent the p^{th} seismic wave, m_{ij} is the time delay that observed at the i^{th} sensor. The parameter

a_{dij} is the attenuation of the p^{th} wave on the d^{th} component of the i^{th} sensor and $b_{di}(m)$ is uncorrelated white Gaussian noise.

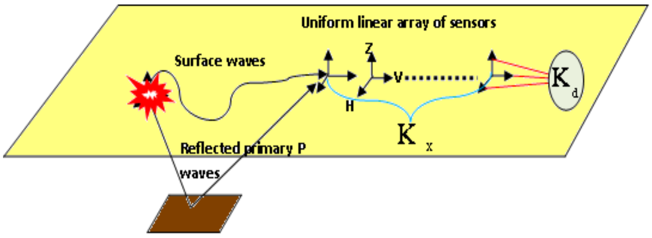

Figure 1-recording the seismic wavefield using a linear multicomponent geophones array

Therefore, the multicomponent seismic data sets that shown in Figure. 2 can be represented as $\underline{\underline{Y}}_T \in E^{K_d \times K_x \times K_t}$

$$\underline{\underline{Y}}_T = \{ y_{dit} = y_{di}(m) \mid 1 \le d \le K_d, 1 \le i \le K_x, 1 \le t \le K_t \} \qquad (2)$$

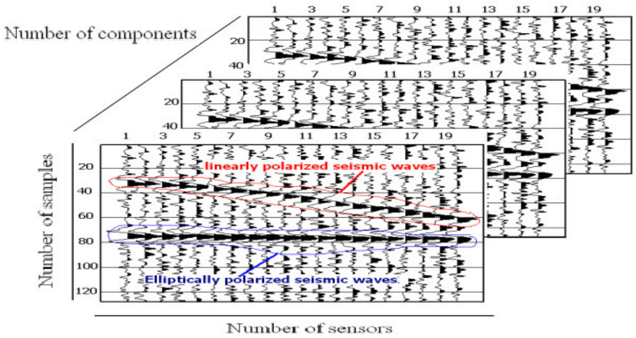

Figure 2- Multi-component wave-field seismic data set

By applying the Fourier transform to equation (2), the multicomponent seismic model can be represented as a set of instantaneous mixture of traces.

$$\underline{\underline{Y}} = FT\{\underline{\underline{Y}}_T\} \in d^{\,m} \qquad (3)$$

Where, $m = K_d K_x K_f$ and K_f is number of frequency bins. Hence, the available information in equation (3) can be rearranged in the long vector as follow [10]:

$$\mathbf{y} = [\mathbf{h}(f_1)^T .. \ \mathbf{h}(f_{K_f})^T, \mathbf{v}(f_1)^T .. \ \mathbf{v}(f_{K_f})^T, \mathbf{z}(f_1)^T .. \ \mathbf{z}(f_{K_f})^T]^T \quad (4)$$

Where, $\mathbf{h}(f_{K_f}), \mathbf{v}(f_{K_f}), \mathbf{z}(f_{K_f})$ are vectors of size (K_x), which correspond to the f^{th} frequency bin of the signals received on each of the K_x sensors, respectively, on components H, V, Z that corresponds to the polarization state of the sources P. The long vector that contains all the frequency bins on all the sensors for each component can be written as:

$$\underbrace{\mathbf{y}}_{(m\times 1)} = \underbrace{\mathbf{X}}_{(m\times P)} \underbrace{\mathbf{a}}_{(P\times 1)} + \underbrace{\mathbf{b}}_{(m\times P)} \qquad (5)$$

Where,

I. The vector $\mathbf{a} = [a_1, a_2, \ldots, a_p]^T$ corresponds to the random wave amplitude

II. $\mathbf{b}$ is the noise vector .

III. $\mathbf{X} = [\mathbf{x}_1 \ldots \mathbf{x}_p \ldots \mathbf{x}_P]$ with $p = 1, \cdots, P$

$$\text{Where, } \mathbf{x}_P = \begin{bmatrix} \underline{x}_p \\ \alpha_p e^{j\varphi_p} \underline{x}_p \\ \beta_p e^{j\psi_p} \underline{x}_p \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha_p e^{j\varphi_p} \\ \beta_p e^{j\psi_p} \end{bmatrix} \otimes \underline{x}_p \qquad (6)$$

