A TEST STATISTIC FOR SEQUENTIAL IDENTIFICATION OF CO-CHANNEL DIGITAL SIGNALS USING A DEFLATION APPROACH

Lars K. Hansen[1] *Guanghan Xu*[1] *

[1]Department of Electrical and Computer Engineering,
University of Texas at Austin,
Austin, Texas 78712-1084, USA

ABSTRACT

When sequentially separating a linear combination of co-channel digital signals, it is necessary at each step to test the validity of the currently estimated signal prior to proceeding to extract the next one. We describe a procedure for use with sequential algorithms which uses a deflation-based approach combined with a simple test statistic. The deflation step removes the contributions of the currently identified signals. The simple test statistic takes into account the error terms introduced into the data by the deflation. The method has been successfully applied in an existing sequential estimation algorithm.

1 INTRODUCTION AND PROBLEM FORMULATION

The general source separation problem has been studied for some time, see [1] for a good discussion and bibliography. More recently, the problem has been specialized to the case of digital signals, [2, 3, 4, 5, 6] for example. These types of algorithms have applications to demodulation of multiple co-channel digital signals which arise in spatial-division-multiple-access and cellular telephony. All of these algorithms simultaneously estimate all of the multiple co-channel signals.

In this paper, we consider a problem motivated by an algorithm which *sequentially* separates an instantaneous mixture of digital signals. Suppose that we have an algorithm which can estimate one signal from the mixture of digital signals. (The algorithms in [7, 8] are the motivating examples for this paper.) As each signal is estimated, it is "removed" from the instantaneous mixture and a new "deflated" problem of lower dimension is formed, and the process is repeated. At any given stage, the estimated signal may be faulty; in particular, the fault may be due to convergence of the estimation algorithm to a non-global optima. Therefore, it is necessary to test each estimate for validity. This testing is complicated by the fact that correlations among the signals introduce an additional error term into the deflated problem. In this paper, we describe a deflation method and derive a simple validity test for each sequential estimate which takes this signal correlation into account.

We consider the problem

$$\mathbf{X} = \mathbf{AS} + \mathbf{WV}$$

where $\mathbf{X}$ is a $d \times N$ matrix of known data, $\mathbf{A}$ is a full-rank $d \times d$ matrix, $\mathbf{S}$ is a deterministic full-rank $d \times N$ signal matrix drawn from a common digital alphabet, $\mathbf{W}$ is a *known* full-rank $d \times m$ matrix, and $\mathbf{V}$ is a $m \times N$ matrix of i.i.d. Gaussian variables with zero mean and known variance σ^2. We have that $d \leq m \ll N$. The matrix $\mathbf{W}$ is known because it was formed to whiten and reduce a dataset that was originally was of size $m \times N$. We assume that all quantities are real variables; no additional insight is gained by considering the complex case.

We define the following vectors and sub-matrices for future reference:

$$\mathbf{A} \equiv \quad [\mathbf{a}_1 \cdots \mathbf{a}_d] \quad = [\mathbf{A}_k \ \mathbf{A}_{d-k}]$$

where

$$\mathbf{A}_k \equiv [\mathbf{a}_1 \cdots \mathbf{a}_k] \qquad \text{and} \qquad \mathbf{A}_{d-k} \equiv [\mathbf{a}_{k+1} \cdots \mathbf{a}_d],$$

$$\text{and} \quad \mathbf{S} \equiv \begin{bmatrix} \mathbf{s}_1^T \\ \vdots \\ \mathbf{s}_d^T \end{bmatrix} = \begin{bmatrix} \mathbf{S}_k \\ \mathbf{S}_{d-k} \end{bmatrix}$$

where

$$\mathbf{S}_k \equiv \begin{bmatrix} \mathbf{s}_1^T \\ \vdots \\ \mathbf{s}_k^T \end{bmatrix} \qquad \text{and} \qquad \mathbf{S}_{d-k} \equiv \begin{bmatrix} \mathbf{s}_{k+1}^T \\ \vdots \\ \mathbf{s}_d^T \end{bmatrix}.$$

