Multiple Signal Extraction by Multiple Interference Attenuation in the Presence of Random Noise in *Seismic* Array Data[1]

By

Magdy T. Hanna[2]
Department of Engineering Mathematics and Physics
Cairo University / Fayoum Branch
Fayoum, Egypt

ABSTRACT

A vector of digital filters is derived for the multichannel processing of the signals acquired by an array of sensors with the objective of extracting *multiple desired signals* by the attenuation of multiple interferences and random noise. The signals and interferences are assumed to have *arbitrary waveforms* with no a priori knowledge of these waveforms. The time duration of the recorded array data is assumed to be long enough to incorporate all time delayed propagated waveforms at the sensors of the array. The derivation is for the general case of an *arbitrary array geometric configuration* and is not confined to the special case of a linear array of equispaced sensors. The rationale adopted in the derivation of the filters is to give first priority *at each discrete frequency* to passing the signals, a second priority to canceling the interferences and a third priority to attenuating the random noise. This rationale well suits the case of *seismic data*, that are dominantly corrupted by strong interferences rather than random noise. Solving a constrained minimization problem derives the vector of array filters. The computation of this vector requires the application of the powerful QR matrix decomposition technique for the detection of any *redundant and/or inconsistent constraints* at each discrete frequency. The simulation results demonstrate the extraction ability of the derived filters in both the multiple input single output and the multiple input multiple output processing schemes.

---

[1] EDICS category: 1-MSAM
[2]Mailing address: Magdy Tawfik Hanna, P.O. Box 568, Alexandria, Egypt.
Fax : 0020-2-5706033 (Giza), Tel : 0020-2-5706033 (Giza), Email: hanna@ieee.org

I. INTRODUCTION

In seismic prospecting for oil, data are acquired by an array of sensors, e.g. geophones or hydrophones. *Seismic data* consist of weak desired signals corrupted by strong coherent interferences and random noise. Desired signals can result from reflections by subsurface oil carrying layers and coherent interferences can arise from different modes of ground roll causing direct propagation in the form of Rayleigh waves from the source of the seismic explosion to the array of sensors [1-3]. Many techniques have been applied for the off-line processing of the recordings of the seismic array for the purpose of signal extraction. These techniques vary drastically in their degree of sophistication and in their extraction ability. The delay-and-sum, the weighted delay-and-sum methods [4] and the Chebyshev technique [5] are simple multichannel methods for combining the array data that have been time-shifted to be aligned on a single desired signal. The optimum array filters of [6] and the absolutely optimum array filters of [7] are multichannel frequency-domain techniques derived for extracting a *single desired signal* by attenuating a single interference without explicitly taking random noise into consideration. Optimum array filters were also derived for extracting a *single desired signal* by suppressing a single interference [8] or multiple interferences [9] while explicitly taking random noise attenuation into account as a secondary objective.

The theme of this paper is the solution of the more general and more challenging problem of frequency-domain multichannel processing of the array data with the objective of extracting *multiple desired signals* by suppressing multiple interferences in the presence of random noise. More specifically, a vector of digital filters will be derived for the multichannel processing scheme of Fig. 1 by giving a first priority *at each discrete frequency* for achieving all-pass conditions for the sum of the signals (if at all possible), a second priority for suppressing the interferences (if not in conflict with the first objective) and a third priority for attenuating the random noise. This rationale, apart from its mathematical elegance, typically suits the case of *seismic array data* that are dominantly corrupted by interferences rather than random noise. It should be mentioned that although the present paper deals with a more general case than those in [6-9] do, it is based on *the same assumptions* (to be given next section). It should be emphasized that this work is aimed at extracting the arbitrary waveforms of the signals rather than finding their directions of arrival. Knowledge or estimation of the delay times of the signals and interferences is required only to be able to apply the array filters with the goal of attenuating the interferences and suppressing the random noise in order to help extracting the waveforms of the signals.

Since there are more than one desired signal and since the beamforming configuration of Fig. 1 has only one output trace, the most that can be done is to extract the sum of the desired signals as they appear on any given trace (a trace is defined as the data recorded by the sensor) to be referred to as the reference trace which can be for example the first trace. Any sensor can be taken as the reference sensor by measuring all delay times of the signals and interferences with respect to that sensor. Adopting this point of view, if the desired signals are nonoverlapping on the reference trace – although being overlapping on the remaining traces – and if the array filters are able to achieve enough attenuation of the interferences and random noise, then the desired signals will be identifiable on the single output trace (although being difficult to identify on the input traces including the reference trace). It should be emphasized that the mathematical derivation of this paper is general and is *not based on an assumption of having nonoverlapping desired signals on a certain trace*. It should also be pointed out that since the signals travel in different directions, they *cannot be replaced by one signal equal to their sum*, and consequently the array filters of [9] cannot be applied for the *multiple signals case*; they can only be applied for extracting one signal at a time. One situation where one is interested in getting the sum of several signals occurs in the case of a horizontal array of geophones receiving reflections from several subsurface layers where the delay times for each layer follow a hyperbolic law [3]. Due to the different depths of the layers the desired signals can be nonoverlapping in the recordings of some of the geophones and their sum on a reference geophone will be of prime importance. Another example can occur in the case of vertical seismic profiling where the geophones of the array are deployed vertically underground rather than being placed horizontally under the surface of the earth [2].

In section II, the mathematical model for the array data is presented and in Section III the array filters are derived. In section IV, the implementation of the array filters is considered where the QR matrix decomposition technique is applied for the *detection of any redundant and/or inconsistent constraints* which is of prime importance for the method of this paper. The QR matrix technique is also applied for the computation of the vector of digital filters. In section V, some simulation results are presented for both the multiple input single output and the multiple input multiple output processing schemes.

II. THE MATHEMATICAL FORMULATION OF THE PROBLEM

After defining the notation to be used in this paper and stating the main assumptions, the model of the time-domain array data will be introduced and then transformed to the frequency domain. Some vectors and matrices will be defined in order to put the problem in a compact form. The algorithm for processing the array data will be provided.

Nomenclature

N Number of sensors of the array

K Length of the recording of each sensor

q_1 Number of desired signals

q_2 Number of coherent interferences

z_n The n th trace of the array (the signal recorded by the n th sensor), $n = 1, \cdots, N$

x_m The m th desired signal, $m = 1, \cdots, q_1$

u_m The m th coherent interference, $m = 1, \cdots, q_2$

w_n The random noise in the recording of the n th sensor

ξ_{nm} The delay time of the m th desired signal at the n th sensor

β_{nm} The amplitude scale factor of the m th desired signal at the n th sensor

τ_{nm} The delay time of the m th coherent interference at the n th sensor

α_{nm} The amplitude scale factor of the m th coherent interference at the n th sensor

σ_n^2 The variance of w_n, $n = 1, \cdots, N$

Assumptions

- The q_1 *waveforms* $x_m(r)$ and q_2 *waveforms* $u_m(r)$ where r is the discrete time, are *arbitrary* and generally different with no a priori knowledge of those waveforms except their lengths in order to select K appropriately; (i.e. the treatment is not confined to the case of narrowband signals).
- The *array geometric configuration is arbitrary (*rather than being restricted to the case of a linear array with equispaced sensors).

- The sampling rate is high enough such that the delay times ξ_{nm} and τ_{nm} are integers.

- The length K of each sensor recording is large enough to incorporate all samples of all signals and interferences after time shifting. If only random noise appears near the end of each trace, one can tell that K is sufficiently large despite the lack of any knowledge about the waveforms of the signals and interferences. (See Figs. 2 and 9). Consequently K is determined by the time duration of all signals and interferences as well as their maximum propagation delays across the array. The length of the FFT to be applied for transforming the input traces to the frequency domain is also K.

- q_1 and q_2 are known and $q_1 + q_2 < N$. In most practical cases q_1 and q_2 can be determined by an expert examination of the recorded seismic traces given knowledge of the physical setting of the seismic experiment. For example in the case of a horizontal array, the weak up-traveling signals are the desired ones and the strong down-traveling signals are the interferences. (See the simulation results).

- The signals x_m and interferences u_m are deterministic and have finite time support.

- The random noise terms $w_n(r)$ are zero-mean temporally white and spatially uncorrelated, i.e. uncorrelated for different n. Those random noise terms model the ambient noise in the medium where the sensors are deployed and the imprecision of the recording process.

- The delay times ξ_{nm} and τ_{nm} are either known or can be estimated from the recorded data using any of the delay time estimation techniques [10-11].

- For each signal $x_m(r)$, $m = 1,\cdots,q_1$, the amplitude scale factors β_{nm}, $n = 1,\cdots,N$ can be estimated from a knowledge of the attenuation properties of the medium and the geometric configuration of the array. In most cases especially for signals arriving from far sources, the β_{nm} are assumed to be equal. The same applies to the amplitude scale factors α_{nm}, $n = 1,\cdots,N$ of the m th interference u_m, $m = 1,\cdots,q_2$.

- The relative values of the variances σ_n^2 of the random noise terms $w_n, n = 1,\cdots,N$ are assumed to be known from knowledge of the nature of the random disturbances in the field where the sensors are deployed. In most cases the σ_n^2's are assumed to be equal. It will be shown in section III that the expression of the derived array filters depends only on the relative values of the σ_n's rather than on their absolute values.

The Mathematical Model

The data recorded by an array of N sensors will be modeled as:

$$z_n(r) = \sum_{m=1}^{q_1} \beta_{nm} x_m\left(r - \xi_{nm}\right) + \sum_{m=1}^{q_2} \alpha_{nm} u_m\left(r - \tau_{nm}\right) + w_n(r) \qquad , n = 1, \cdots, N$$

$$, r = 0, \cdots, K-1 \qquad (1)$$

where $x_m(r)$ and $u_m(r)$ are the m th coherent signal and interference respectively. Here the word 'coherent' is always used in the sense that each source signal leads to coherent sensor signals, and not that the source signals are coherent among themselves. Putting it another way, the m th interference has the *same* arbitrary waveform $u_m(r)$ for all the N sensors of the array apart from the scale factors α_{nm} and delay times τ_{nm}. The same applies to the desired coherent signal $x_m(r)$.