$$\text{Since, } \underline{x}_p = \overbrace{\begin{bmatrix} \mathbf{c}(\theta_p, f_{K_1}) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{c}(\theta_p, f_{K_f}) \end{bmatrix}}^{\mathbf{C}(\theta_P, f_{K_f})} \mathbf{w}_p(f) \qquad (7)$$

$$\text{Where, } \mathbf{c}(\theta_p, f_{K_f}) = \begin{pmatrix} e^{j2\pi f_{K_f} \tau_1(\theta_p)} \\ \vdots \\ e^{j2\pi f_{K_f} \tau_i(\theta_p)} \\ \vdots \\ e^{j2\pi f_{K_f} \tau_{K_x}(\theta_p)} \end{pmatrix} \qquad (8)$$

Then the $\mathbf{x}_P$ vector can be represented using the Kronecker product as follows:

$$\mathbf{x}_P = \left(\begin{bmatrix} 1 \\ \alpha_p e^{j\varphi_p} \\ \beta_p e^{j\psi_p} \end{bmatrix} \otimes \mathbf{C}(\theta_p, f_{K_f}) \right) (1 \otimes \mathbf{w}_p(f)) \qquad (9)$$

It can be shown from the above equations that the matrix $\mathbf{X}$ contains the information about the seismic primary waves and is characterized by

i. The direction of arrival, of the seismic source (θ_p).

ii. The time of propagation between the source and the reference antenna (τ_{K_x}).

iii. The attenuation factor (α_p, β_p).

iv. The parameters (φ_p, ψ_p), which describe the change of phases between H, V, Z components.

v. The emitted seismic wavelet vector $\mathbf{w}_p(f)$.

3. PROPOSED FILTERING TECHNIQUE

All frequencies interactions between different components of directional sensors can be stated in a multi-component covariance a spectral matrix defined by

$$\mathbf{E} = E\{\mathbf{y}\mathbf{y}^{\mathrm{H}}\} \tag{10}$$

Matrix $\mathbf{E}$ has dimensions $(m \times m)$, where E is expectation operator and H is the transpose conjugate operation.Therefore; the presented multi-component covariance spectral matrix is composed of $(\mathrm{K}_d\mathrm{K}_f)^2$ blocks of dimension $\mathrm{K}_x\mathrm{K}_x$. Every block characterizes the correlation between various directional components for the received waves on all sensors at different frequencies. So the structure of the covariance matrix can be expressed in the following way [12].

$$\mathbf{E} = \begin{pmatrix} \mathbf{E}_{H,H} & \mathbf{E}_{H,V} & \mathbf{E}_{H,Z} \\ \mathbf{E}_{V,H} & \mathbf{E}_{V,V} & \mathbf{E}_{V,Z} \\ \mathbf{E}_{Z,H} & \mathbf{E}_{Z,V} & \mathbf{E}_{Z,Z} \end{pmatrix} \tag{11}$$

The component blocks $\mathbf{E}_{H,H}, \mathbf{E}_{V,V}, \mathbf{E}_{Z,Z}$ represent the correlation between each component with it self. With the aim of precise sources de-correlation from noise, jointly spatial and frequency smoothing operator N_s and N_f can be applied to perform an estimation of noninvertible unity rank spectral covariance matrix [13] [14].

$$\hat{\mathbf{E}} = \frac{1}{N} \sum_{n_s=1}^{2N_s+1} \sum_{n_f=1}^{2N_f+1} \mathbf{y}_{n_s,n_f} \mathbf{y}^{H}_{n_s,n_f} \tag{12}$$

$$N = (2N_s + 1)(2N_f + 1) \tag{13}$$

Where the long vector $\mathbf{y}_{n_s,n_f}$ corresponds to a concatenation of seismic waves received on the n_s^{th} sub array and n_f^{th} sub band respectively. According to Figure. 3, it is computationally expensive to diagonalize the whole estimated spectral covariance matrix $\hat{\mathbf{E}}$ of dimension $(G \times G)$ where, $G = \mathrm{K}_d \times (\mathrm{K}_x - 2N_s) \times \mathrm{K}_f$ [11].