Let us describe in words the meanings of the various notations. $\mathbf{S}$ is composed of d rows of data; row $\mathbf{s}_i^T$ represents the signal from the ith user. Since we are considering a sequential identification scheme, the submatrix $\mathbf{S}_k$ respresents the previously identified k signals. Without loss of generality, we assume that the first k rows of $\mathbf{S}$ have been found. The matrix $\mathbf{S}_{d-k}$ represents the currently unidentified $d-k$ signals from $\mathbf{S}$. The matrix $\mathbf{A}_k$ is the submatrix of $\mathbf{A}$ consisting of those columns associated with the k signals $\mathbf{S}_k$; the submatrix $\mathbf{A}_{d-k}$ represents the remaining columns.

*This work was sponsored in part by NSF CAREER Award under Grant MIP-9502695, Office of Naval Research under Grant N00014-95-1-0638, the Joint Services Electronics Program under Contract F49620-95-C-0045, Motorola, Inc., Southwestern Bell Technology Resources, Inc. and Texas Instruments. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

We assume that we have previously estimated $\mathbf{S}_k$ correctly. We use this information to "remove" the contributions of $\mathbf{S}_k$ from $\mathbf{X}$ and deflate the problem to a lower dimension. To perform this deflation, we must estimate $\mathbf{A}_k$ and the range of $\mathbf{A}_{d-k}$. We describe this estimation process in detail in section 2. In section 3, we describe the deflation process and analyze the effects of estimation errors. In section 4, we derive a simple validity test for the hypothesis that the estimate $\hat{\mathbf{s}}_{k+1}^T$ of the next row of $\mathbf{S}$ is correct. We then conclude the paper.

2 ESTIMATION OF $\mathbf{A}_k$ AND RANGE OF $\mathbf{A}_{d-k}$

In this section, we describe and analyze the estimates required by our deflation procedure. The deflation procedure uses an oblique projection which requires estimates of $\mathbf{A}_k$ and the range of $\mathbf{A}_{d-k}$, so we begin by describing how these quantities are estimated. We then analyze sources of bias in these estimates.

To estimate $\mathbf{A}_k$, we use the pseudo-inverse of $\mathbf{S}_k$:

$$\mathbf{S}_k{}^\dagger \equiv \left(\mathbf{S}_k{}^T \left(\mathbf{S}_k \mathbf{S}_k{}^T\right)^{-1}\right).$$

Then $\mathbf{A}_k$ is estimated simply as

$$\hat{\mathbf{A}}_k \equiv \mathbf{X}\mathbf{S}_k{}^\dagger = \mathbf{A}_k + \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{S}_k{}^\dagger + \mathbf{W}\mathbf{V}\mathbf{S}_k{}^\dagger. \quad (1)$$

The expected value of this estimate is

$$\mathcal{E}\{\hat{\mathbf{A}}_k\} = \mathbf{A}_k + \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{S}_k{}^\dagger$$

so the estimate is biased by the unknown correlation between $\mathbf{S}_k$ and $\mathbf{S}_{d-k}$.

Next, we form the quantity

$$\begin{aligned}
\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k &= [\mathbf{A}_k\ \mathbf{A}_{d-k}]\begin{bmatrix}\mathbf{S}_k \\ \mathbf{S}_{d-k}\end{bmatrix} + \mathbf{W}\mathbf{V} \\
&\quad -\mathbf{A}_k\mathbf{S}_k - \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{S}_k{}^\dagger\mathbf{S}_k - \mathbf{W}\mathbf{V}\mathbf{S}_k{}^\dagger\mathbf{S}_k \\
&= \mathbf{A}_{d-k}\mathbf{S}_{d-k}\left[\mathbf{I} - \mathbf{P}_{\mathbf{S}_k}\right] + \mathbf{W}\mathbf{V}\left[\mathbf{I} - \mathbf{P}_{\mathbf{S}_k}\right] \\
&= \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}^\perp + \mathbf{W}\mathbf{V}\mathbf{P}_{\mathbf{S}_k}^\perp
\end{aligned}$$

