Multi-Channel Signal Separation by Decorrelation

Ehud Weinstein, Meir Feder, and Alan V. Oppenheim

***Abstract*—In a variety of contexts, observations are made of the outputs of an unknown multiple-input multiple-output linear system, from which it is of interest to identify the unknown system and to recover the input signals. This often arises, for example, with speech recorded in an acoustic environment in the presence of background noise or competing speakers, in passive sonar applications, and in data communications in the presence of cross-coupling effects between the transmission channels. In this paper we specifically consider the two-channel case in which we observe the outputs of a 2 × 2 linear time invariant system. Our approach consists of reconstructing the input signals by assuming that they are statistically uncorrelated and imposing this constraint on the signal estimates. In order to restrict the set of solutions, additional information on the true signal generation and/or on the form of the coupling systems is incorporated. Specific algorithms are developed and tested. As a special case, these algorithms suggest a potentially interesting modification of Widrow's least-squares method for noise cancellation, when the reference signal contains a component of the desired signal.**

I. Introduction

In a variety of contexts, observations are made of the outputs of an unknown multiple-input multiple-output linear system from which it is of interest to identify the system and to recover its input signals. For example, in problems of enhancing speech in the presence of background noise, or separating competing speakers, multiple microphone measurements will typically have components from both sources, with the linear system representing the acoustic environment. Similar problems occur in passive sonar applications, and in data communication in the presence of cross-coupling effects between the channels.

In this paper we consider specifically the two-channel case, illustrated in Fig. 1, in which we observe the outputs $y_1(t)$ and $y_2(t)$ of a 2 × 2 linear time invariant (LTI) system with inputs $s_1(t)$ and $s_2(t)$, and with frequency response

$$\mathcal{H}(\omega) = \begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{bmatrix} \tag{1}$$

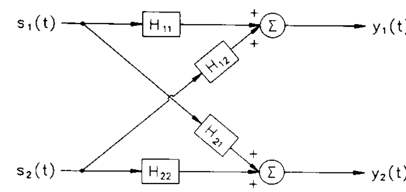


Fig. 1. The two-channel model.

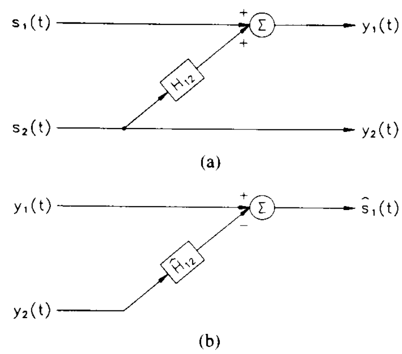


Fig. 2. The least squares method.

where H_{11} and H_{22} represent the transfer functions of each channel separately, and H_{12} and H_{21} represent the cross-coupling effects between the channels.

The most widely used approach to the two-channel signal enhancement, or separation, problem was suggested by Widrow *et al.* [13]. In their approach it is assumed that H_{11} and H_{22} are identity systems and H_{21} is zero, as illustrated in Fig. 2(a) in which case $s_2(t)$ (the interfering signal) is coupled into the first (primary) sensor through the unknown system H_{12} whose input is the signal measured by the second (reference) sensor. It is suggested in [13] that the unknown system be identified by minimizing the average power of the reconstructed, or estimated, signal and use it for cancellation of the interfering signal component at the primary sensor, as illustrated in Fig. 2(b). Minimizing the average power corresponds to identifying, or estimating, the unknown system H_{12} by a least-squares fit of the second (reference) sensor signal to the primary sensor signal. This method will therefore be referred to as the least squares (LS) method. Recursive and sequential/adaptive schemes based on the least mean squares (LMS) and the recursive least squares (RLS) algorithms have been proposed in [13] and in, e.g., [8], respectively.

The LS method has been successful in a wide variety of contexts. However, a critical assumption in that approach

Manuscript received February 9, 1991; revised May 29, 1993. This work was supported in part by the Wolfson Research Awards administered by the Israel Academy of Sciences and Humanities, in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-91-0034, and in part by the Office of Naval Research under Grant N00014-91-J-1628 and Grant N00014-93-1-0686. The associate editor coordinating the review of this paper and approving it for publication was Dr. B. A. Hanson.

Ehud Weinstein is with the Department of Electrical Engineering—Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, 69978, Israel. He is also an adjunct scientist at the Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

Meir Feder is with the Department of Electrical Engineering—Systems, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, 69978, Israel.

Alan V. Oppenheim is with the Department of Electrical Engineering and Computer Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9210882.

is that there is no leakage of the desired signal $s_1(t)$ into the reference sensor. If both signals are coupled into each sensor, then with the LS approach a portion of the desired signal is typically canceled together with the interfering signal. Since the desired signal may be canceled with some delay, or change in phase, it introduces a reverberant distortion in the reconstructed signal.

An approach to the two channel signal enhancement problem when both signals are coupled into each channel is presented in [3]. In that approach it is assumed that $s_1(t)$ (the desired signal) is a Gaussian autoregressive (AR) process, and $s_2(t)$ (the interfering signal) is uncorrelated white Gaussian noise. Both coupling systems H_{12} and H_{21} may be non-zero, and are modeled as discrete-time finite impulse response (FIR) filters. The problem is formulated as a maximum likelihood (ML) estimation problem, and the iterative estimate-maximize (EM) algorithm is used for its solution. Recursive and sequential algorithms based on this approach are developed in [12].

In [2], the specific problem in which the signals are coupled into each channel by a constant gain is considered. This problem occurs in satellite communication, when transmitting independent signals (messages) within the same frequency band and in orthogonal polarization in order to increase the capacity of the communication link. Due to the propagation conditions in the medium, some degree of cross-coupling between the channels will occur. It is suggested in [2] to use a cross-coupled cancellation system, consisting of two decoupling gains, for separating the signals. In order to match the decoupling gains to the unknown coupling gains, it is suggested to use pilot tones (i.e., reference signals) in each channel. Following this work, three different approaches named power–power, correlation–correlation, and power–correlation, were suggested in [1] in order to identify the coupling gains without the need for reference signals. In the power–power approach, the gains are adjusted by iteratively minimizing the average powers of the reconstructed signals; in the correlation–correlation approach, the gains are adjusted by minimizing the cross-correlation between the reconstructed signals at the same time instant (i.e., at lag zero); and the power–correlation approach is the combination of the two. However, as pointed out in [1], all these approaches are insufficient to solve the problem (all of them lead to a single equation that is insufficient to solve for the two unknown gains), and additional processing, referred to in [1] as a "signal discrimination network" is suggested. An alternative approach based on minimizing the cross-correlation at lag zero between nonlinear functions of the reconstructed signals is proposed in [4]. By using two different combinations of nonlinear functionals, it is possible to identify both coupling gains without the need for a discrimination network. However, this approach is also restricted to the special case in which the coupling systems are constant gains.

In this paper, we consider the more general case in which both coupling systems are possibly non-zero frequency dependent unknown LTI filters. Several approaches are presented for solving the problem, and the relation to previous work is indicated. The work presented in this paper has several potential applications as indicated in [10].