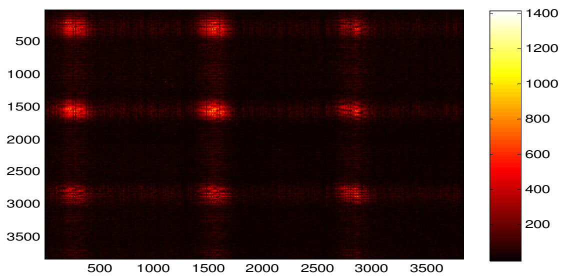

Figure 3- multicomponent spectral covariance matrix

For that reason, a new matrix $\mathbf{R}$ of size $(G \times N)$ that contains concatenated long-vectors resultant from the spatial and frequency smoothing can be proposed as [14].

$$\mathbf{R} = \overbrace{\begin{bmatrix} \vdots & & \vdots & & \vdots & & \vdots \\ \mathbf{y}_{1,1} & \cdots & \mathbf{y}_{2N_s+1,1} & \cdots & \mathbf{y}_{1,2N_f+1} & \cdots & \mathbf{y}_{2N_s+1,2N_f+1} \\ \vdots & & \vdots & & \vdots & & \vdots \end{bmatrix}}^{N=(2N_s+1)(2N_f+1)} \tag{14}$$

The frequency singular value decomposition of $\mathbf{R}$ matrix is

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{H} = \mathbf{U} \begin{bmatrix} \sqrt{\delta_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\delta_N} \\ & & \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \mathbf{V}^{H} \tag{15}$$

Where $\boldsymbol{\Lambda}$ is a pseudo diagonal matrix that holds the singular values of matrix $\mathbf{R}$ noted by $\sqrt{\delta_N}$. $\mathbf{U}$ and $\mathbf{V}^{H}$ are orthonormal matrices that contain the left and right singular vectors in their columns. The complex computation of spectral matrix can be expressed as follow:

$$\begin{aligned} \hat{\mathbf{E}}_1 = \mathbf{R}\mathbf{R}^{H} &= \mathbf{U}\boldsymbol{\Lambda}\ \mathbf{V}^{H}\mathbf{V}\boldsymbol{\Lambda}^{H}\mathbf{U}^{H} \\ &= \mathbf{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{H}\mathbf{U}^{H} \end{aligned} \tag{16}$$

Thus;

$$\hat{\mathbf{E}}_1 = \mathbf{R}\mathbf{R}^{H} = \mathbf{U} \begin{bmatrix} \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_N \end{bmatrix} & 0 \\ 0 & \begin{matrix} \ddots & \\ & 0 \end{matrix} \end{bmatrix} \mathbf{U}^{H} \tag{17}$$

On the other hand, the significant reduction of computing the estimated spectral matrix can be derived as follow:

$$\begin{aligned} \hat{\mathbf{E}}_2 = \mathbf{R}^{H}\mathbf{R} &= \mathbf{V}\boldsymbol{\Lambda}^{H}\mathbf{U}^{H}\mathbf{U}\boldsymbol{\Lambda}\ \mathbf{V}^{H} \\ &= \mathbf{V}\boldsymbol{\Lambda}^{H}\boldsymbol{\Lambda}\mathbf{V}^{H} \end{aligned} \tag{18}$$

Thus, $$\hat{\mathbf{E}}_2 = \mathbf{R}^{H}\mathbf{R} = \mathbf{V} \begin{bmatrix} \delta_1 & & 0 \\ & \ddots & \\ 0 & & \delta_N \end{bmatrix} \mathbf{V}^{H} \tag{19}$$

From equations (14) to (19) the eigenvectors matrix $\mathbf{U}$ of the reduced dimensional estimated spectral covariance matrix can now be expressed as:

$$\mathbf{U} = \mathbf{R}\mathbf{V}^{H} \begin{bmatrix} \sqrt{\delta_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{\delta_N} \end{bmatrix} \tag{20}$$

As a result, the filtering step corresponds to an orthogonal projection of a long vector $\mathbf{y}$ onto the eigenvectors in $\mathbf{U}$ that have the highest eigenvalues.

