A NOVEL TECHNIQUE FOR BROADBAND SUBSPACE DECOMPOSITION

J G McWhirter, P D Baxter, T Cooper and S Redif

QinetiQ Ltd
Malvern Technology Centre, St Andrew's Road, Malvern, WR14 3PS, England
phone: + (44) 01684895384, fax: + (44) 01684896502, email: mcwhirter@signal.qinetiq.com
web: www.qinetiq.com

ABSTRACT

A generalisation of the eigenvalue decomposition (EVD) is proposed for para-Hermitian polynomial matrices. A novel technique for computing this polynomial matrix EVD is outlined. It involves applying a sequence of elementary paraunitary matrices and is referred to as the second order sequential best rotation (SBR2) algorithm. An application of the SBR2 algorithm to broadband subspace identification is briefly illustrated.

1. INTRODUCTION

The eigenvalue decomposition (EVD) is a very important tool for narrowband adaptive sensor array processing. It finds application in areas as diverse as high resolution direction finding, stabilised adaptive beamforming and blind signal separation [1,2]. The EVD decorrelates the signals received from an array of sensors by applying a unitary matrix of complex scalars, which serves to combine the signals with modified phase and amplitude. Because the transformation is unitary, the associated eigenvalues represent the true energy associated with each of the decorrelated components; thus the signal and noise subspaces may sometimes be identified and separated.

In broadband applications, or in a situation where narrowband signals have been convolutively mixed, the received signals cannot be related in terms of simple phase and amplitude factors so instantaneous decorrelation is no longer sufficient. It is necessary to impose decorrelation, not just at the same time instant for all pairs of signals, but over a suitably chosen range of relative time delays. This is referred to as strong decorrelation, and achieving it requires a matrix of suitably chosen finite impulse response (FIR) filters. If each filter is represented in terms of its transfer function, this takes the form of a polynomial matrix.

In this paper we generalise the EVD to broadband adaptive sensor arrays by requiring the strong decorrelation to be implemented using a paraunitary polynomial matrix. A paraunitary polynomial matrix represents a multi-channel all-pass filter and, accordingly, it preserves the total signal power at every frequency [3]. We also present a novel technique for computing the required paraunitary matrix and show how the resulting polynomial matrix EVD algorithm (SBR2) can be used in practice to identify broadband signal and noise subspaces. The algorithm, being highly generic in nature, has potential application to a wide range of important problems. These include broadband adaptive beamforming, broadband blind signal separation [4], multi-channel adaptive noise cancellation, the analysis of multiple-input multiple-output (MIMO) communication channels and the design of optimal filter banks for data compression.

Our approach is quite distinct from other time-domain methods reported to date. Lambert [5] has addressed the problem of broadband blind signal separation in the context of convolutive mixing. He represents the convolutive mixing in terms of DFT filter matrices as well as polynomial matrices. He has developed an EVD for polynomial matrices by generalising some conventional linear algebra and control techniques from the complex number field to the field of rational functions. His method involves the approximate inversion of FIR filters in the frequency domain and is therefore quite distinct from the one proposed here.

Regalia and Huang [6] have addressed the problem of computing a two channel lossless FIR filter for optimal data compaction. This leads to the determination of an optimum paraunitary matrix as required for our polynomial matrix EVD algorithm in the 2×2 case. Their approach exploits the fixed degree parameterisation proposed by Vaidyanathan [3], resulting in a difficult nonlinear optimisation. However, they re-formulate the problem using a state space approach and propose an iterative solution which avoids the problems of local minima associated with gradient descent techniques.

This paper is organised as follows. Section 2 discusses broadband sensor arrays and shows how the convolutive mixing of independent signals may be formulated in terms of polynomial matrices. The concept of a broadband EVD suitable for convolutive mixtures is then introduced. A tractable approach to computing the broadband EVD is described in section 3 and a specific algorithm is then outlined. Section 4 presents the results of some preliminary numerical simulations using the algorithm to perform broadband subspace decomposition. Section 5 contains some concluding remarks.