where we note that the quantity

$$\mathbf{P}_{\mathbf{S}_k} \equiv \mathbf{S}_k{}^\dagger\mathbf{S}_k$$

is a projection matrix and $\mathbf{P}_{\mathbf{S}_k}^\perp$ is the orthogonal projection. Then

$$\begin{aligned}
&\mathcal{E}\left\{\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)^T\right\} \\
= \ &\mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}^\perp\mathbf{S}_{d-k}{}^T\mathbf{A}_{d-k}{}^T \\
&+\sigma^2(N-k)\mathbf{W}\mathbf{W}^T
\end{aligned}$$

where we have used the fact that the elements of $\mathbf{V}$ are i.i.d. zero-mean Gaussian random variables. The important point is that all the terms in the quantity $\sigma^2(N-k)\mathbf{W}\mathbf{W}^T$ are known. Thus

$$\begin{aligned}
&\mathcal{E}\left\{\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)^T\right\} - \sigma^2(N-k)\mathbf{W}\mathbf{W}^T \\
&\quad = \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}^\perp\mathbf{S}_{d-k}{}^T\mathbf{A}_{d-k}{}^T,
\end{aligned}$$

so that the eigendecomposition (with eigenvalues in decreasing order on the diagonal) of

$$\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)^T - \sigma^2(N-k)\mathbf{W}\mathbf{W}^T$$

can be used to estimate the range space of $\mathbf{A}_{d-k}$.

Define the eigendecomposition as follows:

$$\begin{aligned}
&\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)\left(\mathbf{X} - \hat{\mathbf{A}}_k\mathbf{S}_k\right)^T - \sigma^2(N-k)\mathbf{W}\mathbf{W}^T \\
&\equiv \left[\mathbf{U}_{d-k}\mathbf{U}_{d-k}^\perp\right]\begin{bmatrix}\mathbf{\Sigma}_{d-k} & \\ & \mathbf{\Sigma}_{d-k}^\perp\end{bmatrix}\begin{bmatrix}\mathbf{U}_{d-k}^T \\ \mathbf{U}_{d-k}^\perp{}^T\end{bmatrix}
\end{aligned}$$

Then

$$\mathbf{U}_{d-k}\mathbf{\Sigma}_{d-k}\mathbf{U}_{d-k}^T \cong \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}^\perp\mathbf{S}_{d-k}{}^T\mathbf{A}_{d-k}{}^T,$$

so that $\text{Range}(\mathbf{U}_{d-k})$ is an estimate of $\text{Range}(\mathbf{A}_{d-k})$. For consistency with the notation of the papers [7, 8], we use the notation $\hat{\mathcal{N}}_k$ to denote an orthonormal basis spanning $\text{Range}(\mathbf{A}_{d-k})$, so that

$$\begin{aligned}
&\text{Range}(\mathbf{U}_{d-k}) \equiv \text{Range}(\hat{\mathcal{N}}_k) \\
&\cong \text{Range}(\mathcal{N}_k) \equiv \text{Range}(\mathbf{A}_{d-k}).
\end{aligned} \quad (2)$$

3 DEFLATION

In this section, we describe the deflation procedure and analyze the effects of errors in our estimates upon the deflated data. Using two approximations, we arrive at a simple equation describing the deflated data. We will use an equation from reference [9] to form an oblique projection operator $\mathbf{E}$ which will be used to perform the deflation.

We desire the range of $\mathbf{E}$ to be $\text{Range}(\mathcal{N}_k)$, the range of $\mathbf{A}_{d-k}$; however we must use estimate $\text{Range}(\hat{\mathcal{N}}_k)$ instead. We desire the null space of $\mathbf{E}$ to be the range of $\mathbf{A}_k$; we must instead use the estimate $\text{Range}(\hat{\mathbf{A}}_k)$. Plugging these values into the equation from [9] yields

$$\mathbf{E} \equiv \left[\hat{\mathcal{N}}_k\ \mathbf{0}\right]\begin{bmatrix}\mathbf{I} & \hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k \\ \hat{\mathbf{A}}_k^T\hat{\mathcal{N}}_k & \hat{\mathbf{A}}_k^T\hat{\mathbf{A}}_k\end{bmatrix}^{-1}\begin{bmatrix}\hat{\mathcal{N}}_k{}^T \\ \hat{\mathbf{A}}_k^T\end{bmatrix}.$$

