POLYNOMIAL SINGULAR VALUES FOR NUMBER OF WIDEBAND SOURCES ESTIMATION AND PRINCIPAL COMPONENT ANALYSIS

Russell H. Lambert

RF and Advanced Mixed Signal Unit
Broadcom
Pasadena, CA USA
russ@broadcom.com

Marcel Joho and Heinz Mathis

Signal and Information Processing Lab
Swiss Federal Institute of Technology
Zurich, Switzerland
joho,mathis@isi.ee.ethz.ch

ABSTRACT

A multipath enabled singular value decomposition (SVD) algorithm is presented, which will allow computation of wideband (polynomial) singular values, and hence, the signal+noise and noise subspaces. Polynomial singular values are ordered according to total energy. The number of sources can be estimated using the scalar total energy values. Results using both simulated data on the computer and actual speech recorded in a noisy multipath environment are given to demonstrate the usefulness of the techniques shown. After number of sources estimation, only the signal+noise subspace is used to create virtual sensors which have made optimal use of all the sensors. As a final signal copy step, standard blind independent component analysis (ICA) or blind source separation algorithms can be used to recover the original data from the virtual sensors. The number of estimated sources could also be given to a blind algorithm capable of using overdetermined sources and the algorithm can adaptively make use of all sensor data.

1. INTRODUCTION

We address the problem of efficiently recovering M_s sources given M sensors of noisy, multipath mixed data. A key step in the recovery process is estimation of the number of original sources. In earlier work, we have given tools to compute eigenvalues for multipath matrices. Singular value decomposition (SVD) is conceptually an eigenroutine applied to the spectrum of a matrix, $\underline{\mathbf{A}}\,\underline{\mathbf{A}}^H$, in which eigenvalues and associated eigenvectors are ordered according to energy. Herein, we show how to use total energy (over all time extent or, equivalently, over all frequencies) of the eigenvalues for ordering of eigenvalues and eigenvectors and deciding between signal+noise and noise only subspaces.

Narrowband/scalar mixture processes are well described by scalar matrix algorithms, but adding a wideband/multipath dimension to the problem requires a new set of tools. We present a fundamental singular value decomposition technique which uses (finite impulse response) FIR matrices, whose elements are vectors and have time domain (or equivalently, frequency domain) extent.

2. PROBLEM DESCRIPTION

The mixing process is described as

$$\mathbf{x}_t = \underline{\mathbf{A}}\mathbf{s}_t + \mathbf{n}_t \tag{1}$$

where $\mathbf{s}_t = (s_1, \cdots, s_{M_s})_t^T$ contains the samples of the unknown source signals at time t, $\mathbf{x}_t = (x_1, \cdots, x_M)_t^T$ the samples of the M sensor signals at sample time t, $\mathbf{n}_t$ the samples of the sensor noise at time t, with $\underline{\mathbf{A}}^{M \times M_s \times L}$ as the unknown mixing matrix. In the overdetermined case, we have more sensors than source signals ($M > M_s$).

2.1. Notation

The following notation is used throughout: Vectors are written in lower case, matrices in upper case. The sample index is denoted by t. $E\{\cdot\}$ denotes the expectation operator. Vector or matrix dimensions are given in superscript. The Frobenius norm of a matrix is denoted by $\|\cdot\|_F$. Column vectors are bold-face, lower case symbols, as in $\mathbf{x}$, and column vectors with filters as elements are underlined bold-face, lower case symbols, as in $\underline{\mathbf{x}}$. Matrix variables are in bold face, upper case, as in $\mathbf{A}$, and matrices of filters (FIR matrices (see [1]) are underlined bold face, upper case symbols, as in $\underline{\mathbf{A}}$. FIR matrices are usually assumed to be in the frequency domain where convolution is multiplication and standard Linear Algebra notation holds. We convert the matrices to the frequency domain, perform all the computations and then convert back to the time domain at the final step for plotting the result. The time domain or multipath extent of filters and FIR matrices is L taps in length. $\underline{\mathbf{I}}$ is the unit FIR matrix and $\underline{\mathbf{0}}$ is the zero FIR matrix. Inner product, row times column matrix multiplication rules, etc., all apply in FIR matrix algebra, but the scalar matrix term by scalar matrix term multiplication is replaced with element-wise (across all L) multiplication of the frequency domain representations of the corresponding FIR matrix terms.