2. BROADBAND SIGNALS AND CONVOLUTIVE MIXING

The purpose of this paper is to suggest a novel technique for extending the EVD to broadband sensor array signal processing. In the case of a broadband sensor array, the signal received at each element may be represented as a linear superposition of delayed samples of the signals emitted by each source. In the case of q source signals $s_j(t)$ $(j = 1 \cdots q)$ and p

sensor outputs $x_i(t)$ $(i = 1 \cdots p)$ this may be expressed in the form

$$x_i(t) = \sum_{j=1}^{q} \sum_{k=0}^{l} a_{ij}(k) s_j(t-k) + \eta_i(t) \quad (t \in Z). \quad (2.1)$$

In effect, each sensor to signal channel is represented by an individual FIR filter $a_{ij}(t)$, which models the effects of multipath propagation and dispersion. This is generally referred to as convolutive mixing. In (2.1), $\eta_i(t)$ denotes one component of the vector $\boldsymbol{\eta}(t) \in C^p$ which represents additive noise drawn from an i.i.d. process with variance σ^2. We also adopt the vector notation $\mathbf{s}(t) \in C^q$ and $\mathbf{x}(t) \in C^p$ for the transmitted and received signals and assume that $\mathbf{s}(t)$, $\boldsymbol{\eta}(t)$ and $\mathbf{x}(t)$ have zero mean. Using the familiar transfer function notation, (2.1) may be expressed in the form

$$\mathbf{x}(z) = \mathbf{A}(z)\mathbf{s}(z) + \boldsymbol{\eta}(z) \quad (2.2)$$

where $\mathbf{A}(z)$ is a pxq polynomial matrix with elements of the form

$$a_{ij}(z) = \sum_{k=0}^{l} a_{ij}(k) z^{-k}$$

and $\mathbf{s}(z)$, $\boldsymbol{\eta}(z)$, $\mathbf{x}(z)$ denote algebraic power series in z^{-1} exemplifed by

$$\mathbf{s}(z) = \cdots + \mathbf{s}(0) + \mathbf{s}(1)z^{-1} \cdots + \mathbf{s}(t)z^{-t} \cdots.$$

It is assumed that the broadband source signals $s_j(t)$ $(j = 1,2,\cdots q)$ are statistically independent so the cross-correlation at all lags must be zero. Hence, assuming the statistics are wide-sense stationary, the space-time covariance matrix takes the form

$$\mathbf{R}_{ss}(\tau) = \mathrm{E}\{\mathbf{s}(t)\mathbf{s}^{\mathrm{H}}(t-\tau)\} = \mathrm{diag}\{\sigma_1(\tau), \sigma_2(\tau) \cdots \sigma_q(\tau)\}$$

where $\sigma_i(\tau)$ denotes the autocorrelation sequence of the ith signal. It follows that the cross-spectral density matrix is also diagonal and may be written in the form

$$\mathbf{R}_{ss}(z) = \sum_{\tau=-\infty}^{\infty} \mathbf{R}_{ss}(\tau) z^{-\tau} = \mathrm{diag}\{\sigma_1(z), \sigma_2(z) \cdots \sigma_q(z)\}$$

where $\sigma_i(z)$ denotes the z-transform expansion for $\sigma_i(\tau)$. As a result of the mixing process in (2.2), the received signals will generally be correlated and their cross-spectral density matrix, which takes the form

$$\mathbf{R}_{xx}(z) = \mathbf{A}(z)\mathbf{R}_{ss}(z)\tilde{\mathbf{A}}(z) + \sigma^2\mathbf{I}, \quad (2.3)$$

will not generally be diagonal. The tilde operation in (2.3) is used to represent paraconjugation, i.e. the combined operations of matrix transposition, substitution of z^{-1} for z, and complex conjugation of the polynomial coefficients [3].

The first stage of many signal processing algorithms is to filter and recombine the received signals $x_i(t)$ in order to generate signals $y_i(t)$, which (to a good approximation) are uncorrelated over a range of relative time delays. This may be achieved by a number of standard techniques such as multi-channel linear prediction using a least squares lattice filter [1]. However, since these methods do not conserve the spectral power in the signals, they cannot be used to identify the signal and noise subspaces. In order to overcome this limitation, it would be highly desirable to have a suitable broadband EVD algorithm.

