Factorization Properties of Optimum Space-Time Processors in Nonstationary Environments

BENOIT CHAMPAGNE, MEMBER, IEEE, MOSHE EIZENMAN, MEMBER, IEEE, AND SUBBARAYAN PASUPATHY, SENIOR MEMBER, IEEE

Abstract—In this paper, we present a structural analysis of space-time log-likelihood processors (LLP) that applies to arbitrary signal transmission models consisting of N sources, M sensors, a time-varying linear channel, and nonstationary Gaussian source signal and sensor noise processes. Our approach is based on representing the time-varying linear channel as a bounded linear operator L with closed range. By exploiting the properties of such operators and the specific structure of the array covariance function, we show that the classical M-dimensional integral equations defining the LLP can be transformed into equivalent N-dimensional integral equations. As a result, it is always possible to factor the LLP into a cascade of three specialized time-varying subprocessors, namely: a space-time whitening filter, an M-input N-output unitary beamformer (UB), and an N-input quadratic postprocessor (QPP). This decomposition provides a generalization of conventional results on optimum array processing that were previously derived under the assumption of time-invariant transmission channel and signal statistics. Both the UB and the QPP are given an interpretation and their most important features are indicated. The UB, which is closely related to the generalized inverse of the transmission operator L, is independent of the source signal statistics and maximizes the array gain. Moreover, when $N < M$, it can be used advantageously to reduce the number of time functions that need to be quadratically processed. The QPP behaves like an LLP for the output of the UB and, therefore, it admits a number of standard realizations, both causal and noncausal. Specializations of the above results to the cases of low and high signal-to-noise ratios are also considered. Finally, to illustrate the theory, several examples of its application to signal models involving time-varying delays are given.

I. Introduction

THE problem of space-time signal processing consists of extracting useful information from wavefield measurements taken over both space and time coordinates. In many practical applications, with examples in sonar and seismology, the wavefield of interest is generated by one or more distinct sources and is monitored by a spatial array of multiple sensors in the presence of an additive background noise component. The purpose may be to detect the presence of a source or a scatterer, or to estimate the value of an unknown source or transmission parameter. In both cases, one possible strategy for extracting the relevant information is to optimally process the sensor outputs according to a predetermined statistical criterion. In this respect, the log-likelihood processor (LLP) is of fundamental importance because it is optimal (or at least closely related to the optimal scheme) for a variety of detection and estimation criteria [1]–[4]. For this reason, it is frequently referred to as the optimum processor.

Conventional space-time LLP's such as those used for source detection or bearing estimation [2]–[8] are derived under the following assumptions: 1) stationary source signal and sensor noise processes; 2) time-invariant transmission channel between the sources and the sensors; and 3) long observation intervals. Moreover, propagation effects are usually limited to pure time delays, with possibly discrete multipaths. There are situations, however, where such modeling assumptions are unrealistic. For instance, in the presence of source or receiver motion, which is often the case in sonar applications, the delays in the signal components at the array output are functions of time. In this case, the signal transmission cannot be modeled as a time-invariant transformation and, consequently, classical results on optimum array processing are no longer applicable.

Despite these considerations, a literature review indicates that only a few attempts have been made to study the space-time LLP under time-varying conditions, and even these have been rather limited in scope, focusing mainly on specific time delay models. Schweppe [9] studied the least squares array processor for a multiple sources model. Although this analysis applies to nonstationary (discrete-time) source signals, the propagation model is limited to pure, time-invariant delays. Knapp and Carter [10] derived the maximum likelihood processor for estimating a linearly varying time delay between two noisy versions of a common source signal. More recently, Stuller [11] extended their analysis to the case of arbitrary time-varying delay, nonstationary signal and noise processes and arbitrary observation time. In [12], Lourtie and Moura considered a more general problem formulation which includes multiple nonstationary (state-space representable) source signals and time-varying delays.

In this paper, we study the space-time LLP for a very general signal transmission model consisting of N sources, M sensors, a time-varying linear channel, and nonstationary Gaussian signal and noise processes. We base our approach on representing the time-varying linear channel as a bounded linear operator L with closed range, operating on an infinite dimensional Hilbert space of source signals. By proceeding in this way, reference to particular signal representations such as Fourier coefficients or discrete-

Manuscript received April 1, 1989; revised November 9, 1989.
The authors are with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada M5S 1A4.
IEEE Log Number 9038435.

time samples is avoided. This results in considerable simplifications and enables us to uncover the structural properties of the LLP in the more complicated case where nonstationarities are present in the signal model.

In particular, by exploiting the properties of the operator L and the specific structure of the array covariance function, we show that the classical M-dimensional integral equations defining the LLP can be transformed into equivalent N-dimensional integral equations. As a result, it is always possible to factor the LLP into a cascade of three specialized time-varying subprocessors, namely: a space-time whitening filter, an M-input N-output unitary beamformer (UB) and an N-input quadratic postprocessor (QPP). The UB, which is closely related to the generalized inverse of the transmission operator L, is independent of the source signal statistics and maximizes the array gain. Moreover, when $N < M$, it reduces the number of signals that need to be quadratically processed. The QPP behaves like a LLP for the output of the UB and thus admits a number of standard realizations, both causal and noncausal.

Because of the generality of the signal model considered, many well-known space-time LLP's (both time invariant and time varying) can be obtained directly as particular cases of the factored LLP configurations obtained in this paper. Besides their novelty, these configurations thus play an important role in unifying previous results on array processing.

This paper is organized as follows. Section II describes the signal model considered, reviews the integral equations defining the LLP, and establishes an appropriate framework for the application of linear operator theory. Section III shows that the M-dimensional integral equations defining the LLP are indeed equivalent to N-dimensional integral equations of the same types. The canonical LLP configurations that result from this equivalence are presented in Section IV, where the limiting cases of low and high signal-to-noise ratios are briefly discussed. Section V gives three examples of the application of the theory to signal models involving time-varying delays. Finally, Section VI summarizes the main conclusions of this work. Four Appendices at the end of the paper contain the proofs of various properties and formulae.