2.2. Assumptions

In addition to the problem proposed above, we make the following assumptions:

- $A1$ Time-invariant mixing matrix $\underline{\mathbf{A}}$.
- $A2$ $\underline{\mathbf{A}}$ has rank M_s.
- $A3$ Source signals $s_m, m = 1, \dots, M_s$, are mutually independent and iid.
$$\underline{\mathbf{R}}_{ss} = E\{\mathbf{s}_t \mathbf{s}_t^H\} = \sigma_s^2 \, \underline{\mathbf{I}}_{M_s} \tag{2}$$
- $A4$ Source signals s_m (save possibly one) are non-Gaussian. This assumption is required for full independent component analysis (or signal copy), but not for the PCA stage.
- $A5$ All source signals are unknown and have the same power σ_s^2.
- $A6$ There are more sensors than source signals $M > M_s$.
- $A7$ All sensors have the same noise characteristics. The sensor noise is additive white Gaussian noise with power σ_n^2. The sensor noise of the sensors is mutually independent.
$$\underline{\mathbf{R}}_{nn} = E\{\mathbf{n}_t \mathbf{n}_t^H\} = \sigma_n^2 \, \underline{\mathbf{I}}_M. \tag{3}$$
- $A8$ The source signals and the sensor noise are mutually independent.

3. POLYNOMIAL SVD USING TOTAL ENERGY ORDERING

By applying SVD on $\underline{\mathbf{A}}$, we have

$$\underline{\mathbf{A}} = \underline{\mathbf{U}} \, \underline{\mathbf{\Sigma}} \, \underline{\mathbf{V}}^H = \underline{\mathbf{U}} \begin{bmatrix} \underline{\tilde{\mathbf{\Sigma}}} \\ \underline{\mathbf{0}} \end{bmatrix} \underline{\mathbf{V}}^H \tag{4}$$

where $\underline{\mathbf{U}}^{M \times M \times L}$ and $\underline{\mathbf{V}}^{M_s \times M_s \times L}$ are unitary FIR matrices ($\underline{\mathbf{U}} \, \underline{\mathbf{U}}^H = \underline{\mathbf{I}}^{M \times M \times L}$ and $\underline{\mathbf{V}} \, \underline{\mathbf{V}}^H = \underline{\mathbf{I}}^{M_s \times M_s \times L}$). Unitary FIR matrices are isomorphic to the paraunitary matrices defined in [2]. Unitary matrices contain only phase information.

$\underline{\mathbf{\Sigma}}^{M \times M_s \times L}$ and $\underline{\tilde{\mathbf{\Sigma}}}^{M_s \times M_s \times L}$ are diagonal FIR matrices which contain the polynomial singular values of $\underline{\mathbf{A}}$

$$\underline{\tilde{\mathbf{\Sigma}}} = \text{diag}\{\underline{\sigma}_1, \dots, \underline{\sigma}_{M_s}\} \tag{5}$$

with

$$p_1 \geq p_2 \geq \cdots \geq p_{M_s} > 0 \tag{6}$$

where the last inequality comes from the assumption $A2$, and where the scalars p_m are defined as the total summed energy (same in time or frequency domain) in the wideband singular values.