We propose a broadband EVD algorithm of the form

$$\mathbf{y}(z) = \mathbf{H}(z)\mathbf{x}(z)$$

where

$$\hat{\mathbf{R}}_{yy}(z) = \mathbf{H}(z)\hat{\mathbf{R}}_{xx}(z)\tilde{\mathbf{H}}(z) \cong \mathrm{diag}\{d_1(z), \cdots d_p(z)\} \quad (2.4)$$

$\hat{\mathbf{R}}_{xx}(z)$ is a polynomial covariance matrix which serves to estimate $\mathbf{R}_{xx}(z)$ from the available data samples and $\hat{\mathbf{R}}_{yy}(z)$ is the corresponding polynomial matrix for the transformed signals. $\mathbf{H}(z)$ is a polynomial matrix constrained to be paraunitary which means that

$$\mathbf{H}(z)\tilde{\mathbf{H}}(z) = \tilde{\mathbf{H}}(z)\mathbf{H}(z) = \mathbf{I}$$

where tilde again denotes the paraconjugate. The polynomial matrix is constrained to be paraunitary so that the total power of the signals at every frequency is conserved by the transformation [3], i.e. it defines an all-pass filter. This ensures that the power in the resulting broadband signal and noise subspaces has proper physical significance.

For the sake of brevity, the explanation of this algorithm will be restricted to the special case of two signals (assumed to be real) and two sensors. This is sufficient to explain the basic concept. The challenge is to compute a paraunitary matrix $\mathbf{H}(z)$ such that the transformed polynomial covariance matrix in (2.4) is as close to diagonal as possible. In general, it will not be possible to achieve exact diagonalisation since the paraunitary matrix is composed of FIR filters. However, if the number of delay stages in the filter elements of the paraunitary matrix is sufficiently large, the decorrelation can be achieved to a very good approximation.

Since a general polynomial matrix is not necessarily paraunitary, it is vital to ensure that the approximate diagonalisation is carried out over the restricted space of paraunitary matrices. The easiest way of generating a paraunitary matrix is to use a suitably parameterised representation. Vaidyanathan [3] has shown that an arbitrary FIR paraunitary matrix can be decomposed into a set of rotations interspersed by delays. Apart from a scaling factor and a possible channel swap, a two-channel paraunitary matrix $\mathbf{H}_N(z)$ of degree N is decomposed as

$$\mathbf{H}_N(z) = \mathbf{Q}_N \cdots \boldsymbol{\Lambda}(z)\mathbf{Q}_1\boldsymbol{\Lambda}(z)\mathbf{Q}_0.$$

Here, $\boldsymbol{\Lambda}(z)$ denotes a unit delay applied to one channel, i.e.

$$\boldsymbol{\Lambda}(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}$$

and $\mathbf{Q}_i$ represents a 2x2 rotation matrix, which can be parameterised by a single rotation angle. For the purposes of broadband EVD, the challenge is to identify rotation matrices $\mathbf{Q}_0,\cdots\mathbf{Q}_N$ that minimise the output cross-correlation over multiple time lags. Unfortunately, this is a very difficult task since the individual rotations can not be computed independently and a multi-parameter nonlinear optimisation is required.

3. SEQUENTIAL BEST ROTATION ALGORITHM

In order to simplify the problem, we adopt a different formula for generating the paraunitary matrices. This takes the form

$$\mathbf{H}(z)=\mathbf{Q}_L\mathbf{\Lambda}^{d_L}\cdots\mathbf{Q}_1\mathbf{\Lambda}^{d_1} \qquad (3.1)$$