II. Signal Separation by Decorrelation

To simplify the exposition we first consider the case in which H_{11} and H_{22} are unity transformations, i.e., $H_{11}(\omega) = H_{22}(\omega) = 1$ for all ω. Although this case is somewhat restrictive, it represents the important and interesting problem in which the desired signals $s_1(t)$ and $s_2(t)$ are the signals measured at each sensor output in the absence of the other source signal, and the systems H_{12} and H_{21} represent the coupling effects. The more general case will be considered in the sequel. With H_{11} and H_{22} as unity transformations, the observed signals $y_1(t)$ and $y_2(t)$ are the outputs of a 2×2 LTI system with inputs $s_1(t)$ and $s_2(t)$, and with frequency response

$$\mathcal{H}(\omega) = \begin{bmatrix} 1 & H_{12}(\omega) \\ H_{21}(\omega) & 1 \end{bmatrix}. \tag{2}$$

We may assume that

$$1 - H_{12}(\omega)H_{21}(\omega) \neq 0 \quad \forall \omega. \tag{3}$$

If this assumption is not satisfied, then $\mathcal{H}$ is not invertible, in which case $s_1(t)$ and $s_2(t)$ could not be recovered from $y_1(t)$ and $y_2(t)$ even if $\mathcal{H}$ is known. We note that in case of environments subject to reverberant or multipath effects, the inverse of $\mathcal{H}$ may be ill-conditioned.

We shall assume that $s_1(t)$ and $s_2(t)$ are sample functions from statistically uncorrelated wide-sense stationary random processes. To simplify the exposition we shall further assume that $s_1(t)$ and $s_2(t)$ are zero-mean, so that the assumption that they are uncorrelated is equivalent to

$$E\{s_1(t)s_2^*(t-\tau)\} = 0 \quad \forall \tau \tag{4}$$

where $E\{\cdot\}$ stands for expectation, and $*$ denotes the complex conjugate. It should be noted that the derivation and results apply equally to the more general case of non-zero and possibly time-varying means, since they can be phrased in terms of convariances. The important case in which the signals may be nonstationary will be addressed in Section III.

If the systems H_{12} and H_{21} were known, then $s_1(t)$ and $s_2(t)$ could be recovered from $y_1(t)$ and $y_2(t)$ by inverse filtering. However, since in most cases of interest H_{12} and H_{21} are unknown, we need a method or a criterion for identifying or estimating them.

Our approach consists of finding estimates $\hat{H}_{12}$ and $\hat{H}_{21}$ of H_{12} and H_{21}, respectively, so that by performing inverse filtering the estimated, or reconstructed, signals $\hat{s}_1(t)$ and $\hat{s}_2(t)$ are statistically uncorrelated, that is,

$$E\{\hat{s}_1(t)\hat{s}_2^*(t-\tau)\} = 0 \quad \forall \tau. \tag{5}$$

It should be stressed that even if the signals $s_1(t)$ and $s_2(t)$ are assumed to be statistically uncorrelated, it does not imply that a selected estimation criterion such as the minimum mean square error criterion will generate statistically uncorrelated signal estimates. The key idea of our approach is to turn the assumption that the signals are uncorrelated into the estimation criterion. As we will show, the solution based on this criterion reduces exactly to the LS solution in the simplified case considered in [13].

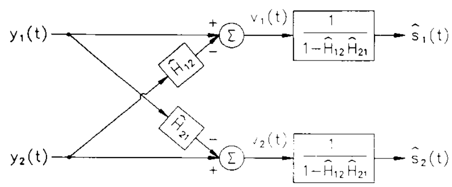


Fig. 3. The reconstruction system.

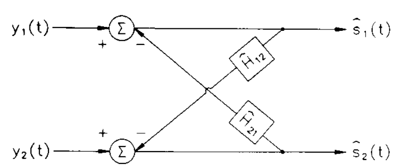


Fig. 4. An alternative implementation of the reconstruction system.

As illustrated in Fig. 3, $\hat{s}_1(t)$ and $\hat{s}_2(t)$ are the output of the 2 × 2 LTI system with inputs $y_1(t)$ and $y_2(t)$, and with frequency response

$$\hat{\mathcal{H}}^{-1}(\omega) = \frac{1}{1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)} \begin{bmatrix} 1 & -\hat{H}_{12}(\omega) \\ -\hat{H}_{21}(\omega) & 1 \end{bmatrix} \tag{6}$$

where, as in (3), it is assumed that

$$1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega) \neq 0 \quad \forall\omega. \tag{7}$$

An alternative system for generating $\hat{s}_1(t)$ and $\hat{s}_2(t)$ is illustrated in Fig. 4.

Using the well-known relationship for the power spectra between inputs and outputs of an LTI system,

$$\begin{bmatrix} P_{\hat{s}_1\hat{s}_1}(\omega) & P_{\hat{s}_1\hat{s}_2}(\omega) \\ P_{\hat{s}_2\hat{s}_1}(\omega) & P_{\hat{s}_2\hat{s}_2}(\omega) \end{bmatrix} = \frac{1}{|1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)|^2} \cdot \begin{bmatrix} 1 & -\hat{H}_{12}(\omega) \\ -\hat{H}_{21}(\omega) & 1 \end{bmatrix} \begin{bmatrix} P_{y_1y_1}(\omega) & P_{y_1y_2}(\omega) \\ P_{y_2y_1}(\omega) & P_{y_2y_2}(\omega) \end{bmatrix} \cdot \begin{bmatrix} 1 & -\hat{H}^*_{21}(\omega) \\ -\hat{H}^*_{12}(\omega) & 1 \end{bmatrix} \tag{8}$$

where $P_{\hat{s}_i\hat{s}_j}(\omega)$ $i, j = 1, 2$ are the auto- and cross-spectra of $\hat{s}_1(t)$ and $\hat{s}_2(t)$, and $P_{y_iy_j}(\omega) i, j = 1, 2$ are the auto- and cross-spectra of $y_1(t)$ and $y_2(t)$.

Performing the matrix multiplication in (8), and observing that the decorrelation condition in (5) implies that $P_{\hat{s}_1\hat{s}_2}(\omega) = 0$ for all ω, we obtain

$$P_{y_1y_2}(\omega) - \hat{H}_{12}(\omega)P_{y_2y_2}(\omega) - \hat{H}^*_{21}(\omega)P_{y_1y_1}(\omega) + \hat{H}_{12}(\omega)\hat{H}^*_{21}(\omega)P_{y_2y_1}(\omega) = 0 \tag{9}$$

where we note that $P_{y_iy_j}(\omega) i, j = 1, 2$ would in practice be estimated from the measured signals $y_1(t)$ and $y_2(t)$. Any combination of $\hat{H}_{12}$ and $\hat{H}_{21}$ that satisfies (9) results in signal estimates $\hat{s}_1(t)$ and $\hat{s}_2(t)$ that are statistically uncorrelated. Clearly, this equation does not specify a unique solution for $\hat{H}_{12}$ and $\hat{H}_{21}$. We could arbitrarily choose $\hat{H}_{21}$, in which case $\hat{H}_{12}$ is specified by

$$\hat{H}_{12}(\omega) = \frac{P_{y_1y_2}(\omega) - \hat{H}^*_{21}(\omega)P_{y_1y_1}(\omega)}{P_{y_2y_2}(\omega) - \hat{H}^*_{21}(\omega)P_{y_2y_1}(\omega)} \tag{10}$$

or we could arbitrarily choose $\hat{H}_{12}$, in which case $\hat{H}_{21}$ is specified by

$$\hat{H}_{21}(\omega) = \frac{P_{y_2y_1}(\omega) - \hat{H}^*_{12}(\omega)P_{y_2y_2}(\omega)}{P_{y_1y_1}(\omega) - \hat{H}^*_{12}(\omega)P_{y_1y_2}(\omega)}. \tag{11}$$

As a special case, if we choose $\hat{H}_{21} = 0$ then (10) reduces to

$$\hat{H}_{12}(\omega) = \frac{P_{y_1y_2}(\omega)}{p_{y_2y_2}(\omega)}. \tag{12}$$

This corresponds exactly to the LS solution for the simplified scenario in which it is assumed that there is no coupling of the desired signal into the reference sensor, i.e., when the actual coupling system H_{21} is zero. Thus, the LS solution is one of many possible solutions of the decorrelation equation. It is consistent with the observation in [7] that the LS method causes the desired signal estimate to be statistically uncorrelated with the reference sensor signal. The LS method has been successful in a wide variety of contexts. However, it is widely recognized that if the assumption of zero coupling is not satisfied, its performance may seriously deteriorate. Eq. (10) therefore suggests a potentially interesting alternative to the LS method that takes into account a pre-specified non-zero coupling.