$$p_m = \Sigma_{\text{all } k} \, \sigma_m^2(k) \tag{7}$$

The SVD of the input correlation matrix $\underline{\mathbf{R}}_{xx}$ gives with (2) and (3)

$$\begin{aligned}
\underline{\mathbf{R}}_{xx} &= E\{\mathbf{x}_t \mathbf{x}_t^H\} && (8)\\
&= \underline{\mathbf{A}} \, \underline{\mathbf{R}}_{ss} \, \underline{\mathbf{A}}^H + \underline{\mathbf{R}}_{nn} && (9)\\
&= \sigma_s^2 \underline{\mathbf{A}} \, \underline{\mathbf{A}}^H + \sigma_n^2 \, \underline{\mathbf{I}}_M && (10)\\
&= \sigma_s^2 \, \underline{\mathbf{U}} \, \underline{\mathbf{\Sigma}} \, \underline{\mathbf{V}}^H \, \underline{\mathbf{V}} \, \underline{\mathbf{\Sigma}}^T \, \underline{\mathbf{U}}^H + \sigma_n^2 \, \underline{\mathbf{I}}_M && (11)\\
&= \sigma_s^2 \, \underline{\mathbf{U}} \, \underline{\mathbf{\Sigma}} \, \underline{\mathbf{\Sigma}}^T \, \underline{\mathbf{U}}^H + \sigma_n^2 \, \underline{\mathbf{I}}_M && (12)\\
&= \sigma_s^2 \, \underline{\mathbf{U}} \, \underline{\bar{\mathbf{\Sigma}}}^2 \, \underline{\mathbf{U}}^H + \sigma_n^2 \, \underline{\mathbf{U}} \, \underline{\mathbf{U}}^H && (13)\\
&= \underline{\mathbf{U}}(\sigma_s^2 \, \underline{\bar{\mathbf{\Sigma}}}^2 + \sigma_n^2 \, \underline{\mathbf{I}}_M)\underline{\mathbf{U}}^H && (14)\\
&= \underline{\mathbf{U}} \, \underline{\mathbf{\Sigma}}_{xx}^2 \, \underline{\mathbf{U}}^H && (15)
\end{aligned}$$

with

$$\underline{\bar{\mathbf{\Sigma}}}^{M \times M \times L} = \text{diag}\{\underline{\sigma}_1, \dots, \underline{\sigma}_{M_s}, 0, \dots, 0\}$$

$$\underline{\mathbf{\Sigma}}_{xx}^2 = \text{diag}\{\underline{\sigma}_{x_1}^2, \dots, \underline{\sigma}_{x_{M_s}}^2, \underline{\sigma}_{x_{M_s+1}}^2, \dots, \underline{\sigma}_{x_M}^2\}$$

$$\underline{\mathbf{\Sigma}}_{xx}^2 = \text{diag}\{\underline{\sigma}_1^2 \sigma_s^2 + \sigma_n^2, \dots, \underline{\sigma}_{M_s}^2 \sigma_s^2 + \sigma_n^2, \underline{\sigma}_n^2, \dots, \sigma_n^2\}.$$

4. SIMULATION

4.1. PCA Results using Speech Recorded in a Noisy Reverberant Environment (M_s=1, M=5)

Five sensors were recorded of a speech signal uttered in an automobile traveling at freeway speeds. The signal-to-noise ratio is a function of frequency, but was very poor, in the subjective range of 5 to -5 dB at each sensor. The multipath correlation matrix for this data corresponding to $\underline{\mathbf{R}}_{xx}$, equation (9) is shown in figure 1. A hundred seconds of data was used, with a sample rate of 8000Hz. A time extent of 16000 samples (2 seconds) was used. The energy order eigenvalues of $\underline{\mathbf{R}}_{xx}$ are shown in figure 2. Note that as in scalar matrix algebra the eigenvalues of $\underline{\mathbf{R}}_{xx}$ will contain spectral magnitude information only (no phase). The eigenvalues shown are symmetric about time zero (completely phaseless). Also, the singular vector matrices $\underline{\mathbf{U}}$ and $\underline{\mathbf{V}}$ are unitary FIR matrices. We demonstrate this property in figure 3 by plotting $\underline{\mathbf{U}} \, \underline{\mathbf{U}}^H = \underline{\mathbf{I}}$.