II. Problem Formulation

A. Signal Model

A typical source-array configuration is shown in Fig. 1. It consists of N distinct sources, a transmission medium, and an array of M sensors connected to a processing device (M, N arbitrary). During a time interval $I = [t_i, t_f]$, each source radiates a signal waveform $a_i(t)$ ($i = 1, \cdots, N$; $t \in I$) which is transmitted through the medium. The resulting wavefield is monitored by the array in the presence of additive noise during a time interval $J = [T_i, T_f]$. Finally, the sensor outputs $x_i(t)$ ($i = 1, \cdots, M$; $t \in J$) are processed to extract the relevant information. In the passive case, where the signals $a_i(t)$ are externally generated, the processor can be used to obtain information about the sources (number of sources, locations, spectral characteristics, etc.). In the active case, where the $a_i(t)$ are known to the observer, the processor can be used to learn about the transmission process or the array itself (presence of a scatterer, array calibration, etc.). From a mathematical point of view, the difference between the two cases amounts to modeling the source signals as random processes with zero mean in the passive case and as deterministic functions in the active case. In this study, we shall focus our attention on the passive case.

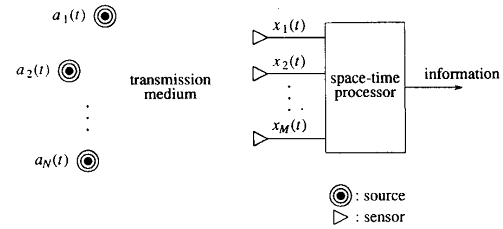


Fig. 1. Generic source-array configuration.

The source signal vector $a(t) = [a_1(t), \cdots, a_N(t)]^T$, where the superscript ^{T} denotes transposition, is generally characterized by a vector θ_a of possibly unknown parameters which can be either random or deterministic in nature. Examples of such parameters might include the bandwidth or the center frequency of a given signal $a_i(t)$. For a fixed value of θ_a, it is assumed that $a(t)$ is a zero-mean Gaussian vector random process with known autocorrelation matrix $R_a(t, u; \theta_a) = E_{\theta_a}[a(t)a^T(u)]$, where $E_{\theta_a}[\ ]$ is to be interpreted as a conditional expectation when θ_a is random and as a conventional expectation based on the signal model with parameter vector θ_a in the deterministic case. (This will only affect the interpretation made of the log-likelihood function.)

The transmission of the signals $a_i(t)$ from the sources to the sensors is represented by a linear transformation satisfying certain general conditions given in Section II-C. This transformation maps the N-vector process $a(t)$ into an M-vector process $s(t)$, consisting of the signal components in the various sensor outputs, according to

$$s(t) = \int_I L(t, u; \theta_l) a(u)\, du, \qquad t \in J \tag{1}$$

where $L(t, u; \theta_l)$ is the $M \times N$ matrix impulse response (possibly a distribution) of the transformation. $L(t, u; \theta_l)$ is completely determined by the geometry of the problem and the nature of the transmission process; it is functionally independent of the process $a(t)$ and the parameter vector θ_a. θ_l is a vector of possibly unknown parameters, either random or deterministic, which characterize the transmission process. Possible examples include the bearing or bearing rate of a given source, or the transmission loss in a given frequency band.

The vector $x(t) = [x_1(t), \cdots, x_M(t)]^T$ of observed sensor outputs is given by

$$x(t) = s(t) + n(t), \qquad t \in J \tag{2}$$

where $s(t)$ is given by (1) and $n(t)$ is an additive noise component modeled as a zero-mean Gaussian vector random process, statistically independent of $a(t)$. In order to simplify the analysis, we assume that $n(t)$ is a white noise process with autocorrelation function

$$R_n(t, u) = E[n(t)n^T(u)] = \delta(t - u)I_{M\times M} \quad (3)$$

where $\delta(t - u)$ is the Dirac delta function and $I_{M\times M}$ is the $M \times M$ identity matrix. The modifications needed for the case of known colored noise are straightforward and described in Section III.

As a first consequence of the linear transformation ir (1), it follows that for a fixed value of the concatenated parameter vector $\theta = (\theta_a, \theta_l)$, $s(t)$ is a zero-mean Gaussian random process and its autocorrelation matrix $R_s(t, u; \theta) = E_\theta[s(t)s^T(u)]$ satisfies

$$R_s(t, u; \theta) = \int_I \int_I L(t, \tau; \theta_l)R_a(\tau, \mu; \theta_a) \cdot L^T(u, \mu; \theta_l)\, d\tau\, d\mu. \quad (4)$$

In this paper, we shall exploit the specific structure of $R_s(t, u; \theta)$ provided by (4) in order to decompose the LLP into simpler subcomponents. To simplify the notation, the dependence in θ_a, θ_l, and θ will generally be omitted unless explicitly needed.

B. Log-Likelihood Processor

We now review the basic equations that define the LLP for the observation model (2), (3). At this point, we do not make use of (4). Since standard derivations of all the results presented below can be found in the literature [11], [13, ch. 2], no proofs are given.

Regardless of whether the problem considered is one of detection or estimation, the LLP always evaluates the log-likelihood function (LLF), $\ln \Lambda(x)$, of the observed data $x(t)$, $t \in J$. In a typical detection problem, the LLF would be computed for each of the possible hypotheses, while in a typical estimation problem, the LLF would be computed for all possible values of the unknown parameters (at least in principle). For the observation model (2), (3), which falls in the "Gaussian signal in Gaussian noise" category, the LFF (or conditional LLF if θ is random) is given by

$$\ln \Lambda(x) = \tfrac{1}{2}\left\{l_1(x) - l_2\right\} \quad (5)$$

$$l_1(x) = \sum_{i=1}^{\infty} \frac{\lambda_i}{\lambda_i + 1}\left\{\int_J \phi_i^T(t)x(t)\, dt\right\}^2 \quad (6)$$

$$l_2 = \sum_{i=1}^{\infty} \ln(1 + \lambda_i). \quad (7)$$

The λ_i and $\phi_i(t)$, which also occur in the Karhunen–Loève expansion [1] of the process $x(t)$, are the eigenvalues and normalized eigenfunctions of $R_s(t, u)$, respectively. They satisfy the following equations:

$$\int_J R_s(t, u)\phi_i(u)\, du = \lambda_i\phi_i(t), \qquad t \in J \quad (8)$$

$$\int_J \phi_i^T(t)\phi_j(t)\, dt = \delta_{ij} \quad (9)$$

where δ_{ij} is the Kronecker delta. It is important to note that the λ_i and $\phi_i(t)$ are functions of the parameter vectors θ_a and θ_l.