To determine the relation between the solutions of the decorrelation equation and the actual (true) coupling filters H_{12} and H_{21}, we use the relation

$$\begin{bmatrix} P_{y_1y_1}(\omega) & P_{y_1y_2}(\omega) \\ P_{y_2y_1}(\omega) & P_{y_2y_2}(\omega) \end{bmatrix} = \begin{bmatrix} 1 & H_{12}(\omega) \\ H_{21}(\omega) & 1 \end{bmatrix} \cdot \begin{bmatrix} P_{s_1s_1}(\omega) & 0 \\ 0 & P_{s_2s_2}(\omega) \end{bmatrix} \begin{bmatrix} 1 & H^*_{21}(\omega) \\ H^*_{12}(\omega) & 1 \end{bmatrix} \tag{13}$$

where $P_{s_1s_1}(\omega)$ and $P_{s_2s_2}(\omega)$ are the power spectra of $s_1(t)$ and $s_2(t)$, respectively. Substituting (13) into (9) and carrying out straightforward algebraic manipulations, we obtain

$$P_{s_1s_1}(\omega)[1 - \hat{H}_{12}(\omega)H_{21}(\omega)][H_{21}(\omega) - \hat{H}_{21}(\omega)]^* + P_{s_2s_2}(\omega)[1 - \hat{H}_{21}(\omega)H_{12}(\omega)]^*[H_{12}(\omega) - \hat{H}_{12}(\omega)] = 0 \tag{14}$$

If $\hat{H}_{21}(\omega) = H_{21}(\omega)$, then $\hat{H}_{12}(\omega) = H_{12}(\omega)$, provided that $P_{s_2s_2}(\omega)$ is strictly positive and that the condition in (3) is satisfied. Similarly, if $\hat{H}_{12}(\omega) = H_{12}(\omega)$, then $\hat{H}_{21}(\omega) = H_{21}(\omega)$, provided that $P_{s_1s_1}(\omega)$ is strictly positive. Thus, if one of the coupling filters is known, then the decorrelation criterion yields the correct solution for the other coupling filter.

There are practical situations in which one of the coupling systems is known *a priori* or can be measured independently. For example, in speech enhancement, either the desired speech signal or the interfering signal may be in a fixed location and therefore the acoustic transfer functions that couple it to the microphones can be measured *a priori*. In such cases either (10) or (11) can be used to find the other coupling system. Another interesting application is the problem of separating competing speakers. By identifying a quiet period for one of the speakers, the acoustic transfer function with respect to the other speaker can be estimated separately and then used to identify the unknown transfer function when both speakers are active.

A common measure of performance is the ratio of the power spectrum of the desired signal component to the power spectrum of the residual, or interfering, signal component. Since the estimates $\hat{s}_1(t)$ and $\hat{s}_2(t)$ are the outputs of the 2×2 system with inputs $s_1(t)$ and $s_2(t)$, and with the frequency response

$$\mathcal{H}(\omega)\hat{\mathcal{H}}^{-1}(\omega) = \frac{1}{1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)} \cdot \begin{bmatrix} 1 - \hat{H}_{12}(\omega)H_{21}(\omega) & \hat{H}_{12}(\omega) - H_{12}(\omega) \\ H_{21}(\omega) - \hat{H}_{21}(\omega) & 1 - H_{12}(\omega)\hat{H}_{21}(\omega) \end{bmatrix} \quad (15)$$

then the post-processing signal-to-interference ratio (S/I) at the first and second sensors are

$$(S/I)_1 = \frac{|1 - \hat{H}_{12}(\omega)H_{21}(\omega)|^2}{|\hat{H}_{12}(\omega) - H_{12}(\omega)|^2} \frac{P_{s_1s_1}(\omega)}{P_{s_2s_2}(\omega)} \quad (16)$$

$$(S/I)_2 = \frac{|1 - H_{12}(\omega)\hat{H}_{21}(\omega)|^2}{|\hat{H}_{21}(\omega) - H_{21}(\omega)|^2} \frac{P_{s_2s_2}(\omega)}{P_{s_1s_1}(\omega)} \quad (17)$$

where in (16) the interference is the signal component involving $s_2(t)$, and in (17) the interference is the signal component involving $s_1(t)$. If $\hat{H}_{12}(\omega) = H_{12}(\omega)$ and $\hat{H}_{21}(\omega) = H_{21}(\omega)$, we obtain infinite S/I at both sensor outputs, as we would expect.

Using the relation in (14), it can easily be verified that

$$(S/I)_1 = (S/I)_2. \quad (18)$$

This result is consistent with the power inversion principle associated with the LS method. If we regard $s_1(t)$ as the desired signal and $s_2(t)$ as the interfering signal, then (18) asserts that the post-processing signal-to-interference ratio at the first (primary) sensor is equal to the interference-to-signal ratio at the second (reference) sensor. Since in the LS approach no processing is applied to the reference sensor output, the post-processing signal-to-interference ratio at the primary sensor is limited by the interference-to-signal ratio at the reference sensor, that is,

$$(S/I)_1 = \frac{1}{|H_{21}(\omega)|^2} \frac{P_{s_2s_2}(\omega)}{P_{s_1s_1}(\omega)}. \quad (19)$$

However, in our decorrelation approach, if H_{21} or a close estimate of it is provided, then using (10) a close estimate of H_{12} can be obtained, and the resulting S/I at the first (primary) sensor can be made very large as indicated by (16), even if H_{21} is non-zero.

We may also consider the more general case illustrated in Fig. 1 in which $y_1(t)$ and $y_2(t)$ are the outputs of a general 2×2 LTI system $\mathcal{H}$ with inputs $s_1(t)$ and $s_2(t)$, and with the frequency response given by (1). In this case we need to identify the four systems $H_{ij} i,j = 1,2$. We assume that $\mathcal{H}$ is invertible, and that H_{11} and H_{22} are also invertible. The estimated, or reconstructed, signals $\hat{s}_1(t)$ and $\hat{s}_2(t)$ are the outputs of the inverse filter $\hat{\mathcal{H}}^{-1}$ whose frequency response is

$$\hat{\mathcal{H}}^{-1}(\omega) = \frac{1}{\hat{H}_{11}(\omega)\hat{H}_{22}(\omega) - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)} \cdot \begin{bmatrix} \hat{H}_{22}(\omega) & -\hat{H}_{12}(\omega) \\ -\hat{H}_{21}(\omega) & \hat{H}_{11}(\omega) \end{bmatrix} \quad (20)$$