4.2. PCA Results using Computer Generated Data (M_s=3, M=6)

Next, we perform a controlled test using simulated uniformly distributed source and white Gaussian noise. Three sources of white uniformly distributed data (from $-.5$ to $.5$)were convolved with the mixture matrix $\underline{\mathbf{A}}$ given by:

$$\begin{bmatrix} .4 & -.1 + .1z^{-1} & -.2 \\ -.1z^{-1} & -.2 + .2z^{-1} & .4 + .23z^{-1} \\ .1 - .2z^{-1} & .1 + .5z^{-1} & .3 - .1z^{-1} \\ .3 & .3z^{-1} & -.4z^{-2} \\ -.5z^{-1} & .1 & -.25z^{-1} + .15z^{-2} \\ .45 & -.38z^{-1} - .2z^{-2} & -.3z^{-1} + .1z^{-2} \end{bmatrix} \tag{19}$$

Then, white Gaussian noise of $\sigma_n = .1$ was added. Results of PCA follow as above, see figures 4, 5, 6, 7, 8.

4.3. ICA Results using MBLMS and Computer Generated Data (M_s=3, M=6)

The same data from the previous PCA simulation was applied to the multichannel blind LMS algorithm ([1]) for ICA (full true-phase signal copy). The algorithm must be told the number of sources to look for (this case 3). Figure 9 shows a high quality separation using a scale invariant measure ISI.

5. CONCLUSION AND IMPLICATION TO SOURCE SEPARATION OF OVERDETERMINED MULTIPATH MIXTURES

Using FIR matrix tools, we have presented a multipath extension of SVD for signal subspace processing for estimation of number of sources and principal component analysis. As the FIR matrix singular values are vectors with time/frequency extent, we use the scalar total energy of each singular value for ordering purposes. We have demonstrated the extended SVD and number or sources estimation on noisy reverberant speech data and simulated systems.

Recently in the area of source separation, there has been interest to find the best ways to use extra (overdetermined) sensors, see [3, 4, 5] . This has been especially true in light of the fact that key algorithms which use RLS-type or relative gradient updates are restricted to dealing with square ($M_s = M$) mixtures. Our proposal is to first estimate the number of original sources using the presented PCA/SVD technique. Then, one has the option of collapsing the data to virtual sensors containing only the principal sources and proceeding to the ICA step were one can now use algorithms limited to square mixtures. Alternatively, one can use an LMS type update algorithm (not restricted to square mixtures) which is instructed to recover the estimated number of original sources.

6. REFERENCES

[1] R. H. Lambert and C. L. Nikias, "Blind deconvolution of multipath mixtures," in *Unsupervised Adaptive Filtering*, S. Haykin, Ed. 2000, vol. 1, pp. 377–436, John Wiley & Sons.

[2] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, 1993.

[3] S. C. Douglas, S.-I. Amari, and S.-Y. Kung, "Adaptive paraunitary filter banks for spatio-temporal principal and minor subspace analysis," in *ICASSP*, Phoenix, AZ, March 1999, vol. 2, pp. 1089–1092.

[4] L.-Q. Zhang, A. Cichocki, and S. Amari, "Natural gradient algorithm for blind separation of overdetermined mixture with additive noise," *IEEE Signal Processing Letters*, pp. 293–295, November 1999.

[5] M. Joho, H. Mathis, and R. H. Lambert, "Overdetermined blind source separation: Using more sensors than source signals in a noisy mixture," in *Proc. International Conference on Independent Component Analysis and Blind Signal Separation*, Helsinki, Finland, June 19–22, 2000, pp. 81–86.

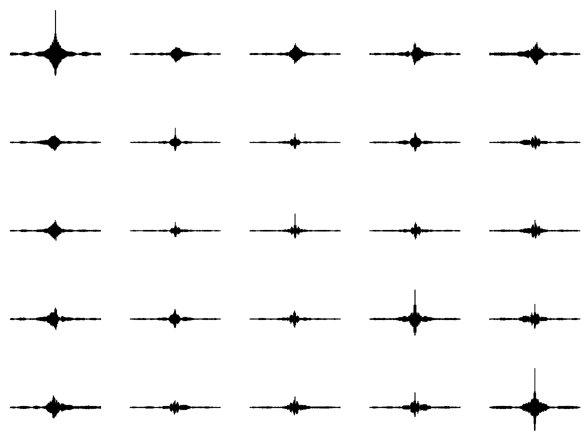

Fig. 1. $\underline{\mathbf{R}}_{xx}$ of acoustic recording with one true speech source in noise.