After obliquely projecting the data with $\mathbf{E}$ it will reside in a $(d-k)$ dimensional subspace, so we will reduce the dimension of the problem by multiplying by $\hat{\mathcal{N}}_k{}^T$. This yields

$$\hat{\mathcal{N}}_k{}^T\mathbf{E} = [\mathbf{I}\ \mathbf{0}]\begin{bmatrix}\mathbf{I} & \hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k \\ \hat{\mathbf{A}}_k^T\hat{\mathcal{N}}_k & \hat{\mathbf{A}}_k^T\hat{\mathbf{A}}_k\end{bmatrix}^{-1}\begin{bmatrix}\hat{\mathcal{N}}_k{}^T \\ \hat{\mathbf{A}}_k^T\end{bmatrix} \quad (3)$$

so it will only be necessary to calculate the quantities in the first row of the inverse matrix in equation 3. We can use a formula for the inverse of a block matrix from [10]. To do so requires the computation of the Schur complement Δ. In our case,

$$\Delta \equiv \hat{\mathbf{A}}_k^T\hat{\mathbf{A}}_k - \hat{\mathbf{A}}_k^T\hat{\mathcal{N}}_k\hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k \equiv \hat{\mathbf{A}}_k^T\mathbf{P}_{\hat{\mathcal{N}}_k}^\perp\hat{\mathbf{A}}_k.$$

The (1,1) term of the block matrix is $(\mathbf{I} + \hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k\Delta^{-1}\hat{\mathbf{A}}_k^T\hat{\mathcal{N}}_k)$. The (1,2) term of the block matrix is $(-\hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k\Delta^{-1})$. Thus equation 3 becomes

$$\hat{\mathcal{N}}_k{}^T\mathbf{E} = \hat{\mathcal{N}}_k{}^T - \hat{\mathcal{N}}_k{}^T\hat{\mathbf{A}}_k\Delta^{-1}\hat{\mathbf{A}}_k^T\mathbf{P}_{\hat{\mathcal{N}}_k}^\perp.$$

We now make two approximations. The first approximation is that the error terms in $\hat{\mathbf{A}}_k$ are dominated by the correlation between $\mathbf{S}_{d-k}$ and $\mathbf{S}_k$ so that we can neglect the noise term in equation 1. This yields

$$\hat{\mathbf{A}}_k \cong \mathbf{A}_k + \mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{S}_k{}^{\dagger}.$$

The second approximation has been previously introduced in equation 2 and merely states that Range($\hat{\mathcal{N}}_k$) is a good approximation to Range($\mathbf{A}_{d-k}$). Thus we get the approximations

$$\begin{aligned} \mathbf{P}^{\perp}_{\hat{\mathcal{N}}_k}\mathbf{A}_{d-k} &\cong \mathbf{0}, \\ \mathbf{P}^{\perp}_{\hat{\mathcal{N}}_k}\hat{\mathbf{A}}_k &\cong \mathbf{P}^{\perp}_{\hat{\mathcal{N}}_k}\mathbf{A}_k, \\ \text{and} & \\ \Delta &\cong \mathbf{A}_k{}^T\mathbf{P}^{\perp}_{\hat{\mathcal{N}}_k}\mathbf{A}. \end{aligned}$$