Taking the K-point discrete Fourier transform (DFT) [12] of (1), one gets[3]:

$$Z_n(k) = \sum_{m=1}^{q_1} \beta_{nm} e^{-j\omega\xi_{nm}} X_m(k) + \sum_{m=1}^{q_2} \alpha_{nm} e^{-j\omega\tau_{nm}} U_m(k) + W_n(k) \qquad , n = 1, \cdots, N$$

$$, k = 0, \cdots, K-1 \qquad (2)$$

where ω is the Fourier frequency[4]:

$$\omega = \frac{2\pi k}{K} \qquad , k = 0, \cdots, K-1. \qquad (3)$$

In order to express (2) concisely one defines the following N-dimensional vectors :

$$\mathbf{Z(k)} = \begin{pmatrix} Z_1(k) \\ \vdots \\ Z_N(k) \end{pmatrix} \quad , \quad \mathbf{W(k)} = \begin{pmatrix} W_1(k) \\ \vdots \\ W_N(k) \end{pmatrix} \quad ,$$

---

[3]Here the linear shift is the same as circular shift since K is assumed to be large enough to incorporate all samples of all signals and interferences after time-shifting.

[4]Since ω depends on k, it should have been written as ω_k; however, the subscript k has been dropped to simplify the notation.

$$\mathbf{b_m(k)} = \begin{pmatrix} \beta_{1m} e^{j\omega\xi_{1m}} \\ \vdots \\ \beta_{Nm} e^{j\omega\xi_{Nm}} \end{pmatrix} \quad , \quad \mathbf{a_m(k)} = \begin{pmatrix} \alpha_{1m} e^{j\omega\tau_{1m}} \\ \vdots \\ \alpha_{Nm} e^{j\omega\tau_{Nm}} \end{pmatrix} \tag{4}$$

and the following q_1-dimensional vector of transformed desired signals and the q_2-dimensional vector of transformed interferences:

$$\mathbf{X(k)} = \begin{pmatrix} X_1(k) \\ \vdots \\ X_{q_1}(k) \end{pmatrix} \quad , \quad \mathbf{U(k)} = \begin{pmatrix} U_1(k) \\ \vdots \\ U_{q_2}(k) \end{pmatrix} . \tag{5}$$

One also defines the following $N \times q_1$ and $N \times q_2$ matrices:

$$\mathbf{B(k)} = \begin{pmatrix} \mathbf{b_1(k)} & \cdots & \mathbf{b_{q_1}(k)} \end{pmatrix} \tag{6}$$

and

$$\mathbf{A(k)} = \begin{pmatrix} \mathbf{a_1(k)} & \cdots & \mathbf{a_{q_2}(k)} \end{pmatrix} . \tag{7}$$

Using (4)-(7), one gets - from (2) - the following vector form for the transformed array data[5] :

$$\mathbf{Z(k)} = \mathbf{B^*(k)X(k)} + \mathbf{A^*(k)U(k)} + \mathbf{W(k)} . \tag{8}$$

Typically the beamforming configuration of Fig. 1 is used for processing the array data. The discrete Fourier Transform (DFT) of the output of this multichannel processing scheme is :

$$Y(k) = \mathbf{Z'(k)F(k)} \tag{9a}$$

where $\mathbf{F(k)}$ is the column vector whose components are the Discrete Fourier Transform (DFT) representation of the array filters to be derived next section, i.e.

$$\mathbf{F(k)} = \begin{pmatrix} F_1(k) & \cdots & F_N(k) \end{pmatrix}' \tag{9b}$$

Combining (8) and (9a), one obtains:

$$Y(k) = \mathbf{X'(k)}\left(\mathbf{B^+(k)F(k)} \right) + \mathbf{U'(k)}\left(\mathbf{A^+(k)F(k)} \right) + \mathbf{F'(k)W(k)} . \tag{10}$$

---

[5]The superscripts * , ' , + denote the complex conjugate, the transpose and the Hermitian transpose, respectively. The three symbols are used since they are needed in different situations.

In the model (1): $x_m(r)$ and $u_m(r)$ are deterministic, and the random noise term $w_n(r)$ is a K-point segment of a sample of a discrete-time random process. Since the noise terms $w_n(r)$ are zero-mean, subtracting from (10) its expected value remembering that $\mathbf{W(k)}$ is the only random term on the right hand side of (10), one gets:

$$Y(k) - E[Y(k)] = \mathbf{F'(k)W(k)} \qquad (11)$$

where E[ ] is the expectation operator. ($E[\ ]$ is an ensemble average operator and not an average operator over k). Therefore the variance of Y(k) is given by:

$$E\left[|Y(k) - E[Y(k)]|^2\right] = \mathbf{F'(k)}E\left[\mathbf{W(k)W^+(k)}\right]\mathbf{F^*(k)}\ . \qquad (12)$$

Let $\mathbf{w}(\mathbf{r})$ be the N-dimensional discrete-time vector random process whose elements are $w_n(r),\ n = 1,\cdots,N$. Consider K samples $\{\mathbf{w}(\mathbf{0}),\cdots,\mathbf{w}(\mathbf{K-1})\}$ of $\mathbf{w}(\mathbf{r})$. Let $\mathbf{S_w}(\mathbf{k}), k = 0,\cdots,K-1$ be samples of the power spectral density matrix of $\mathbf{w}(\mathbf{r})$ [13-15]. Similarly let $S_v(k)$ be samples of the power spectral density of the scalar random process $v(r)$ defined by its DFT $V(k)$ as:

$$V(k) = Y(k) - E[Y(k)]\ . \qquad (13)$$

Both $\mathbf{S_w}(\mathbf{k})$ and $S_v(k)$ are evaluated at the discrete frequency k corresponding to the Fourier frequency ω of (3). It can be proved that for a vector random process:

$$\lim_{K\to\infty} \mathbf{S_w(k)} = \lim_{K\to\infty} E\left[\frac{1}{K}\mathbf{W(k)W^+(k)}\right]. \qquad (14)$$

This result is a generalization of its counterpart for the scalar case [15]:

$$\lim_{K\to\infty} S_v(k) = \lim_{K\to\infty} E\left[\frac{1}{K}|V(k)|^2\right]. \qquad (15)$$

The approximation employed here is to substitute (14) and (15) in (12) although the length K is finite, to get:

$$S_v(k) = \mathbf{F'(k)S_w(k)F^*(k)}. \qquad (16)$$

Since it is assumed that the random processes from which the samples $w_n(r)$ are drawn are spatially uncorrelated, matrix $\mathbf{S_w}(\mathbf{k})$ will be diagonal; and since those random processes are temporally white, this diagonal matrix will be constant [13,14]. Therefore $\mathbf{S_w}(\mathbf{k})$ reduces to:

$$\mathbf{S}_w(\mathbf{k}) = \mathbf{G} = Diag\left\{\sigma_1^2, ..., \sigma_N^2\right\} \quad (17)$$

where σ_n^2 is the variance of the discrete-time white random noise process corresponding to the term $w_n(r)$. From (16) and (17) one gets the following expression for the output power spectral density in response to the random noise:

$$S_v(k) = \mathbf{F}^+(\mathbf{k})\mathbf{G}\mathbf{F}(\mathbf{k}). \quad (18)$$

The Signal Extraction Algorithm:

1. Select the reference trace.
2. Identify the number of signals and interferences and their delay times (with respect to the reference trace) and amplitude scale factors.
3. Compute the DFT of each of the input traces.
4. For each discrete frequency k = 0 , ... , K-1 compute the vector $\mathbf{F(k)}$ of array filters (to be derived below) and the output Y(k) using (9a).
5. Take the inverse DFT of Y(k) to get the time domain single output trace y(n) of the multichannel array processing scheme.

III. DERIVATION OF THE VECTOR OF ARRAY FILTERS

The objective here is to derive the vector F(k) of array filters to be applied to the array data to get the output of the multichannel processing scheme using (9a). The philosophy to be followed in approaching the solution will be first stated, the set of independent and consistent constraints will be identified, and a constrained optimization problem will be solved to get the optimal vector $\mathbf{F(k)}$. Finally a comparison with other beamforming methods will be provided.

Philosophy

The 3 terms, which appear on the right hand side of (10), can be interpreted as the outputs in response to the desired signals, the interferences and the random noise respectively. Stated another way, the terms $\mathbf{B}^+(\mathbf{k})\mathbf{F}(\mathbf{k})$ and $\mathbf{A}^+(\mathbf{k})\mathbf{F}(\mathbf{k})$ represent the processing of the multiple signals and multiple interferences respectively, while the last term is due to the random sensor noise. Since the multichannel scheme of Fig. 1 has only one output trace, the most that can be done in extracting the signals is to try to extract their sum as they appear on a reference trace as was delineated in Section I. This can be achieved by making Y(k) as close as possible to $\boldsymbol{\mu}'\mathbf{X(k)}$ at each discrete frequency k, where $\boldsymbol{\mu}$ is the summing vector defined by:

$$\boldsymbol{\mu} = \begin{pmatrix}1 & \cdots & 1\end{pmatrix}' . \tag{19}$$

This can be accomplished by trying to achieve (if at all possible) the following goals :

1. An all-pass processing of the desired signals by imposing the set of q_1 linear constraints:

$$\mathbf{B}^{+}\mathbf{(k)F(k)} = \boldsymbol{\mu} \ . \tag{20}$$

2. Suppression of the interferences by imposing the set of q_2 linear constraints:

$$\mathbf{A}^{+}\mathbf{(k)F(k)} = \mathbf{0} \ . \tag{21}$$

3. Attenuation of the random noise by the minimization of the output power spectral density (18) in response to the random noise.

The philosophy to be followed here is to rank these goals in the above order. Since extracting the desired signals that have arbitrary waveforms is the main objective, the first goal is of prime importance; and since typically seismic array data are corrupted by interferences rather than random noise, the second goal ranks rationally before the third one.