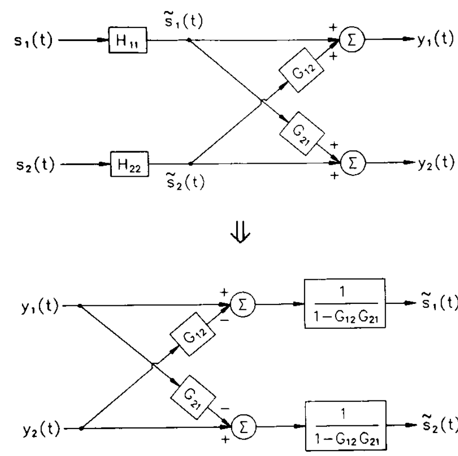


Fig. 5. The system for generating $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$.

where $\hat{H}_{ij}(\omega) i,j = 1,2$ denote the estimates of $H_{ij}(\omega) i,j = 1,2$. Requiring $\hat{s}_1(t)$ and $\hat{s}_2(t)$ to be uncorrelated, and following a development similar to that which led to (9), we obtain

$$P_{y_1y_2}(\omega) - \hat{G}_{12}(\omega)P_{y_2y_2}(\omega) - \hat{G}_{21}^*(\omega)P_{y_1y_1}(\omega) + \hat{G}_{12}(\omega)\hat{G}_{21}^*(\omega)P_{y_2y_1}(\omega) = 0 \quad (21)$$

where

$$\hat{G}_{12}(\omega) = \hat{H}_{12}(\omega)/\hat{H}_{22}(\omega) \quad (22)$$

and

$$\hat{G}_{21}(\omega) = \hat{H}_{21}(\omega)/\hat{H}_{11}(\omega). \quad (23)$$

Equation (21) is identical in form to (9), except that $\hat{H}_{12}$ and $\hat{H}_{21}$ are replaced by $\hat{G}_{12}$ and $\hat{G}_{21}$, respectively. Thus, even if $\hat{H}_{ij} = H_{ij}$ $i,j = 1,2$, i.e., $\mathcal{H}$ is known, then using the decorrelation method we can only identify the ratios $G_{12} = H_{12}/H_{22}$ and $G_{21} = H_{21}/H_{11}$. Consequently, we can only identify the signals $\tilde{s}_1(t) = H_{11}\{s_1(t)\}$ (i.e. the output of H_{11} with input $s_1(t)$) and $\tilde{s}_2(t) = H_{22}\{s_2(t)\}$, as illustrated in Fig. 5. This result is intuitively reasonable. If $s_1(t)$ and $s_2(t)$ are statistically uncorrelated, then $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ are also uncorrelated. Therefore, using the decorrelation criterion we cannot distinguish the pair $s_1(t)$ and $s_2(t)$ from the pair $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$. However, in some problems the recovery of the desired signals up to the shaping filters H_{11} and H_{22} may be sufficient. For example, in the problem of separating competing speakers, if one of the speakers is near one microphone and the other speaker is near the other microphone, then H_{11} and H_{22} are nearly unity transformations, and the recovery of $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$ may be sufficient for intelligibility.

III. Utilizing Additional Information for Signal Separation

If both coupling systems are unknown, the decorrelation criterion is insufficient to solve the problem, and some additional information or constraints are needed on the true signal generation and/or on the form of the coupling systems.

As discussed above, any solution resulting from the decorrelation criterion must satisfy (14). One possible solution is the desired solution:

$$\hat{H}_{12}(\omega) = H_{12}(\omega), \quad \hat{H}_{21}(\omega) = H_{21}(\omega) \tag{24}$$

in which case $\hat{s}_1(t) = s_1(t), \hat{s}_2(t) = s_2(t)$ and the signals are exactly recovered, provided that $\mathcal{H}$ is given by (2). Another desired solution is

$$\hat{H}_{12}(\omega) = 1/H_{21}(\omega), \qquad \hat{H}_{21}(\omega) = 1/H_{12}(\omega) \tag{25}$$

where we have assumed that H_{12} and H_{21} are invertible. It is easy to verify that in this case $\hat{s}_1(t) = H_{12}\{s_2(t)\}$ and $\hat{s}_2(t) = H_{21}\{s_1(t)\}$, and the source signals can be recovered either by applying $\hat{H}_{21} = H_{12}^{-1}$ to $\hat{s}_1(t)$ and $\hat{H}_{12} = H_{21}^{-1}$ to $\hat{s}_2(t)$, or by performing $s_1(t) = y_1(t) - \hat{s}_1(t)$ and $s_2(t) = y_2(t) - \hat{s}_2(t)$.

While the solutions given by (24) and (25) are both acceptable since they allow signal recovery, in general, unless additional information is given, other non-desired solution may exist as well. We shall present several approaches that may lead to the desired solutions.

A. Signal Separation based on Statistical Independence

In Section II we have assumed that the signals $s_1(t)$ and $s_2(t)$ are statistically uncorrelated. If the signals come from independent sources, then it may in fact be reasonable to assume that they are statistically independent, which is a stronger condition if the signals are not jointly Gaussian processes. By imposing statistical independence between the reconstructed signals $\hat{s}_1(t)$ and $\hat{s}_2(t)$, we obtain additional constraints involving high-order cross-cumulants/spectra. This idea is developed in [14], where it is shown that the only possible solutions in this case are the desired solutions given by (24) and (25).

B. Separation of Nonstationary Signals

In Section II we have assumed that the signals $s_1(t)$ and $s_2(t)$ are jointly stationary processes. However, in many practical situations, the signals are more typically nonstationary. If $s_1(t)$ and $s_2(t)$ are statistically uncorrelated but nonstationary we require that

$$E\{s_1(t)s_2^*(t-\tau)\} = 0 \quad \forall t, \tau. \tag{26}$$

This imposes a stronger condition on the reconstructed signals. Specifically, let us make the simplifying but often realistic assumption that the signals are quasi-stationary, so that if we divide the observation interval into sub-intervals then the cross-correlation is stationary (i.e., independent of the time origin t) over each sub-interval. By imposing the decorrelation condition over each sub-interval we obtain, in accordance with (9), the following set of equations:

$$\begin{aligned} &P_{y_1y_2}^{(m)}(\omega) - \hat{H}_{12}(\omega)P_{y_2y_2}^{(m)}(\omega) - \hat{H}_{21}^*(\omega)P_{y_1y_1}^{(m)}(\omega) \\ &\quad + \hat{H}_{12}(\omega)\hat{H}_{21}^*(\omega)P_{y_2y_1}^{(m)}(\omega) = 0 \\ &\quad m = 1, 2, \cdots, M \end{aligned} \tag{27}$$

where $P_{y_iy_j}^{(m)}(\omega) i, j = 1, 2$ are the auto- and cross-spectra of $y_1(t)$ and $y_2(t)$ associated with the mth sub-interval, and M is the number of sub-intervals. If $M = 2$, we obtain two equations that can be solved for the two unknown filters $\hat{H}_{12}$ and $\hat{H}_{21}$. If $M > 2$, we have more equations than unknowns, and we may use a least-squares solution. In practice, $P_{y_iy_j}^{(m)}(\omega)$ are estimated from the observed data so that if M is too large, each sub-interval may be too small, in which case we have an overdetermined set of equations, but each equation may have large statistical variance. Also, the decorrelation equations associated with adjacent time segments may be statistically related. To reflect that, we may use a weighted least-squares solution in which we give more weight to information provided by time segments that are far apart. Recursive and time-adaptive solutions can also be obtained in which the cross-correlation associated with the most current data segments is weighted more heavily. This may be useful in situations where the coupling filters are also slowly time varying, e.g., when there is relative motion between sources and receivers, and we want an adaptive algorithm that is capable of tracking the varying characteristics of the channel. All these issues must be explored in depth.