$$\mathbf{y}_{signal} = \sum_{i=1}^{P} \langle \mathbf{y}, \mathbf{u}_i \rangle . \mathbf{u}_i = \sum_{i=1}^{P} \frac{\mathbf{u}_i^H \mathbf{y} \mathbf{u}_i}{\|\mathbf{u}_i\|_2^2} \qquad (21)$$

These eigenvectors corresponding to the signal subspace $\hat{\mathbf{E}}_x$ where the eigenvectors that have the lowest eigenvalues related to the noise subspace $\hat{\mathbf{E}}_b$.

$$\begin{aligned} \hat{\mathbf{E}} &= \hat{\mathbf{E}}_x + \hat{\mathbf{E}}_b \qquad (22) \\ &= \mathbf{X} E[\mathbf{a}\mathbf{a}^H] \mathbf{X}^H + \sigma_b^2 \mathbf{I} \\ &= \sum_{i=1}^{P} \lambda_i \mathbf{u}_i \mathbf{u}_i^H + \sum_{i=P+1}^{N} \lambda_i \mathbf{u}_i \mathbf{u}_i^H \end{aligned}$$

With $\lambda_i = \delta_i$ for $1 \le i \le N$

Figure.4 shows the amplitude of 25 eigenvalues $\lambda_{i=1,2,\ldots,25}$ of the reduced dimensional covariance spectral matrix that related to the eigenvector matrix $\mathbf{U}$ in equation (20), as $N = 25$ where $(N_s = 2, N_f = 2)$. According to this figure, it can be shown that the first four eigenvalues which have the highest amplitude corresponding to the signal subspace, where the last lowest 21 eigenvalue corresponding to the noise subspace.

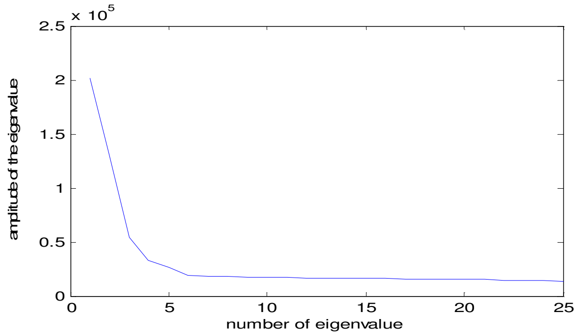

Figure 4- Eigenvalues of estimated reduced dimension covariance spectral matrix

The last steps consist of firstly, rearranging the signal subspace long vector $\mathbf{y}_{signal}$ into multicomponent form given in equation (2). Secondly, an inverse Fourier transform is applied into $\underline{\underline{\mathrm{Y}}}_{signal} \in \mathrm{E}^{K_d \times K_x \times K_f}$ to change back to the time domain. In terms of complexity, the decomposition of reduced dimensional spectral covariance matrix $\hat{\mathbf{E}}_2$ of size $(N \times N)$ rather than the spectral covariance matrix $\hat{\mathbf{E}}_1$ of size $(G \times G)$ where ($G >> N$), will dramatically decrease the computational processing power required for the spectral covariance matrix estimation. In our case this reduction is extremely vast as the estimated spectral covariance matrix size has been reduced from approximately (3700×3700) to (25×25).

4. SIMULATION RESULT

The proposed algorithm has been tested on a simulated wave-field seismic data which represents a multi-component array that is composed of 20 sensors, each of which is made up of three components. The first component relates to the geophone H, the second component to the geophone V, and the third component to the geophone Z. The recording time corresponds to 128 time samples. Two types of waves have been used, the first type linear polarization, the second type has elliptical polarization. Figure.5 shows the initial multi-component seismic wave-field data set that is recorded on the H, V, Z components respectively. The signal to noise ratio is nearly equal for all components and is relatively low. Figure.6 represents the purely signal subspace part separated from the noise subspace by projecting the initial data on the eigenvectors that is related to the dominant Eigen values with a spatial and frequency smoothing order equal to two. The Hodogram of the initial multi-component seismic wave field data set is presented in Figure.7. It corresponds to the layout of the amplitude of one component such as (H) versus the amplitude of others components (V, Z) for all traces. Since the initial seismic data is extremely noisy, Figure .7 shows that it is not possible to distinguish the wave-field information on the Hodogram. Figure.8 shows that the proposed algorithm has an excellent capability of separating the noise from the seismic wavefield by recovering both the elliptical polarization of the first waves as well as the linear polarization of the second waves**.** As shown in Figure.9, four tests have been conducted to evaluate the performance of proposed algorithm. As a result the mean square error (MSE) for the proposed technique shows an improvement in accuracy of 38.02% over the SVD algorithmin [7]. Furthermore, the proposed algorithm provides an extreme improvement in terms of computing time saving by 45.8% over the wide band spectral matrix approach in [10].