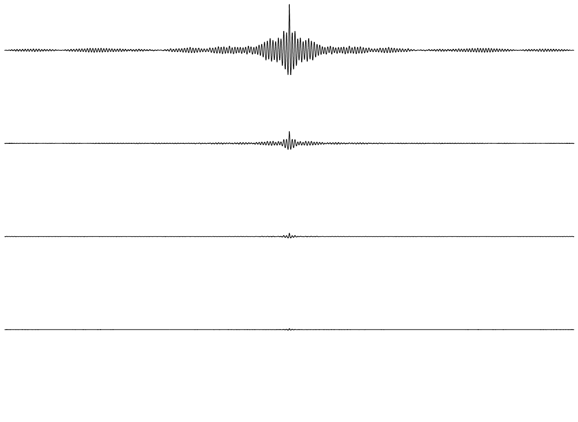

Fig. 2. Energy ordered singular values of $\underline{\mathbf{R}}_{xx}$ for acoustic recording.

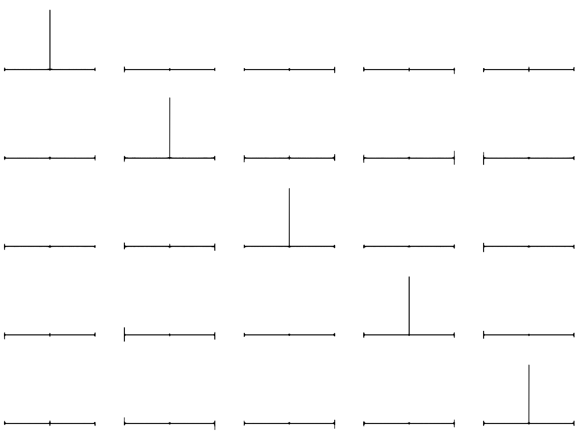

Fig. 3. The result of $\underline{\mathbf{U}}\,\underline{\mathbf{U}}^H$ for acoustic recording.

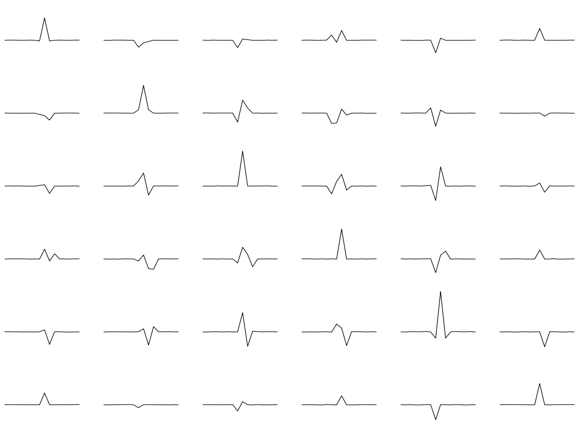

Fig. 4. $\underline{\mathbf{R}}_{xx}$ of computer data with 3 sources and 6 sensors.

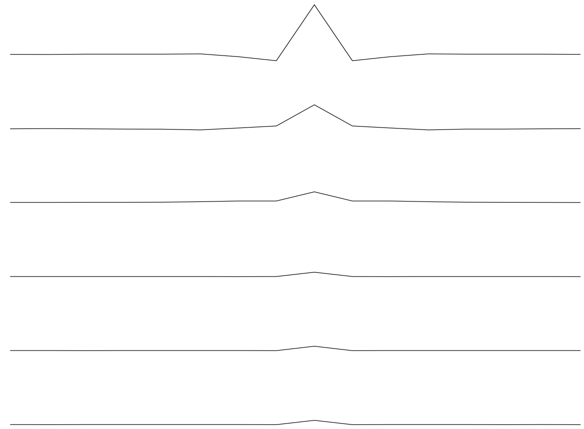

Fig. 5. Energy ordered singular values of $\underline{\mathbf{R}}_{xx}$ of computer data with 3 sources and 6 sensors.