where the integer parameters d_i can be negative or positive. It can be seen that any polynomial matrix generated by (3.1) is paraunitary since each term is paraunitary. Equation (3.1) introduces the important new concept of an "elementary paraunitary matrix". This takes the form $\mathbf{G}(z)=\mathbf{Q}\mathbf{\Lambda}^d$ and comprises a number of delays (possibly negative) applied to one channel, followed by a rotation [4]. It is elementary in the sense that it only involves one rotation, but it does not necessarily have degree one. The second order sequential best rotation (SBR2) algorithm seeks to generate a paraunitary matrix according to (3.1) by calculating and applying an iterative sequence of suitably chosen elementary paraunitary matrices. This sequence is designed to minimise the strong decorrelation measure

$$g=\max\{|\hat{r}_{12}(\tau)|\,|\,\tau\in Z\}$$

where $\hat{r}_{12}(\tau)$ denotes the estimated correlation between the two transformed signals at lag τ, i.e. the off-diagonal element of the estimated polynomial covariance matrix for lag τ. At the initial stage, this is typically given by

$$\hat{\mathbf{R}}_{xx}(z)=\sum_{\tau=-W}^{W}\hat{\mathbf{R}}_{xx}(\tau)z^{-\tau} \qquad (3.2)$$

where

$$\hat{\mathbf{R}}_{xx}(\tau)=\sum_{t=0}^{T-1}\mathbf{x}(t)\mathbf{x}^{\mathrm{H}}(t-\tau)/T \ . \qquad (3.3)$$

It is assumed that $\hat{\mathbf{R}}_{xx}(\tau)\cong 0$ for $|\tau|>W$. This reflects the fact that for broadband signals, the space-time correlation function is negligibly small if $|\tau|$ is large compared to the coherence time. In practice, the value of W is often measured experimentally. It is also assumed that $T>>W$.

The SBR2 algorithm for two signals may be summarised as follows:

1) Compute an estimate of the space-time covariance matrix using, for example, the formula in (3.3).
2) Apply a relative delay between the two signals so that their instantaneous cross-correlation is maximised. This corresponds to the value of τ for which $|\hat{r}_{12}(\tau)|$ is greatest.
3) Rotate the realigned signals through the smallest angle θ which satisfies

$$\tan(2\theta)=\frac{2\hat{r}_{12}(0)}{\hat{r}_{11}(0)-\hat{r}_{22}(0)} \qquad (3.4)$$

This drives their zero-lag cross-correlation to zero.
4) Update the polynomial covariance matrix accordingly:

$$\hat{\mathbf{R}}_{xx}(z)\leftarrow\mathbf{G}(z)\hat{\mathbf{R}}_{xx}(z)\tilde{\mathbf{G}}(z) \qquad (3.5)$$

where $\mathbf{G}(z)$ denotes the elementary paraunitary matrix defined by steps 2 and 3.
5) Repeat steps 2 to 4 (which constitute one iteration) until the strong decorrelation measure g is sufficiently small.

Each iteration applies a single elementary paraunitary matrix, chosen to remove as much cross-correlation as possible at that stage. At first sight, this might not seem to be a sensible strategy since the successive elementary paraunitary matrices do not commute and applying a rotation doesn't just affect the current state but also the potential future gains of the algorithm. Unlike the narrowband case, applying a poorly chosen rotation is likely to make the problem more difficult by increasing the order of the mixing polynomial for no good reason. However, the freedom to choose an optimum delay for each stage makes this process much more meaningful.

In order to explain how the algorithm achieves its objective, we introduce the following set of measures relating to a given pair of signals $x_1(t)$ and $x_2(t)$:

$$N_1=\hat{r}_{11}(0)^2+\hat{r}_{22}(0)^2$$

$$N_2=\sum_{i,j=1}^{2}\hat{r}_{ij}(0)^2$$

and

$$N_3=N_2-N_1=2\hat{r}_{12}(0)^2\,.$$

N_3 is simply the cross-correlation between the two signals at zero lag while N_2 constitutes the squared norm of the zero-lag correlation matrix. It is easy to show that N_2 is invariant to a rotation of the two signal channels and obvious that N_1 is invariant to a delay applied to either channel. As N_2 is invariant to a rotation, and constitutes the sum of N_1 and N_3, any rotation which leads to a reduction in the value of N_3 must increase N_1 by the same amount. Each stage of the algorithm is designed to maximise the value of N_1. Now for any pair of signals, the value of N_3 can be driven to zero by rotating them through an angle θ which satisfies (3.4). It follows that the best delay for increasing N_1 is the one for which N_3 is greatest prior to the rotation, i.e. the one which maximises the zero-lag correlation between the shifted signals. Since the value of N_1 is unaffected by subsequently delaying either channel, it can be seen that successive steps of the SBR2 algorithm must lead to a monotonic increase in the value of N_1. It follows fairly directly, that this sequence of operations converges to a solution which achieves strong decorrelation in the sense that $g\rightarrow 0$. However, the full proof is not included here due to limitations on space.

In this paper we have chosen to describe the SBR2 algorithm in terms of elementary paraunitary matrices applied directly to the received signals. However, it can also be viewed in terms of elementary paraunitary matrices and their paraconjugates applied to the polynomial covariance matrix as evident from (3.5). The SBR2 algorithm should be regarded more fundamentally, as a paraunitary technique for diagonalising any para-Hermitian polynomial matrix (one that is identical to its paraconjugate).

Vaidyanathan [7] has shown that a paraunitary matrix designed for optimal subband coding must achieve strong decorrelation, and also impose spectral majorisation on the output signals. It is worth pointing out that the SBR2 algorithm, by virtue of its cost function, tends to impose spectral majorisation on the output signals provided this is consistent with the requirement for strong decorrelation. This enables it to be used to good effect for broadband subspace decomposition. Unfortunately, due to limited space it is not possible to explain that property here.

In this short paper we have only presented the SBR2 algorithm for the relatively simple case of two signal channels. This is sufficient to explain the key features of our approach. However, the method may be generalised to multiple channels in several ways. One of these may be viewed as a generalisation of the classical Jacobi algorithm for matrix diagonalisation. The results presented in the next section were produced using this method.

4. RESULTS

In order to demonstrate the computational effectiveness of the SBR2 algorithm, we present the results of a simple computer simulation experiment relating to broadband subspace decomposition. With reference to (2.2), the propagation of three signals onto five sensors was modelled by means of a 5x3 polynomial mixing matrix $\mathbf{A}(z)$ whose entries were order-5 FIR filters with coefficients drawn randomly from a uniform distribution in the range [-1, 1]. The source signals took the form of independent BPSK sequences for which each sample takes the value ± 1 with probability 1/2. Gaussian random noise was added to each simulated sensor output with variance σ^2 chosen to achieve the desired SNR. The experiment was repeated many times keeping the same input data and mixing matrix but with a different level of noise for each trial (independently generated). The number of samples, *T*, used to estimate the space-time covariance matrix in (3.3) was chosen to be 1000. The correlation window parameter, *W*, was set to 5 reflecting the statistics of the data and the order of the mixing matrix. For each chosen value of SNR, the SBR2 algorithm was used to strongly decorrelate the signals by diagonalising the estimated polynomial covariance matrix as indicated in (2.4). The algorithm was allowed to run for 500 iterations in each case.

The signal and noise subspaces were then separated, assuming that the SBR2 algorithm had converged. The signal subspace was simply defined by the three output channels with the highest estimated power. The integrity of the signal and noise subspaces was quantified using a measure of the form $\alpha = \alpha_n / \alpha_s$ where α_s and α_n denote the total expected power of the original signals projected onto the computed signal and noise subspaces respectively. The smaller the value of α the more reliable the subspace estimation. The values of α_s and α_n were computed directly from the paraunitary matrix $\mathbf{H}(z)$ produced by the SBR2 algorithm and the known mixing matrix $\mathbf{A}(z)$. Note that for unit power i.i.d. input signals (in the absence of noise) the cross-spectral density matrix for the decorrelated signals generated using $\mathbf{H}(z)$ is given by

$$\mathbf{R}_{yy}(z) = \mathbf{H}(z)\mathbf{A}(z)\tilde{\mathbf{A}}(z)\tilde{\mathbf{H}}(z) \qquad (4.1)$$

The value of α as a function of SNR is plotted in Fig.1. Each point on the graph represents the value of α for a single trial. It can be seen that for values of SNR greater than -5dB the value of α is less than 0.1 falling to less than 0.01 for SNR values greater than 5dB. This indicates that the algorithm is capable of effective broadband subspace decomposition.