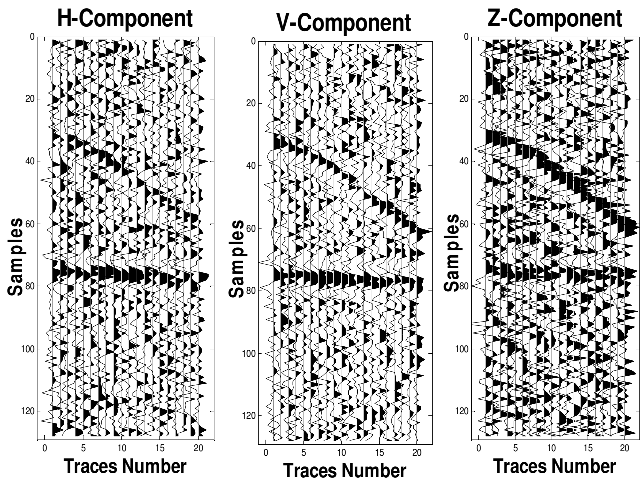

Figure 5-Initial seismic wavefield data set that are recorded on linear array of three component antennas

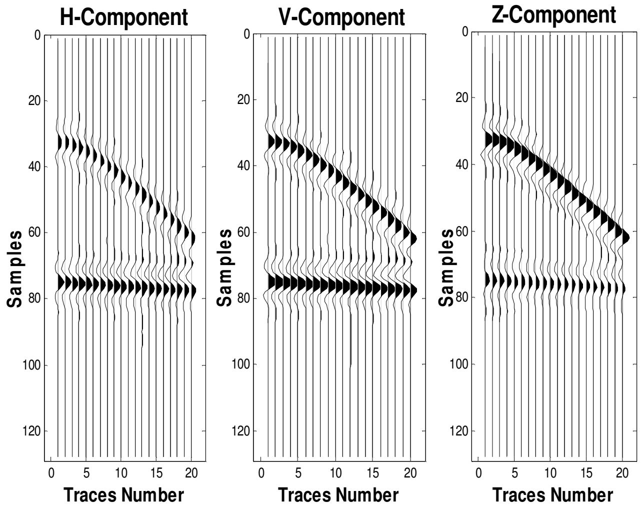

Figure 6-Filtered seismic wavefield data set that are recorded on linear array of three component antennas

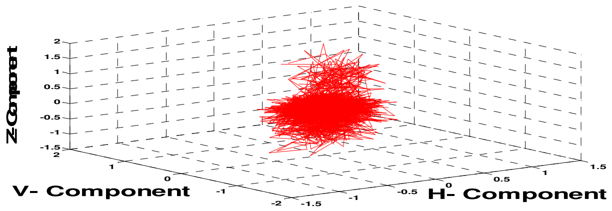

Figure 7- Hodogram of initial multicomponent seismic wavefield data set.