C. Signal Separation by Spectral Matching

In many problems of interest the detailed spectral properties of $s_1(t)$ and $s_2(t)$ may be known. This prior information can be exploited by matching the power spectra of the reconstructed signals to the known spectra, i.e., $P_{\hat{s}_i\hat{s}_j}(\omega) = P_{s_is_j}(\omega) i, j = 1, 2$. Performing the matrix multiplications in (8), we obtain the following set of equations:

$$\begin{aligned} &P_{y_1y_2}(\omega) - \hat{H}_{12}(\omega)P_{y_2y_2}(\omega) - \hat{H}_{21}^*(\omega)P_{y_1y_1}(\omega) \\ &\quad + \hat{H}_{12}(\omega)\hat{H}_{21}^*(\omega)P_{y_2y_1}(\omega) \\ &\quad = |1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)|^2 P_{s_1s_2}(\omega) \end{aligned} \tag{28}$$

$$\begin{aligned} &P_{y_1y_1}(\omega) - \hat{H}_{12}(\omega)P_{y_2y_1}(\omega) - \hat{H}_{12}^*(\omega)P_{y_1y_2}(\omega) \\ &\quad + |\hat{H}_{12}(\omega)|^2 P_{y_2y_2}(\omega) \\ &\quad = |1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)|^2 P_{s_1s_1}(\omega) \end{aligned} \tag{29}$$

$$\begin{aligned} &P_{y_2y_2}(\omega) - \hat{H}_{21}(\omega)P_{y_1y_2}(\omega) - \hat{H}_{21}^*(\omega)P_{y_2y_1}(\omega) \\ &\quad + |\hat{H}_{21}(\omega)|^2 P_{y_1y_1}(\omega) \\ &\quad = |1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)|^2 P_{s_2s_2}(\omega) \end{aligned} \tag{30}$$

If $P_{s_is_j}(\omega)$ $i, j = 1, 2$ are given, then (28)–(30) are sufficient for solving for both coupling filters (note that (29) and (30) are real-valued equations, and (28) is a complex-valued equation, where $\hat{H}_{12}$ and $\hat{H}_{12}$ are complex-valued variables). It is easy to verify that (24) is a solution to (28)–(30). Conditions under which this solution is unique can also be specified.

With this approach it may be possible to work with the more general system model and identify the direct paths H_{11} and H_{22}. We may also invoke the possible nonstationarity of the signals to obtain additional equations and to improve identifiability.

We note that (9) is a special case of (28), obtained by substituting $P_{s_1s_2}(\omega) = 0$. It therefore suggests a modification of the decorrelation approach in case $s_1(t)$ and $s_2(t)$ are not uncorrelated, but have a prespecified cross-correlation function.

D. Signal Separation Using Constraints on the Coupling Systems

An alternative approach to reducing the set of possible solutions to the decorrelation equation is to impose constraints on the form of the coupling systems. For example, if we restrict the coupling systems to be constant gains $H_{12}(\omega) = h_{12}$ and $H_{21}(\omega) = h_{21}$, which is the case considered in [1], [2], and [4], then, as shown in [11], the only two possible solutions to the decorrelation equation are given by (24) and (25) which in this case reduce to

$$\hat{h}_{12} = h_{12}, \quad \hat{h}_{21} = h_{21}$$
$$\hat{h}_{12} = 1/h_{21}, \quad \hat{h}_{21} = 1/h_{12}.$$

In fact, in this simplified case it is sufficient to use the decorrelation criterion in (5) at only two different values of τ to obtain two linearly independent equations for solving for the two unknown gains. The only condition required for the existence of these solutions is that the autocorrelation functions of $s_1(t)$ and $s_2(t)$ are not proportional to each other, i.e. that $s_1(t)$ and $s_2(t)$ can be distinguished based on their normalized autocorrelation functions. However, we note that the decorrelation condition only at $\tau = 0$, as suggested in [1], is insufficient to solve the problem.

A more general and certainly more interesting case is that in which the coupling systems are assumed to be discrete-time finite-impulse-response (FIR) filters. Under the FIR constraint the solution given by (25) is excluded. Sufficient conditions under which the solution in (24) is the only solution can be specified. However, we note that the FIR restriction is essential in order to obtain a unique solution, or at least a reduced set of solutions. As the number of FIR coefficients increases, the solutions become more and more ill-conditioned, and in the limit we may lose identifiability.

There are other structures for coupling systems that may lead to the desired solution. For example, in passive sonar applications, the coupling systems describe the propagation delays of the signals from the sources to the receiving sensors. As long as H_{12} and H_{21} can be described by a finite number of (unknown) parameters, we may obtain the desired solution or, at least, a reduced set of solutions to the decorrelation equation.

IV. Algorithm Development

In this section we present possible algorithms for signal separation based on the decorrelation criterion. In our development we let the decoupling systems $\hat{H}_{12}$ and $\hat{H}_{21}$ be discrete-time causal FIR filters, of the form

$$\hat{H}_{12}(\omega) = \sum_{k=0}^{q_1} a_k e^{-j\omega k} \tag{31}$$

$$\hat{H}_{21}(\omega) = \sum_{k=0}^{q_2} b_k e^{-j\omega k} \tag{32}$$

where q_1 and q_2 are some prespecified filter orders. Similar algorithms can be developed for other structures of the decoupling systems.

With the decoupling systems of the form (31) and (32), the estimated signals $\hat{s}_1(t)$ and $\hat{s}_2(t)$ can be generated either by (see Fig. 4)

$$\hat{s}_1(t) = y_1(t) - \sum_{k=0}^{q_1} a_k \hat{s}_2(t-k) \tag{33}$$

$$\hat{s}_2(t) = y_2(t) - \sum_{k=0}^{q_2} b_k \hat{s}_1(t-k) \tag{34}$$

or by (see Fig. 3)

$$v_1(t) = y_1(t) - \sum_{k=0}^{q_1} a_k y_2(t-k) \tag{35}$$

$$v_2(t) = y_2(t) - \sum_{k=0}^{q_2} b_k y_1(t-k) \tag{36}$$

where $\hat{s}_1(t)$ and $\hat{s}_2(t)$ are generated from $v_1(t)$ and $v_2(t)$ by

$$\sum_{k=0}^{q_1+q_2} c_k \hat{s}_i(t-k) = v_i(t) \quad i = 1, 2 \tag{37}$$

where

$$c_k = \delta_k - \sum_{l=0}^{k} a_l b_{k-l} \quad k = 0, 1, \cdots, (q_1 + q_2) \tag{38}$$

where δ_k is the Kronecker delta function.