Closed-form expressions that do not require the determination of the eigenvalues and eigenfunctions of $R_s(t, u)$ can be obtained for the terms $l_1(x)$ and l_2 of the LLF (5). First, consider $l_1(x)$ (6). Introducing

$$H_2(t, u) = \sum_{i=1}^{\infty} \frac{\lambda_i}{\lambda_i + 1}\phi_i(t)\phi_i^T(u) \quad (10)$$

(the reason for the subscript 2 will become clear later), (6) can be written as

$$l_1(x) = \int_J \int_J x^T(t)H_2(t, u)x(u)\, dt\, du. \quad (11)$$

Although $H_2(t, u)$ is still linked to the λ_i and $\phi_i(t)$ by (10), it can be obtained independently as the unique solution to the $M \times M$ nonhomogeneous Fredholm integral equation of the second kind

$$H_2(t, u) + \int_J H_2(t, v)R_s(v, u)\, dv = R_s(t, u) \qquad t, u \in J. \quad (12)$$

The solution $H_2(t, u)$ of (12), known as the Fredholm resolvent of $R_s(t, u)$, is the impulse response of the noncausal linear filter which provides the minimum mean-squared error (MMSE) estimate of $s(t)$ from $x(u)$, $t, u \in J$ [14].

Next, consider l_2 (7). It can be shown that

$$l_2 = \int_J \mathrm{Tr}\left\{h_2(t, t)\right\} dt \quad (13)$$

where Tr { } is the matrix trace operator, and $h_2(t, u)$ is defined as follows: $h_2(t, u) = 0$ (the $M \times M$ zero matrix) for $T_i \le t < u \le T_f$ and

$$h_2(t, u) + \int_{T_i}^{t} h_2(t, v)R_s(v, u)\, dv = R_s(t, u) \quad (14)$$

for $T_i \le u \le t \le T_f$. The solution $h_2(t, u)$ of (14) has the following interpretation: it is the impulse response of the causal linear filter which provides the MMSE estimate of $s(t)$ from $x(u)$, $T_i \le u \le t \le T_f$ [14].

C. Operator Notation

An appropriate framework for the application of linear operator theory is now established. Moreover, the general conditions imposed on the linear transformation in (1) are stated.

We represent the source signal $a(t)$ as a point in an abstract Hilbert space S_1 of real, N-vector valued functions defined over the interval $I = [t_i, t_f]$, with scalar product and norm, respectively, given by

$$(\alpha, \beta)_1 = \int_I \alpha^T(t)\beta(t)\, dt \quad (15)$$

$$\|\alpha\|_1 = \sqrt{(\alpha, \alpha)_1} \quad (16)$$

for all $\alpha, \beta \in S_1$. By analogy, we represent the array output vector $x(t)$ as a point in a Hilbert space S_2 of real, M-vector valued functions defined over the interval $J = [T_i, T_f]$, with scalar product and norm

$$(\xi, \eta)_2 = \int_J \xi^T(t)\eta(t)\, dt \quad (17)$$

$$\|\xi\|_2 = \sqrt{(\xi, \xi)_2} \quad (18)$$

for all $\xi, \eta \in S_2$. We shall refer to S_1 as the source signal space and to S_2 as the observation space.

Within this framework, we represent the linear transformation in (1) as a linear operator $L: S_1 \rightarrow S_2$ defined by

$$[L\alpha](t) = \int_I L(t, u; \theta_l)\alpha(u)\, du, \qquad t \in J \quad (19)$$

for all $\alpha \in S_1$. In this paper, we shall assume that L is a bounded linear operator with closed range. Recall that L is bounded if there exists a finite positive number m such that

$$\|L\alpha\|_2 \leq m\|\alpha\|_1 \quad (20)$$

for all $\alpha \in S_1$. Physically, this property means that the ratio of the energy contained in $L\alpha$ to the energy contained in α remains bounded for all signals α. The requirement that L have a closed range is more subtle and will be discussed in Section III.

We denote by L^* the adjoint of L. Recall that L^* is a linear operator of $S_2 \rightarrow S_1$ uniquely defined by the condition

$$(L^*\xi, \alpha)_1 = (\xi, L\alpha)_2 \quad (21)$$

required to hold for all $\alpha \in S_1$ and all $\xi \in S_2$. In terms of the impulse response $L(t, u; \theta_l)$, the adjoint operator is given by

$$[L^*\xi](t) = \int_J L^T(u, t; \theta_l)\xi(u)\, du, \qquad t \in I. \quad (22)$$

Proceeding as in (19), we associate the kernels $R_a(t, u)$, $R_s(t, u)$, and $H_2(t, u)$ (12) with operators R_a: $S_1 \rightarrow S_1$, R_s: $S_2 \rightarrow S_2$ and H_2: $S_2 \rightarrow S_2$, respectively. Because $R_a(t, u)$ is a covariance kernel, we have $R_a(t, u) = R_a^T(u, t)$, which implies that R_a is self-adjoint, i.e., $R_a^* = R_a$. The same conclusion applies to R_s; that H_2 is also self-adjoint follows from (12).

With the above notations, some of the previous equations can be written in a simpler way. In particular (denoting by a prime the new form of the equation), we have

$$R_s = LR_aL^* \quad (4')$$

$$R_s\phi_i = \lambda_i\phi_i \quad (8')$$

$$(\phi_i, \phi_j)_2 = \delta_{ij} \quad (9')$$

$$l_1(x) = (x, H_2x)_2 \quad (11')$$

$$H_2(I_2 + R_s) = R_s \quad (12')$$

where, in (12′), I_2 denotes the identity operator in S_2.

III. Factorization Properties of the LLF

In this section, we exploit the specific structure of R_s (4′) and the fact that L is a bounded linear operator with closed range in order to factor the eigenfunctions ϕ_i (8′), (9′) and the Fredholm resolvent H_2 (12′) into simpler components. New expressions for the bias term l_2 (7) of the LLF and other related results are also obtained. The analysis is divided in three parts. In subsection A, we proceed with the construction of a unitary transformation between two fundamental subspaces of S_1 and S_2, and we also provide some motivations for the supplementary assumption of closed range imposed on L. In subsection B, we present and discuss the main properties satisfied by ϕ_i, H_2, and l_2. Finally, in subsection C, we address briefly the case of known colored noise.