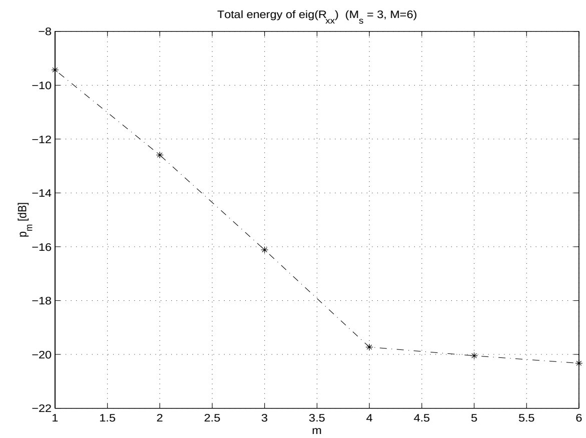

Fig. 6. The terms p_m, scalar summed energy of $\underline{\mathbf{R}}_{xx}$ of computer data with 3 sources and 6 sensors.

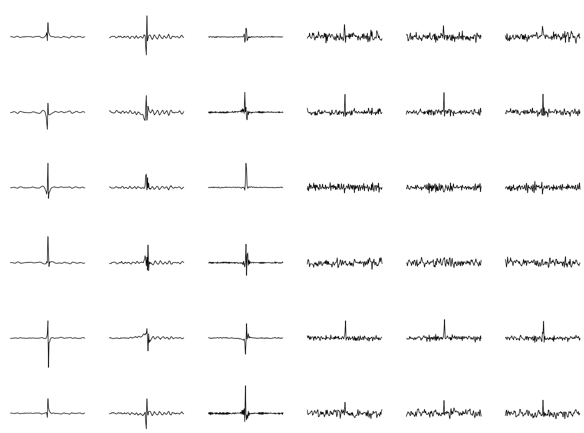

Fig. 7. Unitary singular vector matrix $\underline{\mathbf{U}}$ of computer data with 3 sources and 6 sensors.

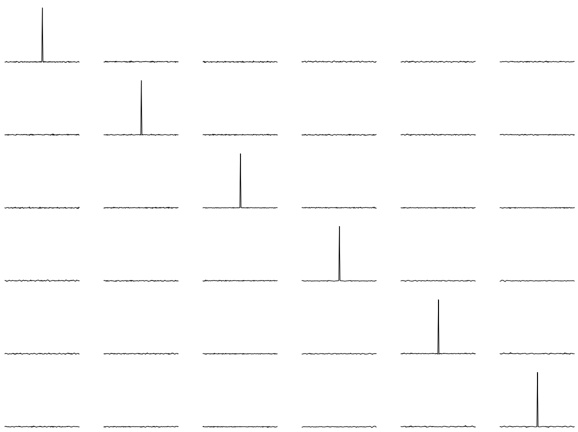

Fig. 8. The result of $\underline{\mathbf{U}}\,\underline{\mathbf{U}}^H$ for computer data with 3 sources and 6 sensors.

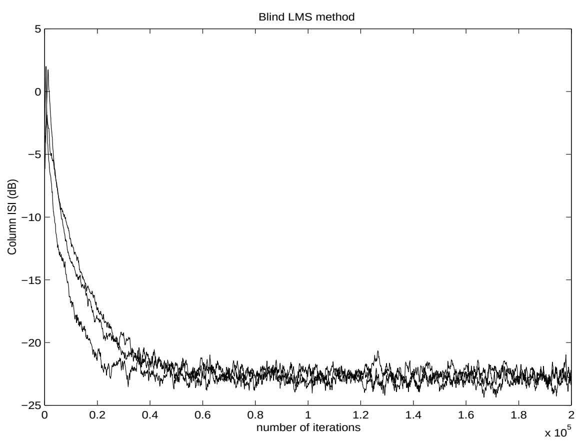

Fig. 9. ICA results for (M_s=3, M=6) multichannel blind LMS adaptation. Column ISI (a scale invariant measure of convergence) taken from the three columns of the global system as convergence proceeds.