We want to adjust the a_k's and the b_k's to satisfy the decorrelation condition in (9). We may consider the following iterative algorithm: For a given set of b_k's, adjust the a_k's to satisfy (10), and for a given set of a_k's, adjust the b_k's to satisfy (11). We shall find it convenient to express these equations in the time domain. To do so, we note that

$$\begin{bmatrix} P_{y_1 v_1}(\omega) & P_{y_1 v_2}(\omega) \\ P_{y_2 v_1}(\omega) & P_{y_2 v_2}(\omega) \end{bmatrix} = \begin{bmatrix} P_{y_1 y_1}(\omega) & P_{y_1 y_2}(\omega) \\ P_{y_2 y_1}(\omega) & P_{y_2 y_2}(\omega) \end{bmatrix} \cdot \begin{bmatrix} 1 & -\hat{H}_{21}^*(\omega) \\ -\hat{H}_{12}^*(\omega) & 1 \end{bmatrix} \tag{39}$$

where $P_{y_i v_j}(\omega)$ denotes the cross-spectrum between $y_i(t)$ and $v_j(t)$. Substituting (39) in (10) and (11), we obtain

$$P_{y_2 v_2}(\omega)\hat{H}_{12}(\omega) = P_{y_1 v_2}(\omega) \tag{40}$$

$$P_{y_1 v_1}(\omega)\hat{H}_{21}(\omega) = P_{y_2 v_1}(\omega). \tag{41}$$

Inverse Fourier transforming, we obtain

$$\sum_{k=0}^{q_1} a_k c_{y_2 v_2}(\tau - k) = c_{y_1 v_2}(\tau) \tag{42}$$

$$\sum_{k=0}^{q_2} b_k c_{y_1 v_1}(\tau - k) = c_{y_2 v_1}(\tau) \tag{43}$$

where $c_{y_i v_j}(\tau)$ is the cross-correlation function between $y_i(t)$ and $v_j(t)$, defined by

$$c_{y_i v_j}(\tau) = E\{y_i(t) v_j^*(t-\tau)\} \tag{44}$$

Expressing (42) for $\tau = 0, 1, \cdots, q_1$ and (43) for $\tau = 0, 1, 2, \cdots, q_2$, and concatenating the equations, we obtain

$$C_{\underline{y}_2\underline{v}_2}\underline{a} = \underline{c}_{y_1\underline{v}_2} \tag{45}$$

$$C_{\underline{y}_1\underline{v}_1}\underline{b} = \underline{c}_{y_2\underline{v}_1} \tag{46}$$

where

$$\underline{a} = (a_0 a_1 \cdots a_{q_1})^T \tag{47}$$

$$\underline{b} = (b_0 b_1 \cdots b_{q_2})^T \tag{48}$$

and

$$C_{\underline{y}_2\underline{v}_2} = E\{\underline{v}_2^*(t)\underline{y}_2^T(t)\} \tag{49}$$

$$\underline{c}_{y_1\underline{v}_2} = E\{\underline{v}_2^*(t)y_1(t)\} \tag{50}$$

$$C_{\underline{y}_1\underline{v}_1} = E\{\underline{v}_1^*(t)\underline{y}_1^T(t)\} \tag{51}$$

$$\underline{c}_{y_2\underline{v}_1} = E\{\underline{v}_1^*(t)y_2(t)\} \tag{52}$$

where

$$\underline{y}_1(t) = [y_1(t)y_1(t-1)\cdots y_1(t-q_2)]^T \tag{53}$$

$$\underline{y}_2(t) = [y_2(t)y_2(t-1)\cdots y_2(t-q_1)]^T \tag{54}$$

$$\underline{v}_1^*(t) = [v_1^*(t)v_1^*(t-1)\cdots v_1^*(t-q_2)]^T \tag{55}$$

$$\underline{v}_2^*(t) = [v_2^*(t)v_2^*(t-1)\cdots v_2^*(t-q_1)]^T \tag{56}$$

Equations (45) and (46) are the time domain equivalents of (10) and (11), respectively. We observe that $\underline{C}_{\underline{y}_2\underline{v}_2}$ and $\underline{c}_{y_1\underline{v}_2}$ depend only on $\underline{b}$, while $C_{\underline{y}_1\underline{v}_1}$ and $\underline{c}_{y_2\underline{v}_1}$ depend only on $\underline{a}$. Therefore, for any given $\underline{b}$, the solution for $\underline{a}$ is

$$\underline{a} = C_{\underline{y}_2\underline{v}_2}^{-1}\underline{c}_{y_1\underline{v}_2} \tag{57}$$

and for any given $\underline{a}$ the solution for $\underline{b}$ is

$$\underline{b} = C_{\underline{y}_1\underline{v}_1}^{-1}\underline{c}_{y_2\underline{v}_1}. \tag{58}$$

By alternating between (57) and (58) we obtain an iterative algorithm for adjusting both filter coefficients. Note that this is not the only algorithm for solving the decorrelation equation. The Gauss method or the Newton-Raphson or some other coordinate-search algorithm may exhibit better convergence behavior. We further note that if both $\underline{a}$ and $\underline{b}$ are unknown, undesired solution may exist and so it is not guarateed that the algorithm converges to the desired solution. In this case additional information/constraints may be incorporated in order to obtain a desired solution, based on the discussion in the previous section. Nevertheless, as we shall see, this algorithm leads to potentially interesting extensions of the least mean squares (LMS) and the recursive least squares (RLS) algorithms for the problem considered here.

Since the correlation functions appearing in (57) and (58) are unknown, they are approximated by their sample estimates:

$$C_{\underline{y}_2\underline{v}_2} \approx \sum_{t=1}^{N} \beta_1^{N-t}\underline{v}_2^*(t)\underline{y}_2^T(t) \tag{59}$$

$$\underline{c}_{y_1\underline{v}_2} \approx \sum_{t=1}^{N} \beta_1^{N-t}\underline{v}_2^*(t)y_1(t) \tag{60}$$

$$C_{\underline{y}_1\underline{v}_1} \approx \sum_{t=1}^{N} \beta_2^{N-t}\underline{v}_1^*(t)\underline{y}_1^T(t) \tag{61}$$

$$\underline{c}_{y_2\underline{v}_1} \approx \sum_{t=1}^{N} \beta_2^{N-t}\underline{v}_1^*(t)y_2(t) \tag{62}$$

where β_1 and β_2 are real numbers between 0 and 1. In (59)–(62) we have omitted the multiplying factors since they cancel each other. To achieve maximum statistical stability, we choose $\beta_1 = \beta_2 = 1$. However, if the signals and/or the coupling systems exhibit nonstationary behavior over time, it may be preferable to choose $\beta_1 . \beta_2 < 1$. In that way we introduce exponential weighting that gives more weight to current data samples, and we have in effect an adaptive algorithm that is capable of tracking the varying characteristics of the underlying system.

Substituting (59) and (60) into (57) and following the development in the Appendix, we obtain the following recursive algorithm for adjusting $\underline{a}$ (for a given $\underline{b}$):

$$\underline{a}(t) = \underline{a}(t-1) + Q(t)\underline{v}_2^*(t)v_1(t;\underline{a}(t-1)) \tag{63}$$

where

$$Q(t) = \frac{1}{\beta_1}\left[Q(t-1) - \frac{Q(t-1)\underline{v}_2^*(t)\underline{y}_2^T(t)Q(t-1)}{\beta_1 + \underline{y}_2^T(t)Q(t-1)\underline{v}_2^*(t)}\right] \tag{64}$$

where $\underline{a}(t)$ is the solution to (57) based on data (i.e., on sample averages) to time t, and $v_1(t:\underline{a}(t-1))$ is the signal in (35) computed at $\underline{a} = \underline{a}(t-1)$.