A. Unitary Transformation Between the Transmission and Reception Signal Subspaces

Let $\boldsymbol{N}$ and $\boldsymbol{R}$ denote the null space and range space of L, respectively (see Fig. 2(a)), and suppose for the moment that both $\boldsymbol{N}$ and $\boldsymbol{R}$ are closed subspaces, so that the decompositions $S_1 = \boldsymbol{N} \oplus \boldsymbol{N}^\perp$ and $S_2 = \boldsymbol{R} \oplus \boldsymbol{R}^\perp$ where $\oplus$ is a direct sum and $^\perp$ indicates the orthogonal complement, are legitimate [15, p. 30]. According to (1) and (2), any signal $x \in S_2$ at the array output can be written in the from $x = La + n$, where $a \in S_1$ is a source signal component and $n \in S_2$ is a noise component. Moreover, due to the above decompositions of S_1 and S_2, we have $a = a_0 + a_1$ for some $a_0 \in \boldsymbol{N}$ and $a_1 \in \boldsymbol{N}^\perp$, and $n = n_0 + n_1$ for some $n_0 \in \boldsymbol{R}$ and $n_1 \in \boldsymbol{R}^\perp$. Hence,

$$x = (La_1 + n_0) + n_1 \quad (23)$$

where $La_1 + n_0 \in \boldsymbol{R}$. In light of (23), we refer to $\boldsymbol{N}^\perp$ as the transmission signal subspace, to $\boldsymbol{R}$ as the reception signal subspace, and to $\boldsymbol{R}^\perp$ as the reference noise-alone subspace. The terminology adopted for $\boldsymbol{R}$ and $\boldsymbol{R}^\perp$ is consistent with that used by Picinbono [16] in finite dimension.

Two important observations can be made in connection with (23). First, no information about the transmitted signal component a_1 is lost when x is projected on $\boldsymbol{R}$ since the result of this operation is precisely $La_1 + n_0$. Second, note that L defines a one-to-one transformation between $\boldsymbol{N}^\perp$ and $\boldsymbol{R}$ (for any $\alpha, \beta \in \boldsymbol{N}^\perp$, $L\alpha = L\beta$ implies $(\alpha - \beta) \in \boldsymbol{N} \cap \boldsymbol{N}^\perp$, which in turn implies $\alpha = \beta$). Hence, it is possible to map $La_1 + n_0$ back into $\boldsymbol{N}^\perp$, again without any loss of information. These observations suggest that it may be possible to realize the optimum space-time processor in three steps, namely (see Fig. 3): 1) orthogonal

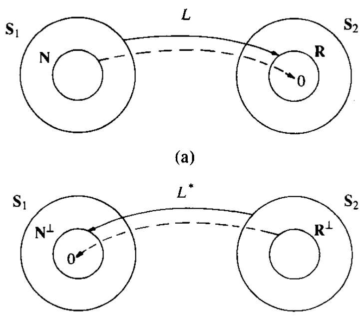


Fig. 2. (a) The operator L and (b) its adjoint L^*.

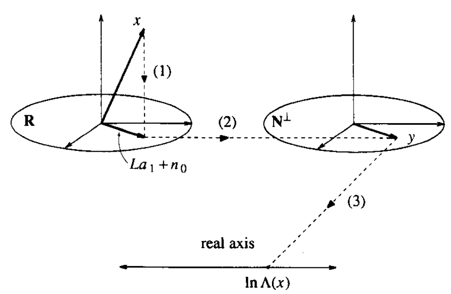


Fig. 3. Three-step approach to optimum space-time processing: (1) orthogonal projection of x onto $\boldsymbol{R}$; (2) one-to-one mapping from $\boldsymbol{R}$ to $N^\perp$; (3) optimum processing of resulting image y.

projection of x onto $\boldsymbol{R}$; (2) mapping of the projection $La_1 + n_0$ back into $N^\perp$ via an appropriate one-to-one transformation; and (3) optimum processing of the resulting image in $N^\perp$. In order to carry out this sequence of operations, we need a proper transformation between $N^\perp$ and $\boldsymbol{R}$. For reasons that will become apparent in the following subsection B, we can advantageously use a unitary transformation. Before proceeding with the construction of this transformation, we digress briefly to comment on the initial assumption that both $\boldsymbol{N}$ and $\boldsymbol{R}$ are closed.

We immediately note that the null space of a bounded linear operator is always closed [17, p. 12] so that we can restrict our attention to $\boldsymbol{R}$. Recall that $\boldsymbol{R}$ is closed if every convergent sequence ξ_n in $\boldsymbol{R}$ converges to a point of $\boldsymbol{R}$. To understand how fundamental is the concept of a closed subspace (in contrast to that of an arbitrary subspace), we point out that geometric notions such as the orthogonal projection of S_2 onto $\boldsymbol{R}$ or the existence of a point in $\boldsymbol{R}$ closest to an arbitrary point in S_2 only make sense when $\boldsymbol{R}$ is closed. Moreover, the assumption that $\boldsymbol{R}$ is closed in the preceeding discussion ensures that the inverse of L, from $\boldsymbol{R}$ to $N^\perp$, is well conditioned (i.e., bounded). Therefore, this assumption can be regarded as a "regularity" condition that simplifies the analysis and the interpretation of the space-time LLP.

A practical way of verifying if L has a closed range is by means of the following property: a bounded linear operator L has closed range if and only if there exists a number $c > 0$ such that

$$\|L\alpha\|_2 \geq c\|\alpha\|_1, \quad \text{all } \alpha \in N^\perp \tag{24}$$

(necessity is proved in [17, p. 14]; sufficiency is proved in [18]). We emphasize that the class of bounded linear operators satisfying the condition (24) is sufficiently large for practical applications and includes most of the "idealistic" transmission models commonly used in the array processing literature.