Independent and Consistent Constraints

It will be assumed that the number of array sensors is larger than the number of signals and interferences, i.e.,

$$q_1 + q_2 < N . \tag{22}$$

However the joint set of $q_1 + q_2$ constraints of (20) and (21) may not be independent or even consistent. Dealing with this situation is the main concern of the method presented here. Let r_1 be the rank of matrix $\mathbf{B(k)}$. If $r_1 = q_1$ then $\mathbf{B}^{+}\mathbf{(k)}$ will have a full row rank and consequently the set of q_1 equations of (20) will be both linearly independent and consistent [16]. If $r_1 < q_1$ then the q_1 equations of (20) can be generally classified into the following 3 sets (where any of the last two can be empty) :

a) a set of $q_{11} = r_1$ independent and consistent equations[6]:

$$\mathbf{B}_{\mathbf{1}}^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) = \boldsymbol{\mu} \tag{23}$$

b) a set of q_{12} consistent equations which are linearly dependent on the first set:

$$\mathbf{B}_{\mathbf{2}}^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) = \boldsymbol{\mu} \tag{24}$$

c) a set of q_{13} inconsistent equations (i.e. cannot be satisfied simultaneously with (23)):

$$\mathbf{B}_{\mathbf{3}}^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) = \boldsymbol{\mu} \tag{25}$$

Matrix $\mathbf{B}_{\mathbf{1}}^{+}(\mathbf{k})$ consists of a set of maximum number of linearly independent rows of $\mathbf{B}^{+}(\mathbf{k})$. The rows of $\mathbf{B}_{\mathbf{2}}^{+}(\mathbf{k})$ and $\mathbf{B}_{\mathbf{3}}^{+}(\mathbf{k})$ are linearly dependent on those of $\mathbf{B}_{\mathbf{1}}^{+}(\mathbf{k})$. Therefore the two sets of constraints (23) and (24) are feasible while the set (25) is infeasible. If the equations of (23) are satisfied, then those of (24) will also be satisfied. Therefore the set of consistent and dependent constraints (24) will be dropped from consideration; however, the signals corresponding to the rows of matrix $\mathbf{B}_{\mathbf{2}}^{+}(\mathbf{k})$ will enjoy the same all-pass condition as those corresponding to the rows of matrix $\mathbf{B}_{\mathbf{1}}^{+}(\mathbf{k})$. Because of the linear dependence, matrix $\mathbf{B}_{\mathbf{3}}(\mathbf{k})$ can be expressed as:

$$\mathbf{B}_3(\mathbf{k}) = \mathbf{B}_1(\mathbf{k})\mathbf{T} \tag{26}$$

where $\mathbf{T}$ is a $q_{11} \times q_{13}$ matrix. Consequently the difference between both sides of (25) which cannot be satisfied as an equality because of the infeasibility, reduces to:

$$\begin{aligned} \mathbf{B}_3^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) - \boldsymbol{\mu} &= \mathbf{T}^{+}\mathbf{B}_1^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) - \boldsymbol{\mu} \\ &= \mathbf{T}^{+}\boldsymbol{\mu} - \boldsymbol{\mu} \end{aligned} \tag{27}$$

where the last equality was obtained by utilizing (23). Therefore the norm $\left\|\mathbf{B}_3^{+}(\mathbf{k})\mathbf{F}(\mathbf{k}) - \boldsymbol{\mu}\right\|$ is a constant independent of vector $\mathbf{F}(\mathbf{k})$ as long as the constraint (23) is satisfied. This implies that nothing can be done to

[6] Although the vectors $\boldsymbol{\mu}$ in (20) and (23)-(25) are all defined by (19), they are not the same vector since they have different dimensions.

help passing the signals corresponding to the rows of matrix $\mathbf{B}_3^+(\mathbf{k})$. The identification of the columns of matrix $\mathbf{B}_1(\mathbf{k})$ from those of matrix $\mathbf{B}(\mathbf{k})$ and the challenging problem of the implementation of the array filters will be treated in detail next section.

Next the feasibility of the constraints of (21) - in charge of suppressing the interferences - will be investigated subject to the requirement that vector $\mathbf{F}(\mathbf{k})$ has been constrained to satisfy (23). The q_2 rows of matrix $\mathbf{A}^+(\mathbf{k})$ of (21) will be most generally classified into the following 3 sets (any of them can be empty):

a) a set of q_{21} rows $\mathbf{A}_1^+(\mathbf{k})$ which are linearly independent and at the same time independent of the rows of matrix $\mathbf{B}_1^+(\mathbf{k})$ of (23).

b) a set of q_{22} rows $\mathbf{A}_2^+(\mathbf{k})$ which are linearly dependent only on the rows of $\mathbf{A}_1^+(\mathbf{k})$.

c) a set of q_{23} rows $\mathbf{A}_3^+(\mathbf{k})$ which are linearly dependent on the rows of $\mathbf{B}_1^+(\mathbf{k})$ and possibly on the rows of $\mathbf{A}_1^+(\mathbf{k})$.

Let the corresponding sets of constraints respectively be :

$$\mathbf{A}_1^+(\mathbf{k})\mathbf{F}(\mathbf{k}) = \mathbf{0} \tag{28}$$

$$\mathbf{A}_2^+(\mathbf{k})\mathbf{F}(\mathbf{k}) = \mathbf{0} \tag{29}$$

$$\mathbf{A}_3^+(\mathbf{k})\mathbf{F}(\mathbf{k}) = \mathbf{0} \tag{30}$$

Since the rows of $\mathbf{A}_2^+$ are linearly dependent on those of $\mathbf{A}_1^+$, Eq. (29) will be satisfied once Eq(28) is so. Therefore the constraints of (28) will be retained and those of (29) will be dropped from consideration; however the interferences corresponding to the rows of (29) will enjoy the same suppression at the k th discrete frequency as those corresponding to the rows of (28).

Based on the above classification, matrix $\mathbf{A}_3(\mathbf{k})$ can be expressed as:

$$\mathbf{A}_3(\mathbf{k}) = \mathbf{B}_1(\mathbf{k})\mathbf{P} + \mathbf{A}_1(\mathbf{k})\mathbf{R} \tag{31}$$

where $\mathbf{P}$ and $\mathbf{R}$ are matrices of dimension $q_{11} \times q_{23}$ and $q_{21} \times q_{23}$ respectively and where $\mathbf{P}$ cannot be zero while $\mathbf{R}$ can be zero. Using (31),(23) and (28), the left-hand side of (30) reduces to:

$$\begin{aligned} \mathbf{A}_3^+(\mathbf{k})\mathbf{F}(\mathbf{k}) &= \mathbf{P}^+\mathbf{B}_1^+(\mathbf{k})\mathbf{F}(\mathbf{k}) + \mathbf{R}^+\mathbf{A}_1^+(\mathbf{k})\mathbf{F}(\mathbf{k}) \\ &= \mathbf{P}^+\boldsymbol{\mu} \end{aligned} \tag{32}$$

Consequently constraint (30) cannot be satisfied except in the very unexpected case when the right hand side of (32) is accidentally the zero vector. Moreover the norm of the difference between both sides of (30) cannot be minimized since by virtue of (32), $\mathbf{A}_3^+(\mathbf{k})\mathbf{F}(\mathbf{k})$ is a constant vector that is independent of $\mathbf{F}(\mathbf{k})$ as long as (23) and (28) are satisfied. Therefore nothing can be done for attenuating the interferences corresponding to the rows of matrix $\mathbf{A}_3^+(\mathbf{k})$. However since the situation of having redundant and/or inconsistent constraints can be different for different values of the discrete frequency k , even those interferences corresponding to the infeasible constraints of (30) will receive some attenuation on the average.

The q_{11} all-pass constraints of (23) and the q_{21} suppression constraints of (28) form a set of consistent and independent constraints that can be compactly expressed as:

$$\mathbf{C}^+(\mathbf{k})\mathbf{F}(\mathbf{k}) = \mathbf{d} \tag{33}$$

where $\mathbf{C}(\mathbf{k})$ is the $N \times (q_{11} + q_{21})$ partitioned matrix:

$$\mathbf{C}(\mathbf{k}) = \left(\mathbf{B}_1(\mathbf{k}) \;\vdots\; \mathbf{A}_1(\mathbf{k})\right) \tag{34}$$

and $\mathbf{d}$ is the $(q_{11} + q_{21})$-dimensional partitioned vector:

$$\mathbf{d} = \begin{pmatrix} \boldsymbol{\mu} \\ \cdots \\ \mathbf{0} \end{pmatrix}. \tag{35}$$

Matrix $\mathbf{C}(\mathbf{k})$ has a full column rank because of the way the matrices $\mathbf{B}_1(\mathbf{k})$ and $\mathbf{A}_1(\mathbf{k})$ have been defined. The linear system (33) is underdetermined based on assumption (22) and consequently it has a family of solution vectors $\mathbf{F}(\mathbf{k})$.

Constrained Optimization

The freedom in the solution space of (33) will be exploited in minimizing criterion (18) that accounts for attenuating the random noise. In the Appendix, the unique *complex* vector $\mathbf{F(k)}$ that minimizes (18) subject to (33), will be derived using the Lagrange multipliers technique [17-19] to get:

$$\mathbf{F(k)} = \mathbf{G}^{-1}\mathbf{C}\left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)^{-1}\mathbf{d} \tag{36}$$

One should notice that the matrix $\left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)$ is nonsingular since $\mathbf{G}$ is nonsingular and $\mathbf{C}$ has a full column rank. Since the above vector $\mathbf{F(k)}$ is invariant under scaling of the diagonal matrix $\mathbf{G}$ of (17), one concludes that $\mathbf{F(k)}$ depends only on the relative values of the variances σ_n^2's rather than on their absolute values.

From (34), one gets:

$$\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C} = \begin{pmatrix} \mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1 & \mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1 \\ \mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1 & \mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1 \end{pmatrix} \tag{37}$$

and using the form of the inverse of a partitioned matrix [20] and substituting (35), one obtains:

$$\left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)^{-1}\mathbf{d} = \begin{pmatrix} \left[\mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1 - \mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1\right)\right]^{-1}\boldsymbol{\mu} \\ -\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1\right)\left[\mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1 - \mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1\right)\right]^{-1}\boldsymbol{\mu} \end{pmatrix} \tag{38}$$

Substituting from (38) and (34) into (36), one gets :

$$\mathbf{F(k)} = \mathbf{G}^{-1}\left[\mathbf{B}_1 - \mathbf{A}_1\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1\right)\right]\left[\mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1 - \mathbf{B}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_1^{+}\mathbf{G}^{-1}\mathbf{B}_1\right)\right]^{-1}\boldsymbol{\mu} \tag{39}$$

This is the optimal filter for extracting *multiple signals* by attenuating multiple interferences and random noise. It extends and generalizes the previous result obtained by the author for the special case of extracting a *single signal* under the same assumptions [9].

A Comparison with Other Beamforming Methods

The signal extraction technique reported in this paper applies mainly to *seismic data* where the processing of the recording of the array of *seismometers* is done off-line. Consequently the descriptor *beamformer* used here is distinct from that used in other work [21] where the beamformer is defined as a signal processor used in conjunction with a set of *antennas* that are spatially separated. By an expert examination of the seismic traces, the geophysicists equipped by their knowledge about the location where seismic data have been acquired can determine the desired signals and interferences. Next they estimate the delay times of the signals and interferences by applying any of the available delay time estimation techniques [10,11].