Fig.2 shows how the magnitude, g, of the dominant off-diagonal coefficient behaves as a function of iteration number when the SNR value was set at 5dB. Within the first 200 iterations, it falls to less than 0.02 although the progress is not monotonic. The non-monotonic behaviour is to be expected since it is always possible for a reduction in the diagonal coefficients at non-zero lag to produce an increase in the off-diagonal coefficients.

Fig.3 relates to the same trial as Fig.2 (5dB) and depicts the power spectral density of the strongly decorrelated output signals produced using SBR2. Each line constitutes the plot, for a given value of *k,* of the k^{th} diagonal element of the cross-spectral density matrix in (4.1), evaluated at $z = e^{i\omega}$ for $0 < \omega \le 2\pi$. This illustrates the tendency of the SBR2 algorithm to generate spectrally majorised output signals as mentioned in section 3. The corresponding plots for the mixed signals prior to applying SBR2 are presented for comparison in Fig.4.

5. CONCLUSIONS

In this paper we have introduced the concept of a polynomial matrix EVD and suggested a tractable approach to performing the necessary computation. One specific algorithm has been outlined and some initial results presented. In many respects, the method presented here may be viewed as a direct extension of the Jacobi algorithm for conventional eigenvalue decomposition. A proof of convergence has been obtained but cannot be presented here. We have only illustrated the relevance of the SBR2 algorithm to signal processing in the context of strong decorrelation and broadband subspace decomposition. However, it could have as wide a range of applications for convolutive (broadband) sensor array signal processing as the conventional EVD or SVD algorithm does for instantaneous (narrowband) sensor array signal processing. It has already been applied successfully to data obtained from real sensor arrays in a number of application areas including sonar and seismology. However, discussion of the specific applications and results is beyond the scope of this paper. It has also been adopted successfully by other researchers for the purpose of designing oversampled filterbanks for channel coding [8] and for second order blind signal separation, applied to polarised signals from a 3-axis seismic sensor array using quaternion (hypercomplex) arithmetic [9].

ACKNOWLEDGEMENTS

This research was sponsored by the United Kingdom Ministry of Defence Corporate Research Programme. The authors are grateful to Joanne Foster of Cardiff University for her help in producing the results reported in section 4.

REFERENCES

[1] S Haykin, "Adaptive Filter Theory", 3rd edition, Prentice Hall (1996).

[2] P Comon, "Independent component analysis, A new concept?", Signal Processing, vol 36, pp 287-314 (1994)

[3] P P Vaidyanathan, "Multirate Systems and Filterbanks", Prentice Hall (1993).

[4] P D Baxter and J G McWhirter, "Blind signal separation of convolutive mixtures", Proc 37th Asilomar Conference on Circuits, Systems and Computers (Nov 2003).

[5] R. Lambert, "Multichannel blind deconvolution: FIR matrix algebra and separation of multipath mixtures", *PhD Thesis*, USC (1996).

[6] P. A. Regalia and D-Y. Huang, "Attainable error bounds in multirate adaptive lossless FIR filters", *Proc. IEEE Conf. on Acoustics, Speech and Signal. Proc.*, pp.1460-63 (1995).

[7] P P Vaidyanathyan, "Theory of optimal orthonormal subband coders", IEEE Trans Signal Processing, Vol 46, No 6, pp 1528-43 (June 1998).

[8] S. Weiss, S. Redif, T. Cooper, C. Liu, P. D. Baxter and J. G. McWhirter, “Paraunitary oversampled filterbank design for channel coding”, accepted for publication in *J. Adaptive Signal Processing*.