We now return to the construction of a unitary transformation between the transmission signal subspace $N^\perp$ and the reception signal subspace $\boldsymbol{R}$. To begin, we note that the null space of L^* is $\boldsymbol{R}^\perp$ and (since $\boldsymbol{R}$ is closed) its range is $N^\perp$ [17, p. 13] (see Fig. 2(b)). This property of the adjoint is fundamentally important to the present discussion. Next, we define the operator W: $N^\perp \to N^\perp$ by

$$W = L^*L. \tag{25}$$

Observe that W is self-adjoint. Moreover, it follows from (25) and (24) that for all $\alpha \in N^\perp$

$$(\alpha, W\alpha)_1 = \|L\alpha\|_2^2 \geq c^2\|\alpha\|_1^2 \tag{26}$$

where $c > 0$. Hence, W is strictly positive definite. As a result, W possesses a well-defined square root, i.e., an operator $W^{1/2}$ such that $W^{1/2}W^{1/2} = W$, and both W and $W^{1/2}$ are invertible, with inverses denoted by W^{-1} and $W^{-1/2}$, respectively. Like W, the operators $W^{1/2}$, W^{-1}, and $W^{-1/2}$ are self-adjoint. It is convenient to extend the domain of the operators W^μ ($\mu = \pm\frac{1}{2}, \pm 1$) to S_1 by setting $W^\mu\alpha = 0$ for $\alpha \in N$.

Finally, consider the operator U: $S_1 \to S_2$ defined by

$$U = LW^{-1/2}. \tag{27}$$

If $\alpha \in N^\perp$, then

$$U^*U\alpha = W^{-1/2}L^*LW^{-1/2}\alpha = W^{-1/2}WW^{-1/2}\alpha = \alpha. \tag{28}$$

If $\xi \in \boldsymbol{R}$, then $\xi = L\alpha$ for some $\alpha \in N^\perp$ and we have

$$UU^*\xi = LW^{-1/2}W^{-1/2}L^*L\alpha = LW^{-1}W\alpha = L\alpha = \xi. \tag{29}$$

Hence U provides a unitary transformation between the subspaces $N^\perp$ and $\boldsymbol{R}$ (i.e., a one-to-one transformation having its own adjoint for inverse). As a consequence of (28), it follows that

$$(U\alpha, U\beta)_2 = (\alpha, \beta)_1, \quad \alpha, \beta \in N^\perp. \tag{30}$$

In other words, U preserves scalar products. As we shall see, U and U^* play a fundamental role in the study of the space-time LLP.

B. Main Results

We mentioned previously that it should be possible to realize the LLP in three steps, namely: projection on $\boldsymbol{R}$, mapping from $\boldsymbol{R}$ to $N^\perp$, and optimum processing in $N^\perp$.

This statement, which is verified in the next section, is a simple consequence of the factorization properties presented below. In essence, these properties assert that the integral equations defining the LLP in S_2 can be transformed into equivalent integral equations in S_1. To understand why such transformations are possible, consider the representation (4′) of R_s, i.e., $R_s = LR_aL^*$. Observe that

$$L = LW^{-1/2}W^{1/2} \tag{31}$$

for if $\alpha \in N^\perp$, $W^{-1/2}W^{1/2}\alpha = \alpha$, while if $\alpha \in N$, $W^{1/2}\alpha = 0$ by definition. Therefore, we can express R_s in the form

$$\begin{aligned} R_s &= (LW^{-1/2}W^{1/2})R_a(LW^{-1/2}W^{1/2})^* \\ &= (LW^{-1/2})(W^{1/2}R_aW^{1/2})(LW^{-1/2})^* \\ &= UKU^* \end{aligned} \tag{32}$$

where U is given by (27) and K: $N^\perp \rightarrow N^\perp$ is defined by

$$K = W^{1/2}R_aW^{1/2}. \tag{33}$$

Equation (32) reminds us of a similarity unitary transformation between two Hermitian matrices of the same dimension. Recall that matrices connected by such transformations have many characteristics in common. In particular, they have the same eigenvalues and their eigenfunctions are related by a unitary transformation. By analogy with the finite dimensional case, we would expect that the operators R_s and K in (32) share the same eigenvalues and have eigenfunctions connected by a unitary transformation. More generally, the LLF associated to R_s in the observation space S_2 should in some sense be equivalent to the LLF associated to K in the source signal space S_1. This is confirmed by Properties 1, 2, and 3 below, whose proofs can be found in Appendix I.

The first property concerns the eigenvalue problem (8′), (9′) for R_s.

Property 1: Let $\{\psi_i\}$ be a complete orthonormal set[1] of eigenfunctions of K (33) in $N^\perp$, with corresponding eigenvalues λ_i. That is

$$K\psi_i = \lambda_i\psi_i \tag{34}$$

$$(\psi_i, \psi_j)_1 = \delta_{ij}. \tag{35}$$

Define

$$\phi_i = U\psi_i. \tag{36}$$

Then, the following is true: a) the functions ϕ_i are normalized eigenfunctions of R_s with eigenvalues λ_i, i.e., they satisfy (8′) and (9′); b) $\{\phi_i\}$ is complete in the reception signal subspace $\boldsymbol{R}$; c) any $\xi \in \boldsymbol{R}^\perp$, the reference noise-alone subspace, is an eigenfunction of R_s with zero eigenvalue.

[1]Such a set will exist if the operator K is compact [15, p. 191]. In the present context, a sufficient condition for K to be compact is that the average energy of the signal component $s(t)$ (1) at the array output be finite, i.e., $E\{\|s\|_2^2\} < \infty$.

According to this property, all the eigenfunctions ϕ_i of R_s having nonzero eigenvalues can be generated through (34)–(36). Since the remaining eigenfunctions of R_s do not explicitly enter the expressions (5)–(7) of the LLF, equations (34), (35) are actually equivalent to (8′), (9′). Equation (36) can also be written in the form $\phi_i = L(W^{-1/2}\psi_i)$, which shows that the eigenfunctions of R_s (with nonzero eigenvalues) have the same "structure" as the signal component $s = La$ at the array output. This is consistent with the fact that s can be written as a linear combination of the ϕ_i by means of the Karhunen–Loève expansion.

Because H_2 is related to the ϕ_i through (10), we expect that the factorization (36) for the ϕ_i translates into a corresponding factorization for H_2. Indeed, we have the following result.

Property 2: The Fredholm resolvent of R_s, H_2 (12′), admits the factorization

$$H_2 = UH_1U^* \tag{37}$$

where H_1: $S_1 \rightarrow S_1$ is a self-adjoint operator given by

$$H_1 = K(I_1 + K)^{-1} \tag{38}$$

and I_1 is the identity operator in S_1.