In a similar way we obtain

$$\underline{b}(t) = \underline{b}(t-1) + R(t)\underline{v}_1^*(t)v_2(t:\underline{b}(t-1)) \tag{65}$$

where

$$R(t) = \frac{1}{\beta_2}\left[R(t-1) - \frac{R(t-1)\underline{v}_1^*(t)\underline{y}_1^T(t)R(t-1)}{\beta_2 + \underline{y}_1^T(t)R(t-1)\underline{v}_1^*(t)}\right] \tag{66}$$

where $\underline{b}(t)$ is the solution to (58) based on data to time t, and $v_2(t;\underline{b}(t-1))$ is the signal in (36) computed at $\underline{b} = \underline{b}(t-1)$. We therefore have an iterative-recursive algorithm in which for a given $\underline{b}$ the vector $\underline{a}$ is adjusted recursively in t using (63) and (64), and for a given $\underline{a}$ the vector $\underline{b}$ is adjusted recursively using (65) and (66).

To convert this iterative-recursive algorithm into a sequential algorithm, we propose replacing $\underline{a}$ and $\underline{b}$ by their current estimates. This corresponds to replacing $v_2(t)$ in (63) and (64) by $v_2(t:\underline{b}(t-1))$, and $v_1(t)$ in (65) and (66) by $v_1(t;\underline{a}(t-1))$.

An alternative approach for deriving a sequential algorithm is obtained by rewriting (45) and (46) in the form

$$\underline{c}_{y_1\underline{v}_2} - C_{\underline{y}_2\underline{v}_2}\underline{a} = E\{\underline{v}_2^*(t)v_1(t)\} \tag{67}$$

$$\underline{c}_{y_2\underline{v}_1} - C_{\underline{y}_1\underline{v}_1}\underline{b} = E\{\underline{v}_1^*(t)v_2(t)\}. \tag{68}$$

Applying Robbins-Monro first-order stochastic approximation methods [5], [9], (67) and (68) give rise to the following sequential algorithm:

$$\underline{a}(t) = \underline{a}(t-1) + \gamma_1(t)\underline{v}_2^*(t:\underline{b}(t-1))v_1(t:\underline{a}(t-1)) \tag{69}$$

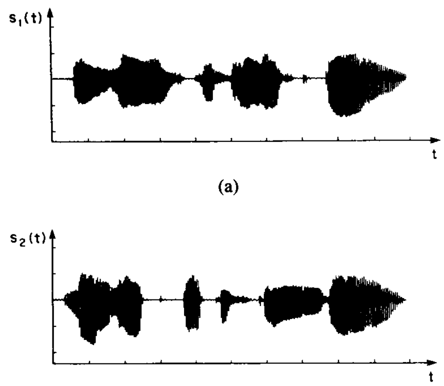


Fig. 6. The speech signals: (a) "He has the bluest eyes." (b) "Line up at the screen door."

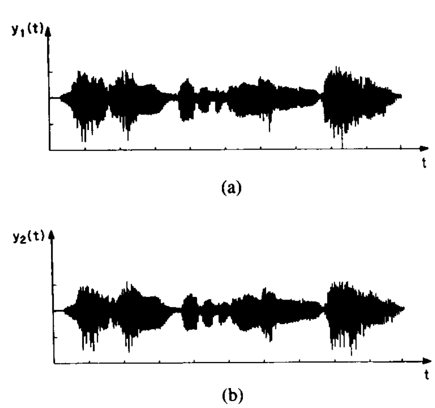


Fig. 7. The measured signals.

$$\underline{b}(t) = \underline{b}(t-1) + \gamma_2(t)\underline{v}_1^*(t;\underline{a}(t-1))v_2(t;\underline{b}(t-1)) \quad (70)$$

where $\gamma_1(t)$ and $\gamma_2(t)$ are some preselected gains that may depend on the time index t. To ensure convergence under stationary conditions, it is recommended to choose $\gamma_i(t)$ $i = 1, 2$ to be positive sequences such that (see, e.g., [6])

$$\lim_{t\to\infty} \gamma_i(t) = 0, \quad \sum_{t=1}^{\infty} \gamma_i(t) = \infty, \quad \sum_{t=1}^{\infty} \gamma_i^2(t) < \infty$$

e.g., $\gamma_i(t) = \gamma_i/t$. However, if the signals and/or the coupling systems exhibit changes in time, and we want an adaptive algorithm, choosing constant gains $\gamma_i(t) = \gamma_i$ $i = 1, 2$ is recommended. This corresponds to an exponential weighting that reduces the effect of past data samples relative to new data in order to track the varying characteristics.

If we substitute $\underline{b}(t-1) = 0$ in (69), we obtain

$$\underline{a}(t) = \underline{a}(t-1) + \gamma_1(t)\underline{y}_2^*(t)v_1(t;\underline{a}(t-1)) \quad (71)$$

which is recognized as the LMS algorithm suggested by Widrow *et al.* [13] for solving the indicated least-squares problem under the assumption that there is no coupling of $s_1(t)$ into $y_2(t)$. Similarly, substituting $\underline{b} = 0$ in (63) and (64), so that $v_2(t) = y_2(t)$, we obtain the RLS algorithm (e.g., see [8]) for solving the indicated least-squares problem. The algorithms developed in this section can therefore be viewed as extensions of the LMS and RLS algorithms for the case where each of the observed signals has components from both sources, and either one or both coupling systems are unknown.

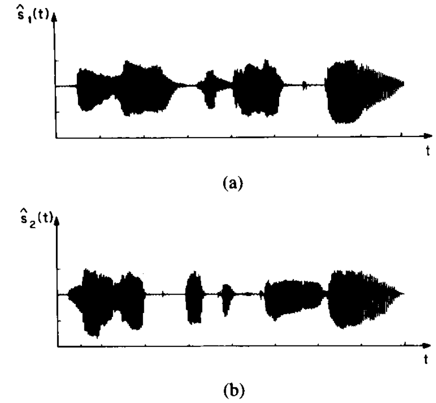


Fig. 8. The recovered signals using the decorrelation method.

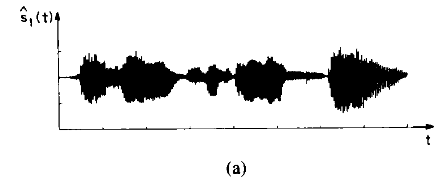


Fig. 9. The recovered signal using the least squares method.

As an illustration of performance, we have considered the following example: $s_1(t)$ and $s_2(t)$ are the sampled speech signals shown in Fig. 6, corresponding to the sentences "He has the bluest eyes" and "Line up at the screen door," respectively. The coupling systems H_{12} and H_{21} are unknown FIR filters of order 10. We only assume prior knowledge of the filter orders. The ratio of average power of the signal to average power of the interference at the first and the second microphones were -1.8 dB and -2 dB, respectively, indicating strong coupling effects. In Fig. 7 we have shown the measured signals $y_1(t)$ and $y_2(t)$. To identify the coupling filters we have implemented the iterative batch algorithm in (57) and (58), where the covariances are replaced by their sample estimates given in (59)–(62), with $\beta_1 = \beta_2 = 1$. The recovered signals are shown in Fig. 8. The post-processing signal-to-interference power ratio at the first and second microphones were 7.5 dB and 8.3 dB, respectively. For purpose of comparison, we have plotted in Fig. 9 the recovered signal at the first (primary) microphone using the LS method, which corresponds to solving (57) under the incorrect choice $\underline{b} = 0$. The post-processing signal-to-interference power ratio in this case was 1.8 dB. Unlike the decorrelation approach that treats the signals as being equally important, the LS method regards $s_2(t)$ as being unwanted interfering signal, and therefore it makes no attempt to estimate it. By actually listening to the recovered speech signals, in the LS method one could hear the reverberant distortion due to the fact that the desired signal is canceled with some delay together with the interfering signal. This reverberant effect does not exist when using the method developed in this paper.