[9] N. Le Bihan, “Diagonalisation de matrices polynomiales quaternioniques: application a la seperation de melanges convolutifs d’ondes polarisees”, *Proc GRETSI Conference*, Louvaine-la-Neuve, Belgium (Sept 2005)

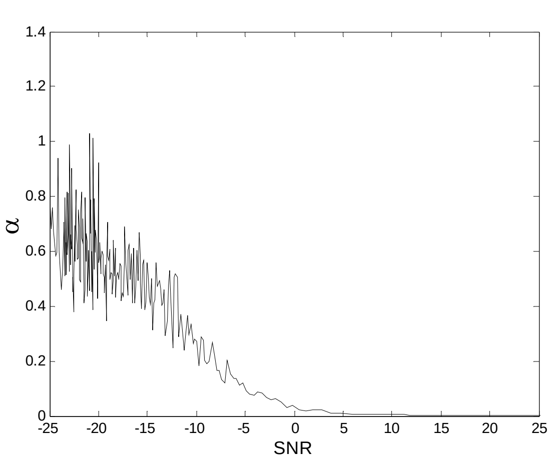

Fig.1. Performance measure for broadband subspace decomposition.

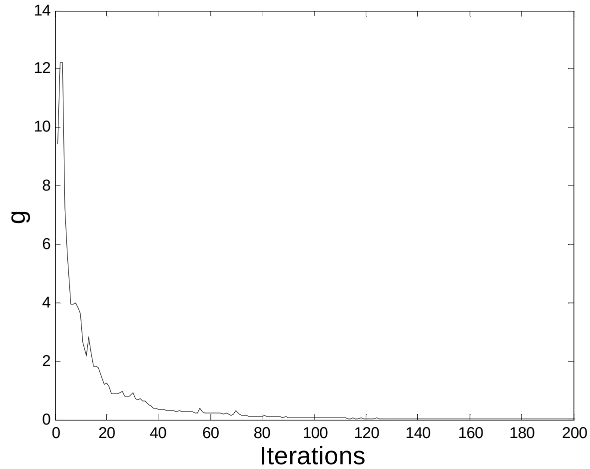

Fig.2. Convergence behaviour observed in case of 5dB SNR.

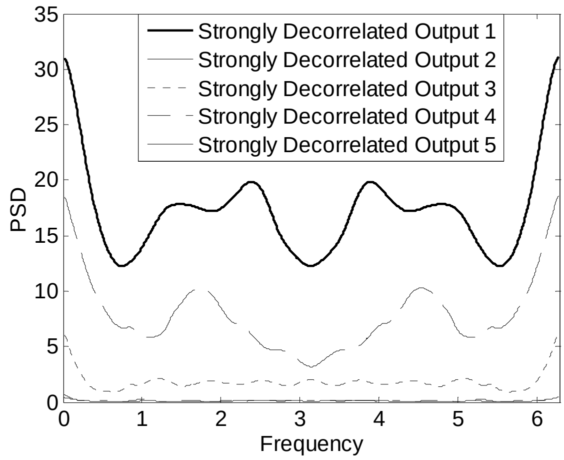

Fig.3. Power spectral density of output signals generated using SBR2 at 5dB SNR.

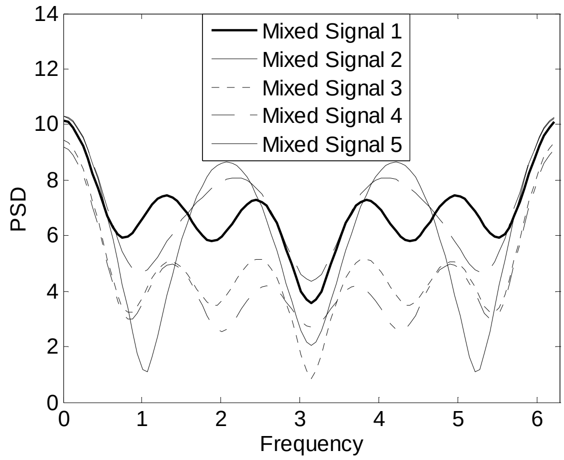

Fig.4. Power spectral density of mixed signals at 5dB SNR before applying SBR2.