According to this property, the solution of the M-dimensional integral equation (12) reduces to the solution of an equivalent N-dimensional integral equation which, in operator notation, takes the form $H_1(I_1 + K) = K$. We note that H_1 is precisely the Fredholm resolvent of K. This important feature of the factorization (37), (38) will later be exploited to obtain a simple interpretation of H_1 in terms of MMSE estimation theory.

The factorization of $h_2(t, u)$ defined in (14) follows immediately as a particular case of (37) and (38). To see this, let us first rewrite (37) in integral notation, making the dependence of the various kernels upon T_f explicit:

$$H_2(t, u; T_f) = \int_{t_i}^{t_f} d\tau \int_{t_i}^{t_f} d\mu\, U(t, \tau; T_f) \cdot H_1(\tau, \mu; T_f)U^T(u, \mu; T_f) \tag{39}$$

where $H_1(t, u; T_f)$ and $U(t, u; T_f)$ are the integral kernels corresponding to H_1 and U, respectively. Upon comparison of (14) with (12), we notice that for $t \geq u$

$$h_2(t, u) = H_2(t, u; t). \tag{40}$$

(In other words, when $T_f = t$, the causal and noncausal filters providing the MMSE estimates of $s(t)$ from $x(u)$, $T_i \leq u \leq T_f$, are identical.) Using (39) in (40), we finally obtain

$$h_2(t, u) = \int_{t_i}^{t_f} d\tau \int_{t_i}^{t_f} d\mu\, U(t, \tau; t) \cdot H_1(\tau, \mu; t)U^T(u, \mu; t). \tag{41}$$

While (41) provides a legitimate factorization of $h_2(t, u)$, its substitution into the expression (13) for the bias term l_2 of the LLF does not lead to any clarifying simplifications (because the kernels $U(\tau, \mu; t)$ do not cancel out). A more natural approach to the evaluation of l_2 is based on the next property.

Property 3: The kernel $h_2(t, u)$ in (14) satisfies the identity

$$\int_J \mathrm{Tr}\{h_2(t, t)\}\, dt = \int_I \mathrm{Tr}\{h_1(t, t)\}\, dt \quad (42)$$

where $h_1(t, u) = 0$ (the $N \times N$ zero matrix) for $t_i \le t < u \le t_f$ and

$$h_1(t, u) + \int_{t_i}^{t} h_1(t, v) K(v, u)\, dv = K(t, u) \quad (43)$$

for $t_i \le u \le t \le t_f$.

According to (42), one can use $h_1(t, u)$ instead of $h_2(t, u)$ to calculate l_2 (13). Because of the special way in which h_1 and H_1 are related to K through (43) and (38), it is possible to show that [19, p. 130]

$$H_1 = h_1 + h_1^* - h_1^* h_1. \quad (44)$$

This important relation, also known as the Krein factorization theorem, will be used in Section IV.

We emphasize the complete equivalence existing between equations (8), (9), (12), (14), and (34), (35), (38), (43), respectively. The former set of equations defines the LLF associated with R_s in the observation space S_2, while the latter set defines the LLF associated with K in the source signal space S_1. The two sets are connected by (36), (37), and (42), so that knowledge of one of these LLF specifies the other, and vice versa.

In some cases, the evaluation of W^{-1} and $W^{\pm 1/2}$ may pose a serious difficulty. For this reason, we complete the discussion by providing alternative results that do not actually involve these quantities.

Property 1': Let P denote the orthogonal projector of S_1 onto the transmission signal subspace $N^\perp$, and let

$$(\alpha, \beta)_w \equiv (\alpha, W\beta)_1 \quad (45)$$

for all $\alpha, \beta \in N^\perp$. Observe that $(\,.\,,\,.\,)_w$ defines a scalar product on $N^\perp$. Let $\{\eta_i\}$ be a complete orthonormal set of eigenfunctions of $PR_a W$ in $N^\perp$, with respect to the scalar product (45), and let λ_i denote the corresponding eigenvalues. That is

$$PR_a W \eta_i = \lambda_i \eta_i \quad (46)$$

$$(\eta_i, \eta_j)_w = \delta_{ij}. \quad (47)$$

Define

$$\phi_i = L\eta_i. \quad (48)$$

Then, the conclusions of Property 1 apply without modification.

Property 2': H_2 admits the factorization

$$H_2 = LG_1L^* \quad (49)$$

where G_1: $S_1 \to S_1$ is a self-adjoint operator given by

$$G_1 = R_a(I_1 + WR_a)^{-1}. \quad (50)$$

Property 3': Suppose that $N = \{0\}$ (i.e., that L is injective), then

$$l_2 = 2 \int_I \mathrm{Tr}\{f(t, t)\}\, dt \quad (51)$$

where $f(t, u) = 0$ for $u > t$ and

$$f(t, u) + \int_{t_i}^{t} dt'\, f(t, t') \int_{t_i}^{t} du'\, W(t', u') R_a(u', u) = \int_{t_i}^{t} du'\, W(t, u') R_a(u', u) \quad (52)$$

for $t_i \le u \le t \le t_f$.

The proofs of Properties 1' and 2' are almost identical to those of Properties 1 and 2; for this reason, they have been omitted. The proof of Property 3', which is not as simple as that of Property 3, is outlined in Appendix II. Finally, a factorization formula similar to (44) can be obtained for G_1 in the case $N = \{0\}$. It is given by

$$G_1 = g + g^* - g^*f - f^*g \quad (53)$$

where g is a causal operator defined by $g(t, u) = 0$ for $u > t$ and

$$g(t, u) + \int_{t_i}^{t} dt'\, g(t, t') \int_{t_i}^{t} du'\, W(t', u') R_a(u', u) = R_a(t, u) \quad (54)$$

for $t_i \le u \le t \le t_f$. The proof of (53) is outlined in Appendix III.