We now use these approximations in the product

$$\begin{aligned} & \hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{A} \\ = \;& \hat{\mathcal{N}}_k{}^T\left[[\mathbf{A}_k \;\; \mathbf{A}_{d-k}] - \hat{\mathbf{A}}_k\Delta^{-1}\hat{\mathbf{A}}_k^T\mathbf{P}^{\perp}_{\hat{\mathcal{N}}_k}[\mathbf{A}_k \;\; \mathbf{A}_{d-k}]\right] \\ \cong \;& \hat{\mathcal{N}}_k{}^T\left[[\mathbf{A}_k \;\; \mathbf{A}_{d-k}] - \hat{\mathbf{A}}_k\Delta^{-1}[\Delta \;\; \mathbf{0}]\right] \\ = \;& \hat{\mathcal{N}}_k{}^T\left[(\mathbf{A}_k - \hat{\mathbf{A}}_k) \;\; \mathbf{A}_{d-k}\right] \\ \cong \;& \hat{\mathcal{N}}_k{}^T\left[(-\mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{S}_k{}^{\dagger}) \;\; \mathbf{A}_{d-k}\right] \\ = \;& \hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k}\left[(-\mathbf{S}_{d-k}\mathbf{S}_k{}^{\dagger}) \;\; \mathbf{I}\right] \end{aligned}$$

Multiplying this quantity by $\mathbf{S}$ yields

$$\begin{aligned} & \hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{A}\mathbf{S} \\ \cong \;& \hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k}\left[-\mathbf{S}_{d-k}\mathbf{S}_k{}^{\dagger} \;\; \mathbf{I}\right]\begin{bmatrix}\mathbf{S}_k \\ \mathbf{S}_{d-k}\end{bmatrix} \\ = \;& (\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k})\mathbf{S}_{d-k} - (\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}). \end{aligned}$$

Finally, we arrive at the product

$$\begin{aligned} & \hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{X} \\ = \;& \hat{\mathcal{N}}_k{}^T\mathbf{E}[\mathbf{A}\mathbf{S} + \mathbf{W}\mathbf{V}] \\ \cong \;& (\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k})\mathbf{S}_{d-k} + (\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})\mathbf{V} \\ & -(\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k}). \end{aligned} \tag{4}$$

Thus, except for the error term $(\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k}\mathbf{S}_{d-k}\mathbf{P}_{\mathbf{S}_k})$, the linear operator $(\hat{\mathcal{N}}_k{}^T\mathbf{E})$ converts the problem $\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{W}\mathbf{V}$ into a lower-dimensional problem of similar form. This lower-dimensional problem can now be input into the signal-identifying algorithm for further processing. Note that the error term is due to the unknown correlation between $\mathbf{S}_{d-k}$ and $\mathbf{S}_k$. We will see in the next section that even though we do not know this correlation, we can still compensate for its effects when testing the next estimated signal $\hat{\mathbf{s}}_{k+1}$.

4 A SIMPLE VALIDITY TEST FOR $\hat{\mathbf{s}}^T_{K+1}$

In this section, we derive a simple test for the hypothesis that we have successfully estimated the next signal $\mathbf{s}^T_{k+1}$. We assume that we have input the reduced problem 4 into an algorithm which returns two quantities: $\boldsymbol{\alpha}^T$ and $\hat{\mathbf{s}}^T_{k+1}$. The quantity $\hat{\mathbf{s}}^T_{k+1}$ is an estimate of one of the rows (signals) of $\mathbf{S}_{d-k}$, where we have assumed without loss of generality that row $k+1$ of $\mathbf{S}$ is found. The quantity $\boldsymbol{\alpha}^T$ is an estimate of the first row of $(\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k})^{-1}$. (We remark that the algorithms described in [7, 8] return exactly these quantities; they are the motivating examples for this paper.)

We now form the hypothesis that $\hat{\mathbf{s}}^T_{k+1} = \mathbf{s}^T_{k+1}$ and that $\boldsymbol{\alpha}^T$ equals the first row of $(\hat{\mathcal{N}}_k{}^T\mathbf{A}_{d-k})^{-1}$. Under this hypothesis we get

$$\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{X} \cong \mathbf{s}^T_{k+1} + (\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})\mathbf{V} - \mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k}.$$

Subtracting the decoded signal $\mathbf{s}^T_{k+1}$ from this quantity yields the residual vector $\mathbf{r}^T$:

$$\mathbf{r}^T = -\mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k} + (\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})\mathbf{V}.$$