In the method presented in this paper, the minimization criterion has been the output power spectral density in response to *only the random noise* as given by (18). Therefore it is distinct from the multiple sidelobe canceler (MSC) method where the weights are chosen by minimizing the expected value of the *total* output power with the possible disadvantage of the cancellation of the desired signal [21-23]. It is also distinct from the reference signal method [22,24] where the weights are evaluated by minimizing the mean square error between the beamformer output and the reference signal necessitating the generation of the latter.

In addition to using a *different minimization criterion*, the technique of this paper is quite distinct from the linearly constrained minimum variance (LCMV) beamforming method [21,22,25] in other aspects. First, identifying *redundant and/or inconsistent constraints* has been a main focus in the derivation presented here while in the LCMV method the linear independence and consistency of the constraints were taken for granted. Although having linearly dependent or conflicting constraints does not arise very often, ignoring this issue can lead to singular matrices and invalid results. The situation is aggravated by the fact that different constraints are used for different discrete frequencies. Second, in the point constraints version of the LCMV method [26], point constraints are used for fixing the beamformer response at points of *both* spatial direction *and* temporal frequency thus consuming many of the degrees of freedom represented by the number of weights of the beamformer and consequently reducing the degrees of freedom left for random noise attenuation [22]. In contrast, in the present method the constraints fix the beamformer response at points of *only* spatial direction since a new vector $\mathbf{F}(\mathbf{k})$ is evaluated for each discrete frequency k. Consequently more degrees of freedom are left for random noise attenuation.

The technique of this paper is distinct from those presented in the book by Johnson and Dudgeon [19] since the main objective here is the *extraction of signals* having arbitrary waveforms. In contrast, the eigenanalysis algorithms of [19] are direction of arrival estimation techniques meant for finding the direction of propagation of the signals rather than estimating their waveforms.

In conclusion, the unique features of the method of this paper are: extracting signals of *arbitrary waveforms* by attenuating interferences of *arbitrary waveforms*, deriving an *explicit expression* for the DFT representation of the vector of array filters $\mathbf{F}(\mathbf{k})$ in a mathematical framework which identifies any *redundant and/or inconsistent* constraints, and working with *arbitrary array geometry* rather than a uniform linear array. Therefore, the approach followed here is distinct from that of other researchers [19,22-29]. It is a clear extension and generalization of the work reported in [7-9] under the same assumptions.

IV. IMPLEMENTATION OF THE ARRAY FILTERS

In order to implement the array filters derived above, vector $\mathbf{F}(\mathbf{k})$ of (36) should be computed for all discrete frequencies $k = 0,\cdots,K-1$ and substituted in (9a) to get Y(k). As a preliminary step for computing vector $\mathbf{F}(\mathbf{k})$ one should extract matrix $\mathbf{B_1}$ from the columns of matrix $\mathbf{B}$ of (6) and extract matrix $\mathbf{A_1}$ from the columns of matrix $\mathbf{A}$ of (7).

First, the QR matrix decomposition technique with column pivoting is applied to matrix $\mathbf{B}$ to determine its rank r_1 and a set of r_1 linearly independent columns, which will form matrix $\mathbf{B_1}$ [30]. More specifically:

$$\mathbf{Q}^{+}\mathbf{BP} = \begin{pmatrix} \mathbf{R_{11}} & \mathbf{R_{12}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{40}$$

where $\mathbf{Q}$ is a unitary matrix of order N, $\mathbf{P}$ is a permutation matrix of order q_1, $\mathbf{R_{11}}$ is a nonsingular upper triangular matrix of order r_1 and $\mathbf{R_{12}}$ is a rectangular matrix which arises only if $\mathbf{B}$ is rank deficient $(r_1 < q_1)$. Strictly speaking, the QR decomposition has been computed for the matrix $\mathbf{BP}$ rather than for the matrix $\mathbf{B}$. Since $\mathbf{P}$ is a permutation matrix (i.e. a matrix obtained from the identity matrix by permuting its columns), each column of $\mathbf{P}$ will have only one nonzero element whose value is the unity. If matrix $\mathbf{B}$ is expressed as the column partitioned matrix $\mathbf{B} = (\mathbf{b_1} \quad \mathbf{b_2} \quad \cdots \quad \mathbf{b_{q_1}})$ and if the j th column of $\mathbf{P}$ is $\mathbf{e_i}$ (the unit vector whose only nonzero element is the i th one), then the j th column of $\mathbf{BP}$ will be $\mathbf{b_i}$ rather than $\mathbf{b_j}$. In this manner one can generate matrix $\mathbf{B_1}$ by selecting the first r_1 columns of $\mathbf{BP}$.

Second, matrix $\mathbf{A_1}$ can be generated by forming the augmented matrix:

$$\mathbf{E} = \left(\mathbf{B_1} \quad \vdots \quad \mathbf{A}\right) \tag{41}$$

and finding its QR decomposition while freezing the first r_1 columns and allowing column pivoting for the remaining columns. Matrix $\mathbf{A_1}$ is then formed by the independent columns of $\mathbf{E}$ apart from those forming $\mathbf{B_1}$.

Having prepared the matrices $\mathbf{B_1}$ and $\mathbf{A_1}$, one turns to the main task of computing vector $\mathbf{F(k)}$ of (36). The accuracy of the entire array filtering process depends on the numerical accuracy and stability of the computational procedure used for evaluating $\mathbf{F(k)}$. A careful look at (36) reveals that the heart of the procedure is the solution of the following linear system of equations :

$$\left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)\xi = \boldsymbol{\eta}. \tag{42}$$

where $\boldsymbol{\eta} = \mathbf{d}$. Because of the diagonal structure of matrix $\mathbf{G}$ of (17), the above linear system can be expressed as:

$$\left(\mathbf{W}^{+}\mathbf{W}\right)\boldsymbol{\xi} = \boldsymbol{\eta} \tag{43}$$

where $\mathbf{W}$ is the N x NC rectangular matrix defined by:

$$\mathbf{W} = Diag\left\{\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_N}\right\}\mathbf{C} \tag{44}$$

and

$$NC = q_{11} + q_{21} \ . \tag{45}$$

Matrix $\mathbf{W}$ has a full column rank since $\mathbf{C}$ has the same property and the premultiplying diagonal matrix in (44) is nonsingular.

In order to accurately solve the system (43), the QR decomposition with the pivoting option is applied to matrix $\mathbf{W}$ to get:

$$\mathbf{W} = \mathbf{QRP}^{+} \tag{46}$$

where $\mathbf{Q}$ is a unitary matrix of order N obtained as the product of NC Householder transformation matrices and $\mathbf{P}$ is a permutation matrix of order NC. The N x NC matrix $\mathbf{R}$ has the form:

$$\mathbf{R} = \begin{pmatrix} \widetilde{\mathbf{R}} \\ \cdots \\ \mathbf{0} \end{pmatrix} \tag{47}$$

where $\widetilde{\mathbf{R}}$ is an upper triangular matrix of order NC. Matrix $\widetilde{\mathbf{R}}$ is nonsingular since matrix $\mathbf{R}$ in (46) like matrix $\mathbf{W}$, has a full column rank. From (46) and (47) and after exploiting the unitarity of $\mathbf{Q}$, one gets:

$$\mathbf{W}^{+}\mathbf{W} = \mathbf{P}\widetilde{\mathbf{R}}^{+}\widetilde{\mathbf{R}}\mathbf{P}^{+} \tag{48}$$

and consequently the linear system (43) can be expressed as:

$$\widetilde{\mathbf{R}}^{+}\widetilde{\mathbf{R}}\left(\mathbf{P}^{+}\boldsymbol{\xi}\right) = \mathbf{P}^{+}\boldsymbol{\eta} \;. \tag{49}$$

Therefore vector $\boldsymbol{\xi}$ can be evaluated by scrambling vector $\boldsymbol{\eta}$ to get $\mathbf{P}^{+}\boldsymbol{\eta}$, solving two triangular systems of equations, and finally unscrambling the solution vector $\left(\mathbf{P}^{+}\boldsymbol{\xi}\right)$ to get $\boldsymbol{\xi}$.

In summary the algorithm for computing $\mathbf{F}(\mathbf{k})$ is:

1. Apply the QR technique to $\mathbf{B}$ in order to identify $\mathbf{B}_1$.
2. Apply the QR technique to $\mathbf{E}$ of (41) in order to identify $\mathbf{A}_1$.
3. Form matrix $\mathbf{C}$ of (34) and $\mathbf{W}$ of (44).
4. Apply the QR technique to $\mathbf{W}$ to get $\widetilde{\mathbf{R}}$.
5. Solve the two triangular systems of (49) with $\boldsymbol{\eta} = \mathbf{d}$ defined by (35) in order to get $\boldsymbol{\xi}$.
6. Compute $\mathbf{F}(\mathbf{k}) = \mathbf{G}^{-1}\mathbf{C}\boldsymbol{\xi}$.

The LINPACK software package [30] was used for performing the computations delineated in this section; and the complex version of the subroutines was chosen because of the nature of the involved matrices. The subroutine CQRDC was used for QR decomposing $\mathbf{B}$ in (40), $\mathbf{E}$ of (41) and $\mathbf{W}$ in (46). Finally two calls to the subroutine CTRSL were needed for solving (49). It should be emphasized that the last step cannot be replaced by a single call to the subroutine CPOSL because the triangular matrix $\widetilde{\mathbf{R}}$ in (49) may have negative diagonal elements since it has been obtained by the QR decomposition of the rectangular matrix $\mathbf{W}$ in (46) rather than by the Cholesky decomposition of the square positive definite matrix $\mathbf{W}^{+}\mathbf{W}$ in (48).

The approximate number of flops (floating point operations), where an operation stands for a multiplication and addition, for the above algorithm is given in Table 1. Since $q_{11} \le q_1$ and $q_{21} \le q_2$, the total number of flops required for computing $\mathbf{F}(\mathbf{k})$ can be approximated by:

$$Nq_1^2 - \frac{1}{3}q_1^3 + (2N+1)(q_1+q_2)^2 - \frac{2}{3}(q_1+q_2)^3 + (2N+1)(q_1+q_2).$$

V. SIMULATION RESULTS

Since the model (1) of array recordings well suits the seismic data where the signals are reflections from subsurface layers and the interferences are different modes of propagation of the ground roll traveling directly from the source of the explosion to the geophones, synthetic seismic traces will be generated and processed in this section. In the first example the data will be processed according to the multiple input single output scheme of Fig. 1 and in the second example the multiple input multiple output scheme, to be explained below, will be used.