C. Factorization in Colored Noise

Instead of being specified by (1)–(3), suppose that the vector process observed at the array output is given by

$$x_c = La + n_c \quad (55)$$

where n_c is a colored noise component with strictly positive autocorrelation operator R_{n_c}. Let

$$Q = R_{n_c}^{-1}. \quad (56)$$

Then, the LLP for the observed signal x_c (55) performs the following operations [1]: first, the whitening operator $Q^{1/2}$ is applied to x_c, resulting in

$$x \equiv Q^{1/2} x_c = s + n \quad (57)$$

$$s \equiv Q^{1/2} La \quad (58)$$

$$n \equiv Q^{1/2} n_c \quad (59)$$

where s is the signal complement of x and n is now a white noise process; second, the LLF for the whitened process x is evaluated. Hence, to apply the results of subsection B to the second step of the procedure, simply *replace L by $Q^{1/2}L$ in all the previous expressions.* Of course, $\boldsymbol{R}$ *has to be redefined as the range of* $Q^{1/2}L$ (the null spaces of L and $Q^{1/2}L$ are the same). We note that such a redefinition of $\boldsymbol{R}$ does not affect the original assumption that it is closed, for if L satisfies (24), so does $Q^{1/2}L$ (with

TABLE I
IMPORTANT DEFINITIONS FOR LLP FACTORIZATION IN COLORED NOISE

Symbol	Definition
a	N-component source process
n_c	M-component sensor-noise process
x_c	M-component sensor-output process ($x_c = La + n_c$)
R_a	autocorrelation operator of a
R_{n_c}	autocorrelation operator of n_c
Q	$R_{n_c}^{-1}$
L	transmission operator
$\mathbf{N}$	null space of L
$\mathbf{N}^{\perp}$	transmission signal subspace
$\mathbf{R}$	reception signal subspace (range of $Q^{1/2}L$)
$\mathbf{R}^{\perp}$	reference noise-alone subspace
x	$Q^{1/2}x_c$
s	$Q^{1/2}La$
n	$Q^{1/2}n_c$
R_s	$Q^{1/2}LR_aL^*Q^{1/2}$
W	L^*QL
U	$Q^{1/2}LW^{-1/2}$
K	$W^{1/2}R_aW^{1/2}$
P	projector onto transmission signal subspace $\mathbf{N}^{\perp}$
ϕ_i	eigenfunctions of R_s ($\phi_i = U\psi_i = Q^{1/2}L\eta_i$)
ψ_i	eigenfunctions of K
η_i	eigenfunctions of PR_aW
H_2	Fredholm resolvent of R_s ($H_2 = UH_1U^* = Q^{1/2}LG_1L^*Q^{1/2}$)
H_1	$K(I_1+K)^{-1}$
G_1	$R_a(I_1+WR_a)^{-1}$
h_1	causal operator related to K through equation (43)
f	causal operator related to R and W through equation (52)
g	causal operator related to R and W through equation (54)

possibly a different constant c). This follows because Q (56) is strictly positive. For future reference, the colored noise versions of the various quantities introduced thus far have been listed in Table I.

IV. Processor Configurations

We now look at the LLP configurations that result from the factorization properties in Section III. For the sake of generality, colored noise is assumed throughout.

A. Canonical Configuration

Using (37), the data dependent term (11′) of the LLF (5) can be written in the form

$$l_1(x) = (y, H_1 y)_1 \tag{60}$$

where

$$y \equiv U^*x = U^*Q^{1/2}La + U^*n \tag{61}$$

(the second equality in (61) follows from (57) and (58)). The resulting canonical LLP configuration is shown in Fig. 4 where, for completeness of presentation, the dependence of the various block components on the parameter vectors θ_a and θ_l have been indicated. This configuration consists of three specialized subprocessors serving very different purposes. The first one is a space-time whitening filter that transforms x_c into $x = Q^{1/2}x_c$ (57). The second one, referred to as a *unitary beamformer* (UB), transforms the M-component signal x into the N-component signal $y = U^*x$ (61). The last one, referred to as a *quadratic postprocessor* (QPP), finally operates on y according to (60) to produce the LLF (5).

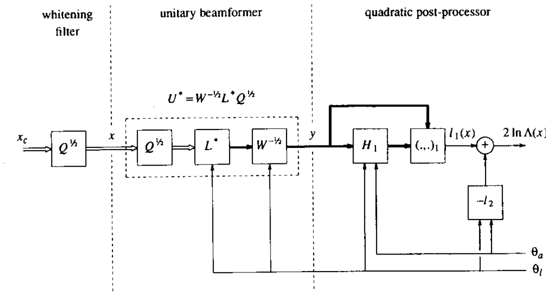


Fig. 4. Canonical space-time LLP configuration (see Table I for description of various operators).

The canonical configuration in Fig. 4 provides a generalization to nonstationary signal models of conventional optimum array processing structures previously derived under the assumptions of (specific) time-invariant transmission channels, stationary signal processes, and long observation intervals [2]–[8]. To recover these structures from Fig. 4, simply replace the various operators by the Fourier transforms of their matrix kernels (assumed to be time invariant); the resulting structure then operates in the frequency domain with abstract operator composition replaced by conventional matrix multiplication. Additional explanations are given in Section V-B, where the optimum structure of [5] is derived as a particular case of a more general time-varying processor.

We note that in Fig. 4, the whitening filter and the UB can actually be reconfigured so that only Q appears at the input of the LLP, and not $Q^{1/2}$. However, we emphasize that Q enters the block components $W^{-1/2}$ and H_1 implicitly because $W = L^*QL$ and $H_1 = K(I_1 + K)^{-1}$ with $K = W^{1/2}R_aW^{1/2}$. Even the bias term l_2 depends on Q because it is a function of the eigenvalues of K. The point is that conventional noise prewhitening affects all stages of the LLP, not only the input. In subsections B and C, we take a closer look at the UB and the QPP.

B. The Unitary Beamformer (UB)

By definition, the UB performs the operation $U^* = W^{-1/2}L^*Q^{1/2}$ on the whitened signal x. Since the null and range spaces of $U = Q^{1/2}LW^{-1/2}$ are $\boldsymbol{N}$ and $\boldsymbol{R}$, respectively, the null space of U^* is $\boldsymbol{R}^{\perp}$ and its range is $\boldsymbol{N}^{\perp}$. We can therefore visualize the operation U^*x in two steps: x is first projected on the reception signal subspace $\boldsymbol{R}$ and then, the resulting projection is mapped into the transmission signal subspace $\boldsymbol{N}^{\perp}$ for further processing by the QPP. This validates the general approach suggested in Fig. 3 where both steps (1) and (2) can be performed by the UB. Note that in $U^* = W^{-1/2}L^*Q^{1/2}$, $Q^{1/2}$ must not be mistaken for a whitening operation (the operator $Q^{1/2}$ at the input of the LLP already acts as a whitening filter). One must rather interpret the combination $L^*Q^{1/2} = (Q^{1/2}L)^*$ as a space-time filter matched to the "transmission" operator $Q^{1/2}L$ in (58).