The expected value of the sum of the squared residuals is

$$\begin{aligned} \mathcal{E}\{\mathbf{r}^T\mathbf{r}\} \;=\; & \mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k}\mathbf{s}_{k+1} \\ & +(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})\mathcal{E}\{\mathbf{V}\mathbf{V}^T\}(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})^T. \end{aligned}$$

We have arrived at the final important point: we know (or assume by hypothesis) the quantities $\mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k}\mathbf{s}_{k+1}$ and $(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})$. Note that the quantity $(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})\mathbf{V}$ is a row vector of zero-mean i.i.d. Gaussian random variables with variance $\sigma^2(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})^T$. We can thus form the test statistic

$$\frac{\mathbf{r}^T\mathbf{r} - \mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k}\mathbf{s}_{k+1}}{\sigma^2(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})(\boldsymbol{\alpha}^T\hat{\mathcal{N}}_k{}^T\mathbf{E}\mathbf{W})^T} \tag{5}$$

which under our hypothesis will be distributed as $\chi^2(N)$. Thus, even though we do not know the correlation between $\mathbf{S}_k$ and $\mathbf{S}_{d-k}$, we can form a simple test statistic to determine whether the signal-extracting algorithm has converged to a correct estimate of $\mathbf{s}^T_{k+1}$.

5 DISCUSSION AND CONCLUSION

The deflation approach and test statistic described in this paper was used successfully in the algorithms described in references [7, 8]. The combined algorithm was faster than existing simultaneous-estimation algorithms while paying only a modest penalty in BER for using its sequential approach. The simulations performed in references [7, 8] demonstrate that the test statistic was robust with respect to modest violations of the assumptions used in this paper. For example, in this paper we assumed that the previous k signals $\mathbf{S}_k$ were estimated without error; in practice, however, the approach will still yield approximate results when $\hat{\mathbf{S}}_k$ contains some symbol errors. We also assumed that the noise term in equation 1 was negligible with respect to the correlation error term. Violation of this assumption in practice is not a problem since in that case the correction term $\mathbf{s}^T_{k+1}\mathbf{P}_{\mathbf{S}_k}\mathbf{s}_{k+1}$ in equation 5 is small with respect to the total residual power $\mathbf{r}^T\mathbf{r}$. In conclusion, the deflation approach and test statistic described in this paper can be successfully combined with a sequential-estimation algorithm to create a fast and effective digital source separation algorithm.

REFERENCES

[1] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36:287–314, 1994.

[2] A. Belouchrani and J-F. Cardoso. Maximum likelihood source separation for discrete sources. In *Proc. EUSIPCO*, pages 768–771, Edinburgh, September 1994.

[3] S. Talwar, M. Viberg, and A. Paulraj. Blind estimation of multiple co-channel digital signals using an antenna array. *IEEE Signal Processing Letters*, 1(2):29–31, February 1994.

[4] S. Talwar, M. Viberg, and A. Paulraj. Blind separation of synchronous co-channel digital signals using an antenna array—part I: Algorithms. *IEEE Transactions on Signal Processing*, 44(5):1184–1197, May 1996.

[5] K. Anand, G. Mathew, and V. U. Reddy. Blind separation of multiple co-channel BPSK signals arriving at an antenna array. *IEEE Signal Processing Letters*, 2(9):176–178, September 1995.

[6] B. Halder, B. C. Ng, A. Paulraj, and T. Kailath. Unconditional maximum likelihood approach for blind estimation of digital signals. In *Proc. ICASSP'96*, pages 1081–1084, Atlanta, GA, May 1996.

[7] L. K. Hansen and G. Xu. A fast sequential source separation algorithm for digital co-channel signals. Accepted for publication in the *IEEE Signal Processing Letters*.

[8] L. K. Hansen and G. Xu. A fast algorithm for the blind separation of digital co-channel signals. In *Proc. 30^{th} Asilomar Conf. on Signals, Systems, and Comp.*, Pacific Grove, CA, November 1996.

[9] R. T. Behrens and L. L. Scharf. Signal processing applications of oblique projection operators. *IEEE Transactions on Signal Processing*, 42(6):1413–1424, June 1994.

[10] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1980.