A Multiple Input Single Output Example

The 16 synthetic seismic traces of Fig. 2 have been generated by combining the three signals of Fig. 3 - shown in their locations on the first trace (reference trace) - and the three interferences of Fig. 4 (also shown in their locations on the first trace) in addition to random noise according to the model of Eq. (1). The three desired signals are assumed to be traveling up and to have commensurate delay times corresponding to intersensor travel times of 2, 4, and 8 respectively where all times are in terms of the sampling period. The three interferences are assumed to be traveling down and to have commensurate delay times corresponding to intersensor travel times of 8, 6, and 4 respectively. All interferences have the same energy level which is 6 dB above that of any of the desired signals which all have the same energy level. The amplitude scale factors β_{nm} of the desired signals are all unity while their counterparts α_{nm} of the interferences are given in Table 2. The random sensor noises have been produced by a generator of pseudorandom numbers from a standard normal distribution and then scaled to be 13.979 dB below the level of any of the desired signals (in order for the amplitude level of the random noise to be 0.2 times that of any desired signal). The length of each trace is K = 800 samples which is adequate to incorporate all signals and interferences irrespective of their delay times as can be seen from Fig. 2. Figure 5a shows all the desired signals together as they should appear on the first trace (reference trace) and Fig. 5b shows the first trace (having signals, interferences and random noise) drawn to the same scale as Fig. 5a. One can hardly recognize the desired signals in the traces of Fig. 2.

After transforming the input array traces by the FFT with the same length K = 800, the resulting array data have been processed by the filters derived in Section III and the output is shown in Fig. 6. By comparing Figs. 5a and 6, it is obvious that the array filters have been able to extract the three desired signals. Actually the desired signals are quite identifiable on the output trace of Fig. 6 and not identifiable on the input reference trace of Fig. 5b. In order to interpret the discrepancy between Figs. 6 and 5a, one starts by noticing that having commensurate delay times for the signals implies that in (1):

$$\xi_{nm} = -(n-1)d_m \qquad , n = 1 , \ldots , N \tag{50}$$

where $d_1 = 2$, $d_2 = 4$ and $d_3 = 8$. Since the amplitude scale factors β_{nm}'s are all unity, Eqs. (4) and (6) imply that matrix B(k) is a 16 x 3 Vandermonde matrix where all elements of the first row are unity. Consequently the right-hand side vector of (20) is equal to the first column of the coefficient matrix $B^{+}(k)$, implying that the linear system of equations (20) is consistent [16]. Hence the set of inconsistent equations of the form of (25) is empty for all values of the discrete frequency k. Therefore the discrepancy between the output in Fig. 6 and the sum of the signals in Fig. 5a is only due to random noise and residual interference. In order to verify the heuristic claim that this discrepancy is mainly due to random noise, the same synthetic traces were generated without adding random noise and the result of filtering is shown in Fig. 7 which astonishingly compares perfectly with Fig. 5a. The interpretation is that using nonequal amplitude factors α_{nm}'s for the interferences in (1) as given in Table 2 tends to decrease the chance of having a nonempty subsystem of the form of (30) which is responsible for the residual interference in the output. In order to further verify this heuristic interpretation, the same random noise free input traces were generated using $\alpha_{nm} = 1$ (instead of the values of Table 2) and the result of filtering is shown in Fig. 8 where the residual interference is noticeable.

A Multiple Input Multiple Output Example

Since the recording of a single sensor of the array is a set of temporal samples and since at each recording instant the recording of the array is a set of spatial samples, the acquired array data are two dimensional in nature.

If the output of the filtering process is to be used for an application such as seismic migration where having 2D data is a requirement, the single-output processing scheme of Fig. 1 will no longer be adequate. Multiple output can be generated by employing a sliding window that can only cover M traces of the available N traces (M << N). Initially, the window covers traces 1, 2,..., M and the corresponding output is computed as in the single output scheme (using M rather than N traces), then the window slides one location to cover traces 2, 3,..., (M+1) and so on until eventually the window covers traces N-M+1, N-M+2,..., N. The resulting processing scheme will have N input and (N - M + 1) output traces [31]. There are two factors to be taken into account in the selection of the window length M. First since the filter $\mathbf{F}(\mathbf{k})$ of (36) will have M instead of N components, constraint (22) implies that $M > q_1 + q_2$; actually the larger the value of M the better will the filter be able to attenuate the interferences and random noise. Second since the number of output traces is only (N-M+1) compared to the N input traces, one should have M << N. Of course it is recommended to have a large value of N when using the multiple input multiple output scheme.

The 24 synthetic input traces of Fig. 9 have been generated by combining the first two signals of Fig. 3 and the first two interferences of Fig. 4 in addition to random noise. The two signals are up traveling with intersensor delay times of 11 and 4 units and equal β_{nm}'s; and the two interferences are down propagating with intersensor delay times of 8 and 6 units and nonequal α_{nm}'s. A sliding window of length M = 8 was used for generating the output traces of Fig. 10. It is obvious that the array filters have been able to extract the two desired signals.

VI. CONCLUSION

A vector of array filters has been derived for processing the recordings of an array of sensors in a multichannel scheme with the objective of extracting *multiple signals* by attenuating multiple interferences and random noise. The rationale adopted in the derivation is based on the following ranking of objectives: achieving all-pass conditions (whenever possible) for the signals, imposing complete suppression (whenever feasible) for the interferences, and attenuating the random sensor noise. This model well suits the *seismic data* which are dominantly corrupted by interferences rather than random noise. The derived filters have been successfully applied in both the multiple input single output and the multiple input multiple output processing schemes where in the latter a window is slid on the input traces and for each window position the output is obtained as in the single output scheme.

APPENDIX

Statement of the Problem :

Find vector $\mathbf{F}$ which minimizes

$$J = \mathbf{F}^{+}\mathbf{GF} \tag{A1}$$

where $\mathbf{G}$ is a nonsingular Hermitian matrix[7] subject to the linear constraints :

$$\mathbf{C}^{+}\mathbf{F} = \mathbf{d} \tag{A2}$$

where matrix $\mathbf{C}$ has a full column rank.

Solution :

The complex constraints of (A2) can be augmented to the real criterion of (A1) through the complex vector $\mathbf{p}$ of Lagrange multipliers to get the augmented criterion :

$$J_a = J - \mathbf{p}^{+}\left(\mathbf{C}^{+}\mathbf{F} - \mathbf{d}\right) - \left(\mathbf{C}^{+}\mathbf{F} - \mathbf{d}\right)^{+}\mathbf{p} \tag{A3}$$

Substituting (A1) into (A3), one gets :

$$J_a = \mathbf{F}^{+}\mathbf{GF} - \mathbf{q}^{+}\mathbf{F} - \mathbf{F}^{+}\mathbf{q} + \mathbf{p}^{+}\mathbf{d} + \mathbf{d}^{+}\mathbf{p} \tag{A4}$$

where

$$\mathbf{q} = \mathbf{Cp} \tag{A5}$$

The minimizer of J_a with respect to the complex vector $\mathbf{F}$ is [18]:

$$\mathbf{F} = \mathbf{G}^{-1}\mathbf{q} \tag{A6}$$

Substituting this vector in (A2) and using (A5), one gets:

$$\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{Cp} = \mathbf{d}\ . \tag{A7}$$

The matrix $\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}$ is nonsingular since $\mathbf{G}$ is nonsingular and $\mathbf{C}$ has a full column rank. Consequently the Lagrange multipliers vector is given by:

---

[7] Although matrix $\mathbf{G}$ in (18) is real and diagonal, it has been allowed to be any Hermitian matrix in the derivation of this appendix for the sake of generality.

$$\mathbf{p} = \left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)^{-1}\mathbf{d} \ . \tag{A8}$$

Substituting (A5) and (A8) in (A6), one gets:

$$\mathbf{F} = \mathbf{G}^{-1}\mathbf{C}\left(\mathbf{C}^{+}\mathbf{G}^{-1}\mathbf{C}\right)^{-1}\mathbf{d} \ . \tag{A9}$$

REFERENCES

[1] J.H. Justice, "Array processing in exploration seismology," in S. Haykin, Ed., *Array Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985.

[2] M. Simaan, Ed., *Advances in Geophysical Data Processing, Volume 1: Vertical Seismic Profiles*. Greenwich, CT: JAI Press, 1984.

[3] E.A. Robinson and T.S. Durrani, *Geophysical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[4] D.E. Dudgeon, "Fundamentals of digital array processing," *Proceedings of the IEEE*, vol. 65, no. 6, pp. 898 - 904, June 1977.

[5] M. Holzman, "Chebyshev optimized geophone arrays," *Geophysics*, vol. 28, no. 2, pp. 145 - 155, April 1963.

[6] M. Simaan, "Optimum array filters for array data signal processing," *IEEE transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-31, no. 4, pp. 1006 - 1015, August 1983.

[7] M.T. Hanna and M. Simaan, "Absolutely optimum array filters for sensor arrays," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 6, pp. 1380 - 1386, December 1985.

[8] M.T. Hanna and M. Simaan, "Array filters for attenuating coherent interferences in the presence of random noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 4, pp. 661 - 668, August 1986.

[9] M.T. Hanna, "Array filters for attenuating multiple coherent interference," *IEEE transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-36, no. 6, pp. 844 - 853, June 1988.

[10] G.C. Carter, Guest Ed., Special Issue on Time Delay Estimation, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-29, June 1981.

[11] M. Simaan, "A frequency-domain method for time-shift estimation and alignment of seismic signals," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-23, no. 2, pp. 132-138, March 1985.

[12] A.V. Oppenheim, R.W. Schafer and J.R. Buck, Discrete-Time Signal Processing, 2nd edition, Upper Saddle River, NJ: Prentice-Hall, 1999.

[13] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York : Mc-Graw-Hill, 3rd edition, 1991.

[14] Petre Stoica and Randolph L. Moses, *Introduction to Spectral Analysis*, Upper Saddle River, NJ : Prentice-Hall, 1997.

[15] Henry Stark and John Woods, *Probability and Random Processes with Applications to Signal Processing*, 3rd Ed., Englewood Cliffs, NJ: Prentice-Hall, 2002.

[16] G. Strang, *Linear Algebra and its applications*, 2nd. Edition, New York : Academic Press, 1980.

[17] Todd K. Moon and Wynn C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 2000.

[18] D.H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings, vol. 130, Parts F and H*, no. 1, pp. 11-16, February 1983.