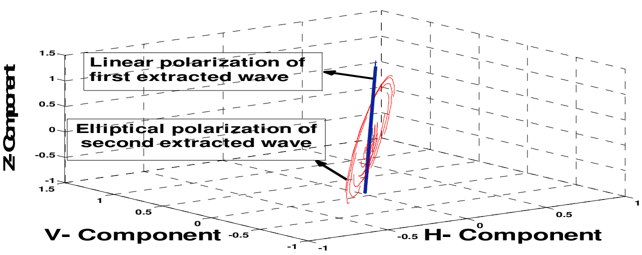

Figure 8**:** Hodogram of seismic wavefield after filtering

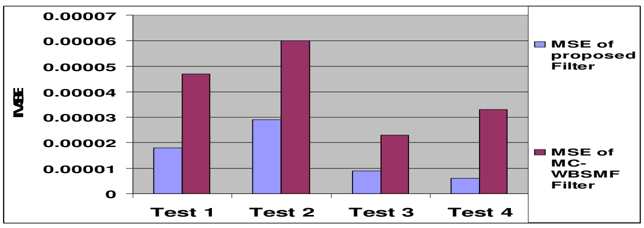

Figure 9-MSE of the proposed filter and WBSMF filter

5. CONCLUSION

A mathematical model of the noisy seismic wavefield that is received on a linear array of three component sensors has been fully developed and used as a framework to implement the proposed approach. This paper has proposed a new method for enhancing the SNR of the primary seismic waves. A significant reduction of spectral covariance matrix dimension has been proposed. The eigenvalue decomposition of the reduced dimension covariance spectral matrix has been derived from the SVD. Furthermore, the Hodogram has clearly demonstrated the success of extracted primary seismic waves from noisy multicomponent seismic wavefield. Comparing with other conventional algorithms, the proposed method has been extremely capable to reduce the computing time by 45.8% over the wide band spectral matrix algorithm as well as improving the accuracy by 38.02% over the SVD algorithm.

REFERENCES

[1] Aws. Al-Qaisi, W. l. Woo and S. S. Dlay, “Novel statistical approach to blind recovery of earth signal and source wavelet using independent component analysis,” *Wseas Trans on Signal Processing,* vol. 4, no.4, April 2008.

[2] M. Hanna, “Velocity filters for multiple interference attenuation in geophysical array data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 741–748, Nov. 1988.

[3] D.J. Foster and C.C. Mosher, “Suppression of multiple reflections using the Radon transform,” *Geophysics,* vol.57, no. 3, pp. 386–395, Mar.1992.

[4] P. Embree, J. P. Burg, and M. M. Backus, “Wide band velocity filtering the pie-slice process,” *Geophysics*, vol. 28, no. 6, pp. 948–974, Dec. 1963.

[5] A. Al-anboori, , M.V. D Baan and J. M. Kendall, “Approximate separation of pure mode and converted waves in 3-C reflection seismic by τ -p transform,” *Geophysics*, vol. 70, pp. 81-86, 2005.

[6] S. L. M. Freire and T. J. Ulrych, “Application of singular value decomposition to vertical seismic profiling,” *Geophysics*, vol. 53, no. 6, pp. 778–785, Jun. 1988.

[7] N. Le Bihan and J.I. Mars, “Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing,” *Signal Process*, vol. 84, pp. 1177–1199, June. 2004.

[8] V. D. Vrabie, J. I. Mars, and J.-L. Lacoume, “Modified singular value decomposition by means of independent component analysis,” *Signal Process*, vol. 84, no. 3, pp. 645–652, Mar. 2004.

[9] Natasha Hendrick, “Multi-component seismic wavefield separation via spectral matrix filtering” *Australian earth science convection* (AESC2006), Melbourne, Australia.

[10] C. Paulus and J. I. Mars, “New multicomponent filters for geophysical data processing,” *IEEE Trans. Geosci. Remote Sens*, vol.44, no.8, August. 2006.

[11] C. Paulus, P. Gounon, and J. I. Mars, “Wideband spectral matrix filtering for multicomponent sensor array,” *Signal Process,* vol. 85, no. 9, pp. 1723–1743, Sep. 2005.

[12] R.Kirlin, “Data covarianc matrices in seismic signal processing,” *Can SEGRec*, vol.26, no.4, pp.18-24, April.2001

[13] B. D. Rao and K. V. S. Hari, “Weighted subspace methods and spatial smoothing: Analysis and comparison,” *IEEE Trans. Signal Process*, vol. 41, no. 2, pp. 788–803, Feb. 1993

[14] S. U. Pillai and B. H. Kwon, “Fowrward backward spatial smoothing techniques for coherent signal identification,” *IEEE Trans. Acoust Speech Signal Process*, vol. 37, no. 1, pp. 8–15, Jan. 1989.