Appendix: Derivation of the Recursive Algorithm in (63) (64)

Define

$$Q(t) = \left[\sum_{k=1}^{t} \beta_1^{t-k} \underline{v}_2^*(k) \underline{y}_2^T(k)\right]^{-1}$$

$$- [\beta_1 Q^{-1}(t-1) + \underline{v}2^*(t) \underline{y}_2^T(t)]^{-1}$$

$$= \frac{1}{\beta_1}\left[Q(t-1) - \frac{Q(t-1)\underline{v}_2^*(t)\underline{y}_2^T(t)Q(t-1)}{\beta_1 + \underline{y}_2^T(t)Q(t-1)\underline{v}_2^*(t)}\right]$$

$$\underline{q}(t) = \sum_{k=1}^{t} \beta_1^{t-k} \underline{v}_2^*(k) y_1(k) = \beta_1 \underline{q}(t-1) + \underline{v}_2^*(t) y_1(t)$$

Then, by (57),

$$\underline{a}(t) = Q(t)\underline{q}(t) = Q(t)[\beta_1 \underline{q}(t-1) + \underline{v}_2^*(t) y_1(t)]$$

$$= Q(t)[\beta_1 Q^{-1}(t-1)\underline{a}(t-1) + \underline{v}_2^*(t) y_1(t)]$$

$$= Q(t)\left\{\beta_1 \cdot \frac{1}{\beta_1}[Q^{-1}(t) - \underline{v}_2^*(t)\underline{y}_2^T(t)]\underline{a}(t-1) + \underline{v}_2^*(t) y_1(t)\right\}$$

$$= \underline{a}(t-1) + Q(t)\underline{v}_2^*(t)[y_1(t) - \underline{y}_2^T(t)\underline{a}(t-1)]$$

$$= \underline{a}(t-1) + Q(t)\underline{v}_2^*(t) v_1(t; \underline{a}(t-1)).$$

Acknowledgment

We would like to thank Dr. Arie Feuer for a number of important and helpful discussions. We would also like to thank the anonymous reviewers for their careful reading and helpful comments.

References

[1] Y. Bar-Ness, J. Carlin, and M. Steinberger, "Bootstrapping adaptive cross-pol canceller for satellite communications," *IEEE Int. Conf. Communications* (Philadelphia, PA), pp. 4F.5.1–4F.5.5, June 1982.

[2] D. Brandwood, "Cross-coupled cancellation system for improving cross-polarisation discrimination," *IEEE Int. Conf. Antennas and Propagation, Part I* (London, England), pp. 41–45, Nov. 1978.

[3] M. Feder, A. Oppenheim, and E. Weinstein, "Maximum likelihood noise cancellation using the EM algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing,* vol. 37, pp. 204–216, Feb. 1989.

[4] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing,* vol. 24, no. 1, pp. 1–10, July 1991.

[5] J. Keifer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.,* vol. 23, pp. 462–466, 1952.

[6] H. Kushner, *Stochastic Approximation Methods of Constrained and Unconstrained Systems.* New York: Springer Verlag, 1978.

[7] M. Larimore, "The adaptive minimum error correlation filter," PhD dissertation, Stanford University, Stanford, CA.

[8] L. Ljung, *System Identification.* New York: Wiley 1988.

[9] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.,* vol. 22, pp. 400–407, 1951.

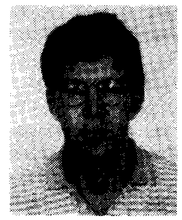
[10] E. Weinstein, M. Feder, and A. V. Oppenheim, "Multi-channel signal separation," patent pending, patent application no. 07/750,917, filing date August 28, 1991.

[11] E. Weinstein, M. Feder, and A. Oppenheim, "Signal separation by decorrelation in a constant inter-channel interference," tech. rep. no. EE-S-92-04, Dept. Electrical Engineering–Systems, Tel-Aviv University, Mar. 1992.

[12] E. Weinstein, A. V. Oppenheim, M. Feder, and J. R. Buck, "Iterative and sequential algorithms for multi-sensor signal enhancvement," accepted for publication to *IEEE Trans. Signal Processing.*

[13] B. Widrow *et al.*, "Adaptive noise canceling: principles and application," in *Proc. IEEE,* vol. 63, pp. 1692–1716, 1975.

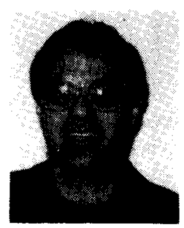
[14] D. Yellin and E. Weinstein, "Criteria for multi-channel signal separation," submitted for publication to *IEEE Trans. Signal Processing.*



Ehud Weinstein received the B.Sc. degree from the Technion-Israel Institute of Technology, and the Ph.D. degree from Yale University, both in electrical engineering, in 1975 and 1978, respectively.

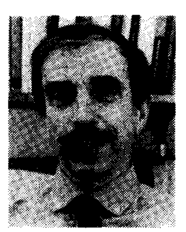
In 1980 he joined the Department of Electrical Engineering–Systems, Faculty of Engineering, Tel-Aviv University, where he is currently a Professor. Since 1986 he has been an Adjunct Scientist at the Woods Hole Oceanographic Institution, and since 1990 he has been a Research Affiliate at the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology. His research interests are in the general areas of estimation theory and statistical signal processing.

Dr. Weinstein is a co-recipient of the 1983 Senior Award of the IEEE Acoustics, Speech, and Signal Processing Society.



Meir Feder received the B.Sc and M.Sc degrees (summa cum laude) from Tel-Aviv University, Tel-Aviv, Israel, and the Sc.D degree from the Massachusetts Institute of Technology, Cambridge, and the Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA, all in electrical engineering, in 1980, 1984, and 1987, respectively.

During 1987–1988 he was a Research Associate in the Department of Electrical Engineering and Computer Science at MIT. He then worked in Elbit Computers, Haifa, Israel. In October 1989 he joined the faculty of the Department of Electrical Engineering–Systems, Tel-Aviv University. He was also a Visiting and Guest Investigator at the WHOI in the summers of 1983 and 1988–1991 as well as a visiting scientist at the Scripps Institute of Oceanography in the summer of 1992. His research interests include signal and image processing, sonar signal processing, data and signal compression, and topics in information theory. Since June 1993 he serves as Associate Editor for Source Coding of the IEEE Transactions on Information Theory.



Alan V. Oppenheim received the S.B. and S.M. Degrees in 1961 and the Sc.D. degree in 1964, all in electrical engineering, from the Massachusetts Institute of Technology.

In 1964, he joined the faculty at MIT, where he currently holds the Distinguished Professor chair in Electrical Engineering. Since 1967 he has also been affiliated with MIT Lincoln Laboratory and since 1977 with the Woods Hole Oceanographic Institution. His research interests are in the general area of signal processing and its applications to speech, image, and under water acoustic data processing. He is coauthor of the widely used textbooks *Discrete-Time Signal Processing* and *Signals and Systems.* He is also editor of several advanced books on signal processing.

Dr. Oppenheim is a member of the National Academy of Engineering, a Fellow of the IEEE, and a member of Sigma Xi and Eta Kappa Nu. He has been a Guggenheim Fellow and a Sackler Fellow. He has also received a number of awards for outstanding research and teaching, including the IEEE Education Medal, the IEEE Centennial Award, and the Society Award, the Technical Achievement Award, and the Senior Award of the IEEE Society on Acoustics, Speech, and Signal Processing.