[19] D.H. Johnson and D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[20] W.L. Brogan, *Modern Control Theory*. Englewood Cliffs, NJ: Quantum Publishers, 1982.

[21] B. Van Veen, "Minimum variance beamforming," in S. Haykin and A. Steinhardt, eds., *Adaptive Radar Detection and Estimation*, New York: John Wiley & Sons, pp. 161-236, 1992.

[22] B.D. Van Veen and K.M. Buckley, "Beamforming: A versatile approach to spatial filtering," *The IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4-24, April 1988.

[23] S.P. Applebaum and D.J. Chapman, "Adaptive arrays with main beam constraints," *IEEE Transactions on Antennas and Propagation*, vol. AP-24, pp. 650-662, Sept. 1976.

[24] B. Widrow, P.E. Mantey, L.J. Griffiths and B.B. Goode, "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, pp. 2143-2159, Dec. 1967.

[25] O.L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, pp. 926-935, Aug. 1972.

[26] E.L. Kelly, Jr. and M.L. Levin, "Signal parameter estimation for seismometer arrays," MIT Lincoln Lab, Technical Report 339, Jan. 1964.

[27] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 5, pp. 1081-1089, Oct. 1986.

[28] C.V. Kimball, P. Lewicki and N.I. Wijeyesekera, ″Error analysis of maximum likelihood estimates of physical parameters from one or more dispersive waves,″ *IEEE Transactions on Signal Processing*, vol. SP-43, no. 12, pp. 2928-2936, Dec. 1995.

[29] C. Chambers, T.C. Tozer, K.C. Sharman and T.S. Durrani, ″Temporal and spatial sampling influence on the estimates of superimposed narrowband signals: When less can mean more,″ *IEEE Transactions on Signal Processing*, vol. SP-44, no. 12, pp. 3085-3098, Dec. 1996.

[30] J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, *LINPACK Users' Guide*. Philadelphia, PA: SIAM, 1979.

[31] M.T. Hanna, "Velocity Filters for Multiple Interference Attenuation in Geophysical Array Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-26, no. 6, pp. 741-748, November 1988.

Table 1: The computational cost for $\mathbf{F}(\mathbf{k})$ based on the algorithm of section IV.

Step	*Number of flops*
1	$Nq_1^2 - \frac{1}{3}q_1^3$
2	$N(q_{11}+q_2)^2 - \frac{1}{3}(q_{11}+q_2)^3$
3	$N(q_{11}+q_{21})$
4	$N(q_{11}+q_{21})^2 - \frac{1}{3}(q_{11}+q_{21})^3$
5	$(q_{11}+q_{21})^2 + (q_{11}+q_{21})$
6	$N(q_{11}+q_{21})$

Table 2: The amplitude scale factors α_{nm} 's of the interferences.

n	α_{n1}	α_{n2}	α_{n3}	n	α_{n1}	α_{n2}	α_{n3}
1	1	1	1	9	1.08	1.16	1.24
2	1.01	1.02	1.03	10	1.09	1.18	1.27
3	1.02	1.04	1.06	11	1.10	1.20	1.30
4	1.03	1.06	1.09	12	1.11	1.22	1.33
5	1.04	1.08	1.12	13	1.12	1.24	1.36
6	1.05	1.10	1.15	14	1.13	1.26	1.39
7	1.06	1.12	1.18	15	1.14	1.28	1.42
8	1.07	1.14	1.21	16	1.15	1.30	1.45

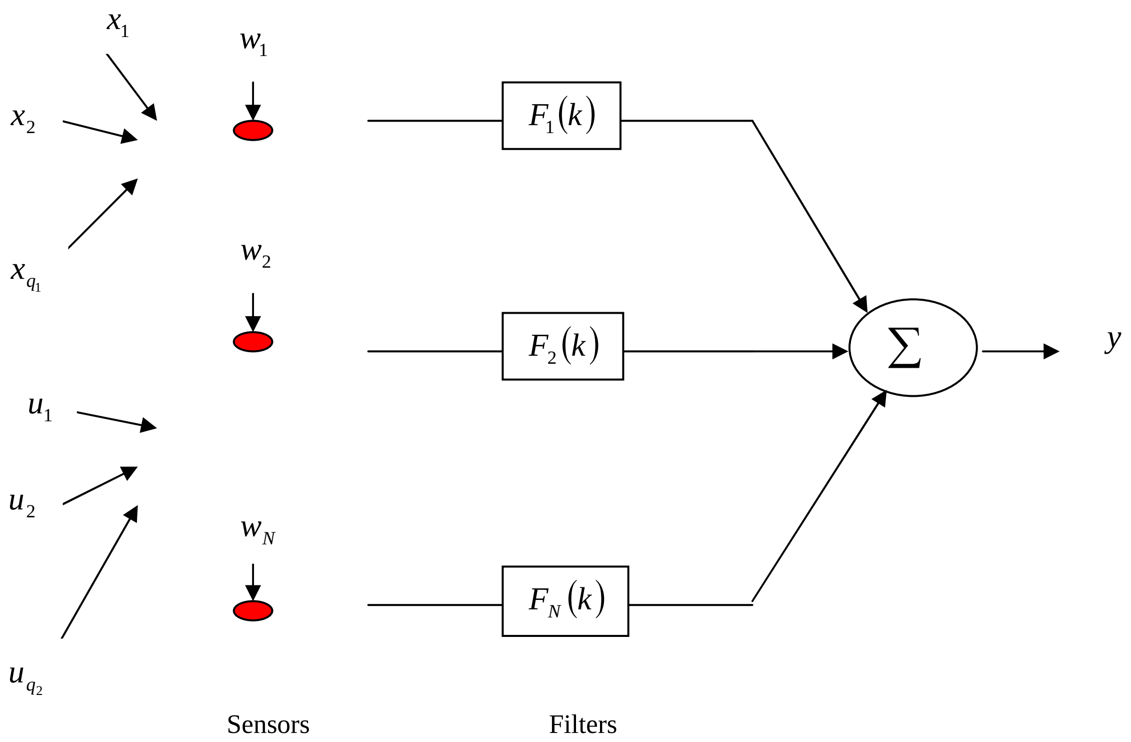

Fig. 1 : Multichannel array processing scheme.

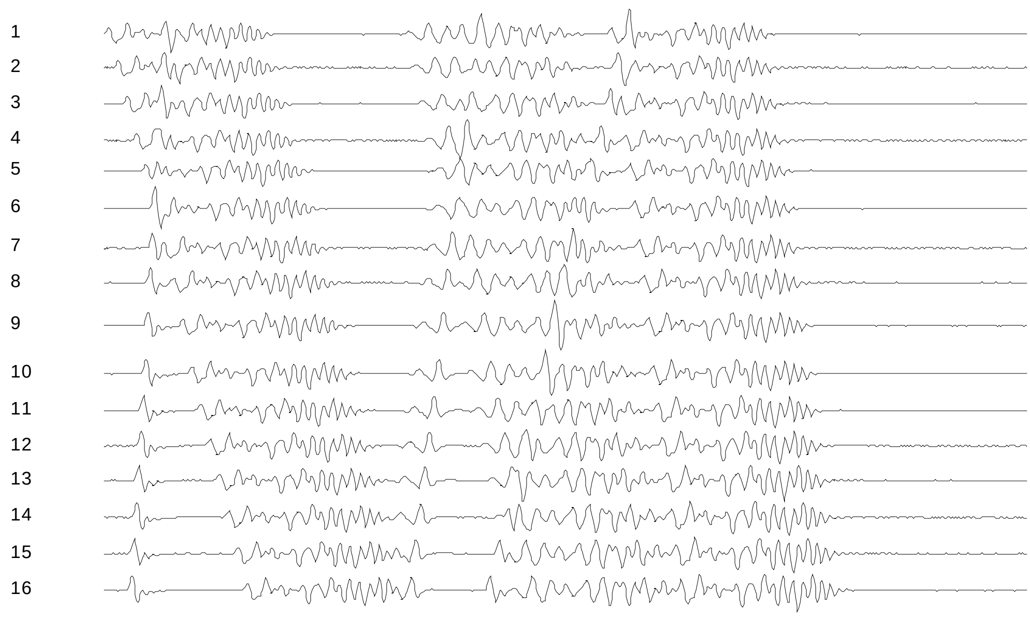

Fig. 2 : Synthetic seismic traces.

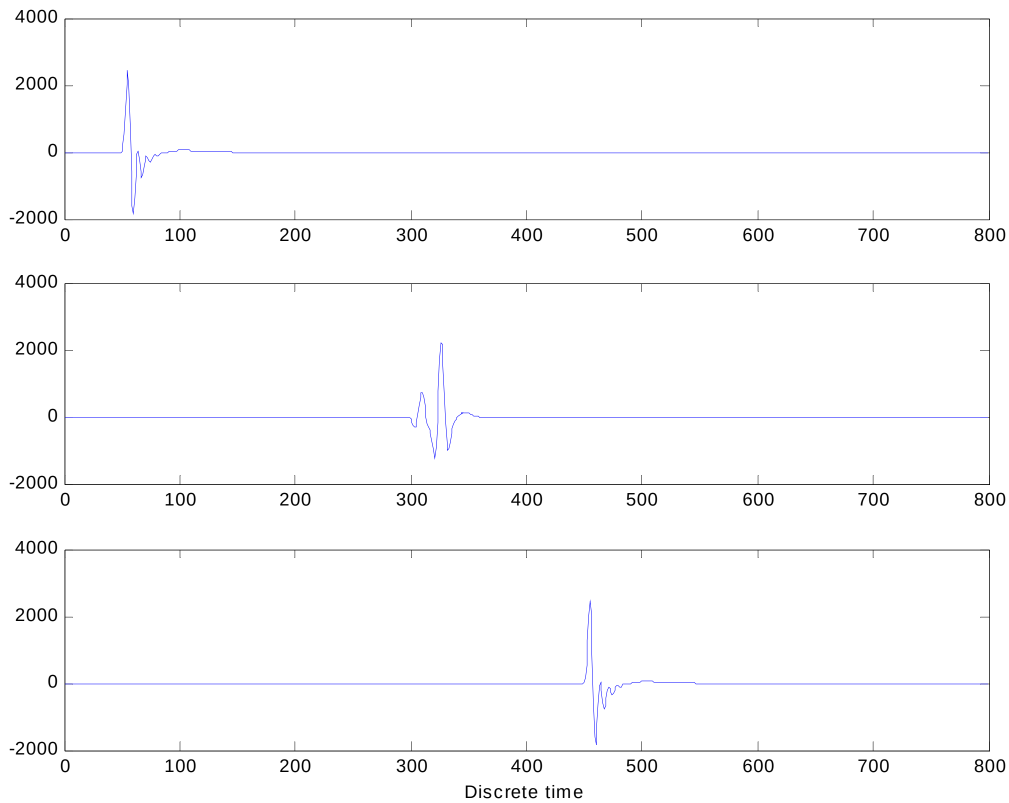

Fig. 3 : The individual desired signals.

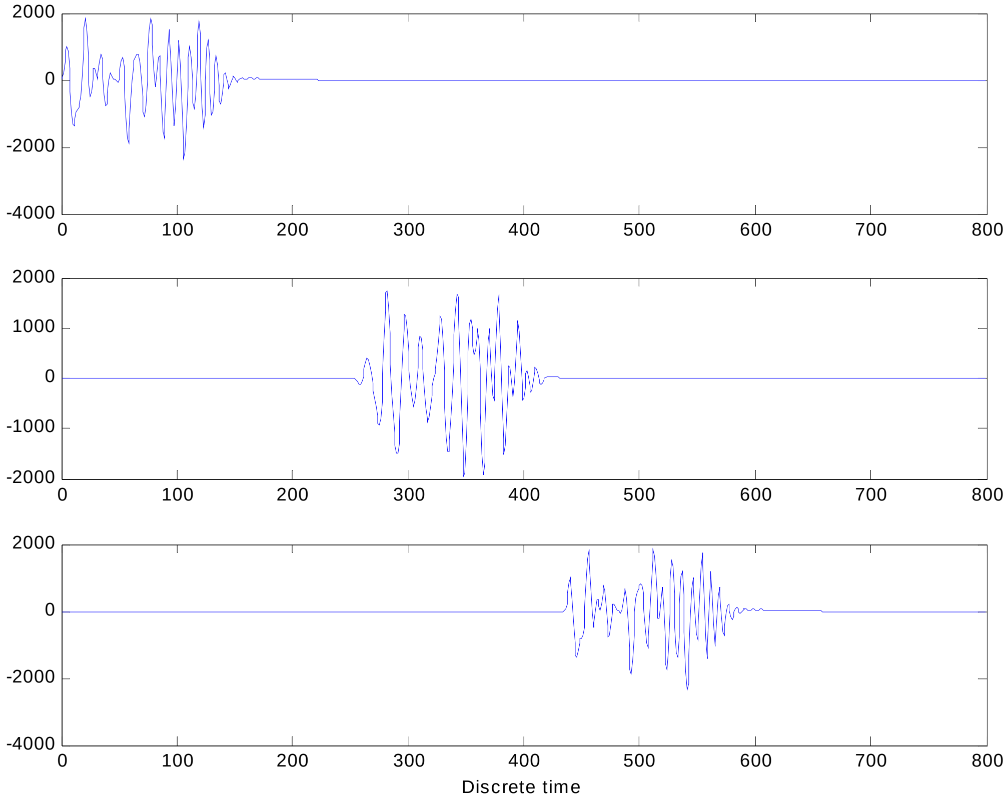

Fig. 4 : The individual coherent interferences.

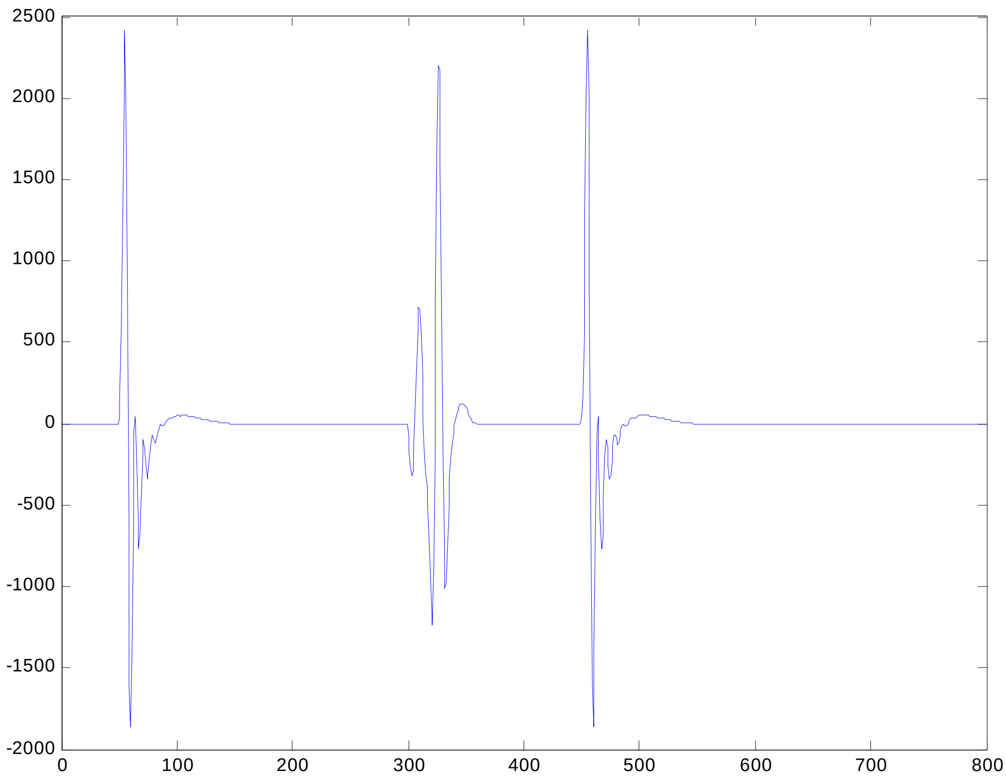

Fig. 5a : The desired signals together as they appear on the first trace.

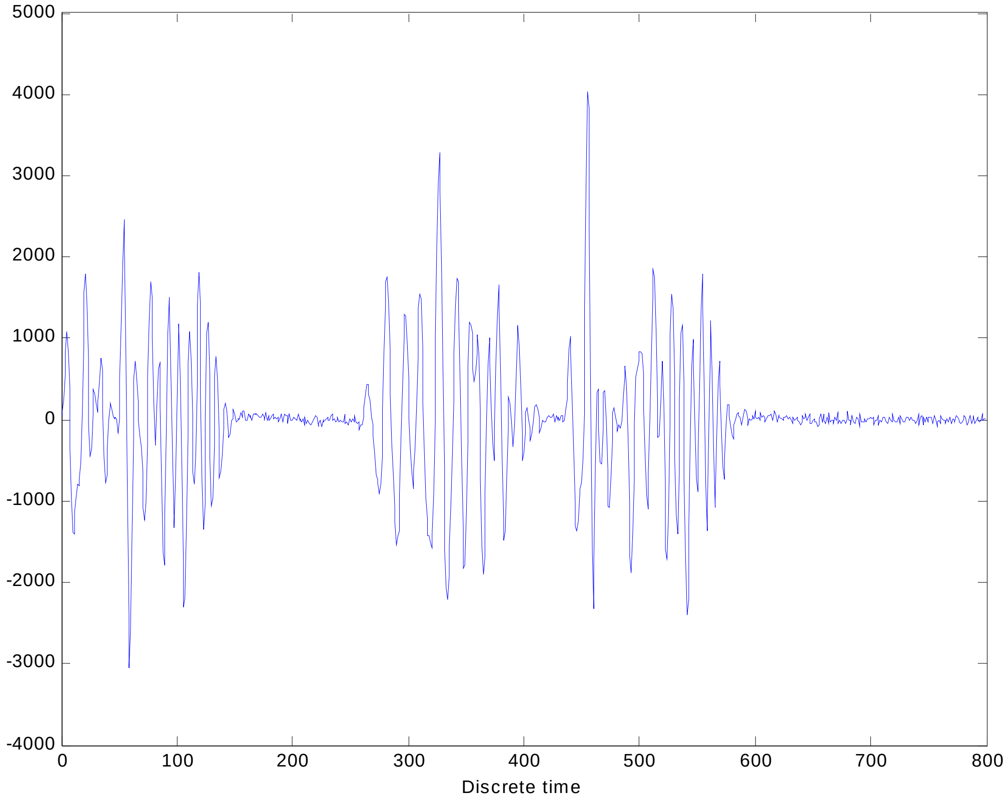

Fig. 5b : The first trace.

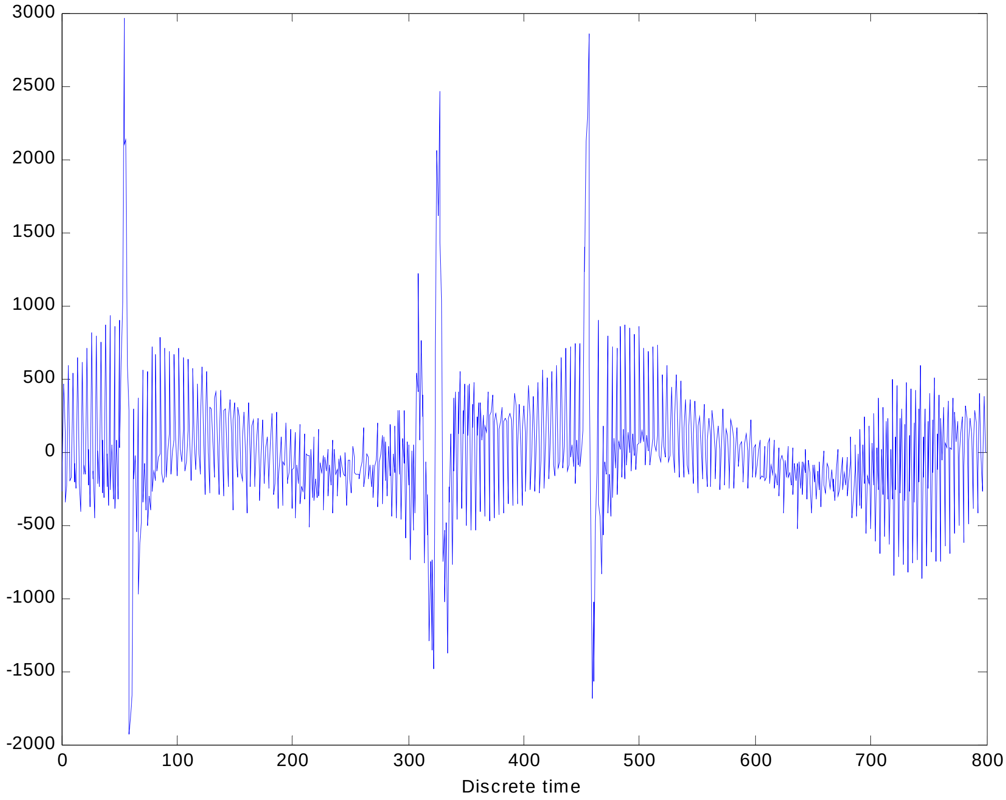

Fig. 6 : The output for noisy input data and N = 16.

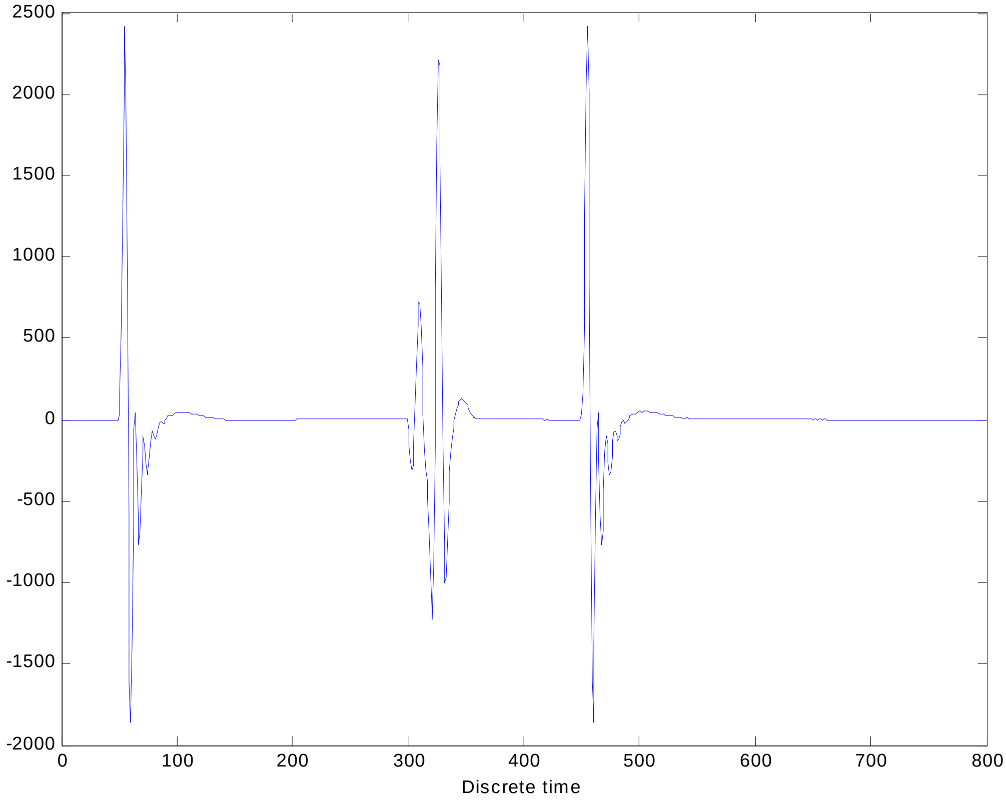

Fig. 7 : The output for noiseless input data and nonequal α_{nm} 's .

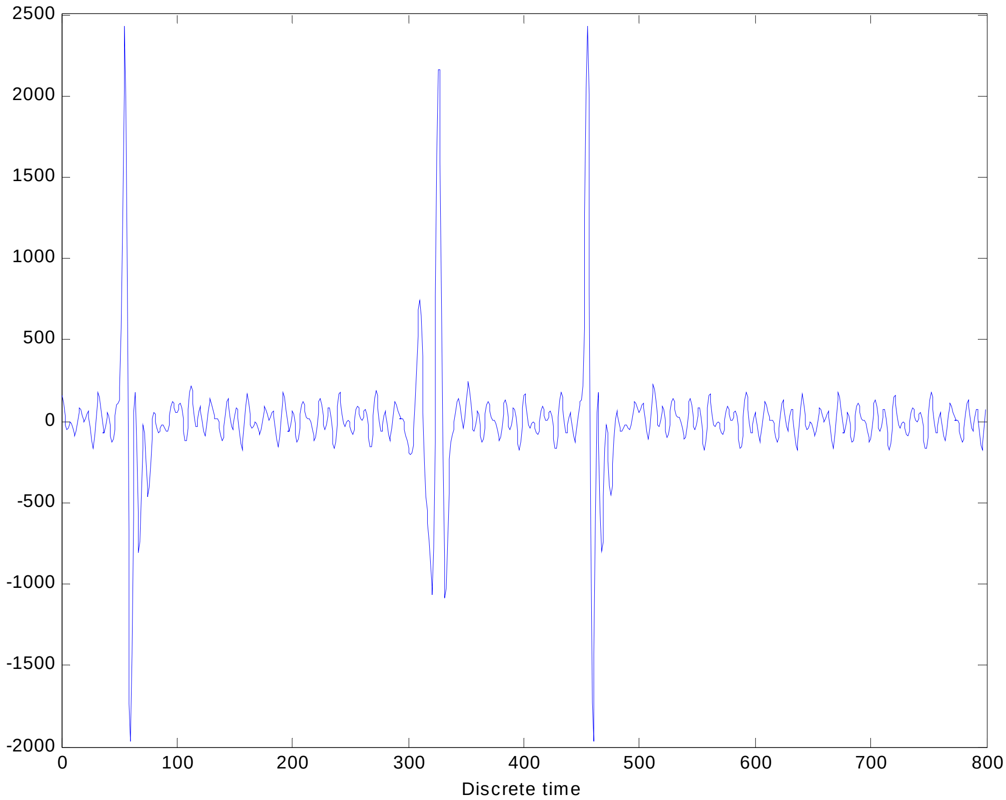

Fig. 8 : The output for noiseless input data and equal α_{nm} 's .

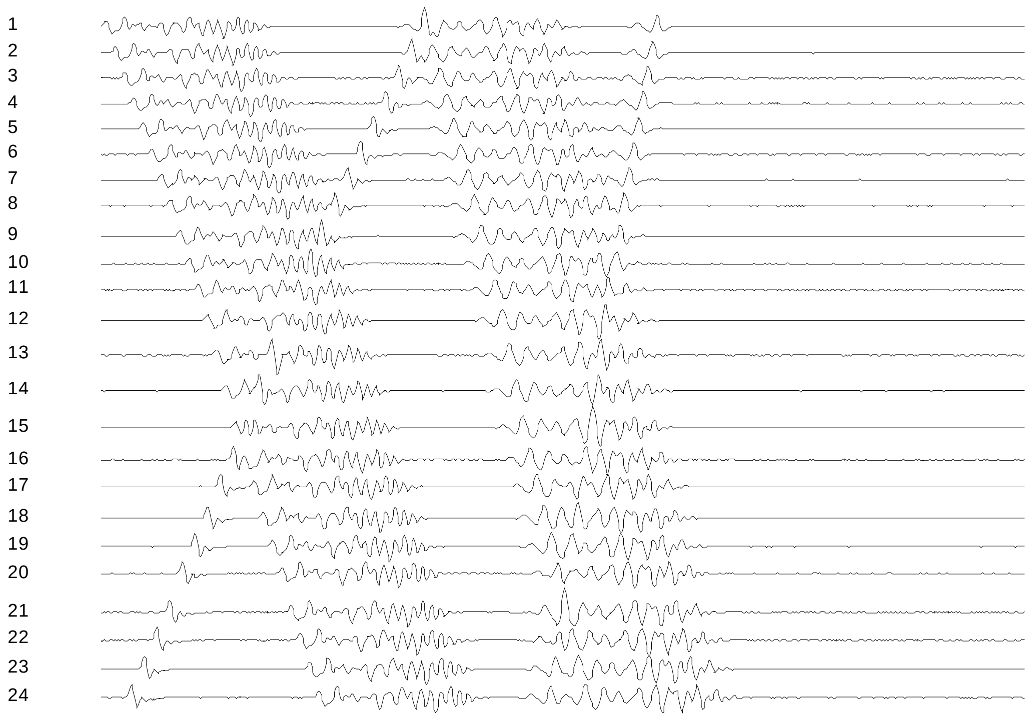

Fig. 9 : The input traces of the multiple input multiple output example.

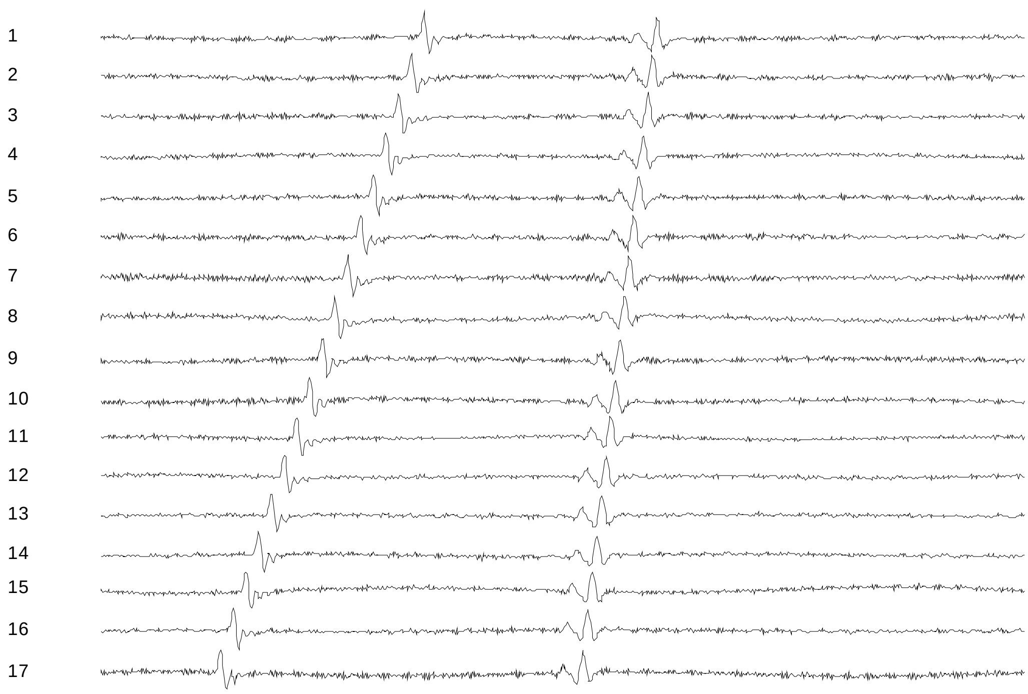

Fig. 10 : The output traces of the multiple input multiple output example.