The UB has a number of very desirable properties. To begin with, it does not depend upon the statistics of the source signal $a(t)$; it is entirely determined by the source-array geometry, the transmission characteristics of the medium and the noise statistics. This property is particularly important for applications in which a complete description of R_a is unavailable, or in estimation problems where $\ln \Lambda(x)$ has to be computed for different values of θ_a. In fact, *for any fixed value of* θ_l, *the output* y *of the UB provides a sufficient statistic for both problems of signal detection and estimation of* θ_a. Another important feature of the UB is that the number of signal components at its output is equal to the number of sources N in the model (which needs not be the same as the exact number of sources really present in the environment (see subsection F)). *Whenever* $N < M$, *the UB can therefore be used to reduce the number of signals that needs to be quadratically processed.*

If n is an S_2 process with autocorrelation operator I_2, then U^*n is an $N^\perp$ process with autocorrelation operator U^*U. Since U^*U acts as the identity in $N^\perp$ (see (28)), we conclude that *the UB is a distortionless processor, in the sense that it maps a white noise process in* S_2 *into a white noise process in* $N^\perp$.

In the class of all operators Ω: $S_2 \rightarrow N^\perp$ satisfying the distortionless condition $\Omega\Omega^*\alpha = \alpha$ for all $\alpha \in N^\perp$, the operator U^* plays a very important role. Indeed, let $y = \Omega x$, where $x = s + n$ (57) is the input to the UB in Fig. 4. Since both x and y consist of a signal component corrupted by additive unit white noise, it is legitimate to define the input and output channel signal-to-noise ratios for the operator Ω as follows:

$$\text{SNR}_{\text{in}} = \frac{1}{M} E[\text{total energy in s}] \tag{62}$$

$$\text{SNR}_{\text{out}} = \frac{1}{N} E[\text{total energy in } \Omega s] \tag{63}$$

where s is given by (58). Then, it can be shown (see Appendix IV) that the choice $\Omega = U^*$ maximizes the ratio $\text{SNR}_{\text{out}}/\text{SNR}_{\text{in}}$, with

$$\left.\frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}}\right|_{\Omega = U^*} = \frac{M}{N}. \tag{64}$$

In other words, *the unitary beamformer maximizes the array gain.*

C. The Quadratic Postprocessor

The following expressions for the data-dependent term (11′) of the LLF can be obtained easily

$$l_1(x) = (H_1^{1/2}y, H_1^{1/2}y)_1 \tag{65}$$

$$l_1(x) = (2y - h_1 y, h_1 y)_1 \tag{66}$$

$$l_1(x) = \sum_i \frac{\lambda_i}{\lambda_i + 1} \{(\psi_i, y)_1\}^2. \tag{67}$$

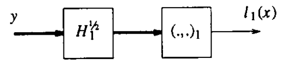


Fig. 5. Filter-squarer postprocessor.

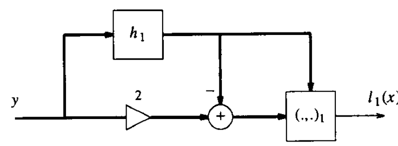


Fig. 6. Causal postprocessor.

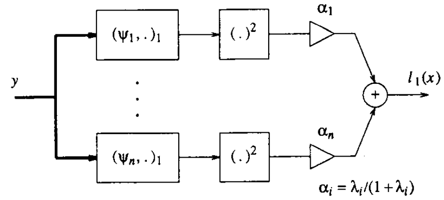


Fig. 7. Eigenvector postprocessor (n-term approximation).

Equation (65) is a simple consequence of (60), (66) follows from (60) and (44), and (67) follows from (6), (36), and (61). The postprocessor configurations corresponding to (65), (66), and (67) are shown in Figs. 5, 6, and 7, respectively. These configurations are quite standard [13, ch. 2] and the terminology adopted should require no explanation. The novelty here lies in that these configurations are not applied directly to the M-component array output vector x_c, but rather to the N-component beamformer output y.

In the present context, simple interpretations can be given to the filters H_1 and h_1 in Figs. 4–6 and to the eigenvalues λ_i and eigenfunctions ψ_i in Fig. 7. To begin with, observe that in the colored noise case, (31) becomes $Q^{1/2}LW^{-1/2}W^{1/2} = Q^{1/2}L$ (with $W = L^*QL$). Hence

$$\begin{aligned} U^*Q^{1/2}La &= U^*Q^{1/2}LW^{-1/2}W^{1/2}a \\ &= U^*UW^{1/2}a \\ &= W^{1/2}a. \end{aligned} \tag{68}$$

Making this substitution into (61) and defining

$$z = W^{1/2}a \tag{69}$$

we obtain

$$y = z + U^*n. \tag{70}$$

Now observe that $U^*U = P$, the orthogonal projector of S_1 onto $N^\perp$ introduced in Property 1′, for if $\alpha \in N^\perp$, $U^*U\alpha = P\alpha = \alpha$, and if $\alpha \in N$, $U\alpha = 0$ and $P\alpha = 0$. Using this result and the definition (33) of K, and observing that a and n are independent, we easily obtain that

$$E[z(t)z^T(u)] = K(t, u) \tag{71}$$

$$E[z(t)y^T(u)] = K(t, u) \tag{72}$$

$$E[y(t)y^T(u)] = K(t, u) + P(t, u) \tag{73}$$

where $K(t, u)$ and $P(t, u)$ are the integral kernels associated with the operators K and P. Using (72), (73), and (38), it can be verified that the orthogonality condition $E\{(z(t) - [H_1 y](t))y^T(u)\} = 0$ is satisfied for all t, $u \in I$. Therefore, $H_1 y$ is a noncausal linear MMSE estimate of z from y. When $P = I_1$ (i.e., when $N = \{0\}$), it can be verified in a similar way that $h_1 y$ is a causal linear MMSE estimate of z from y (however, when $P \neq I_1$, $h_1 y$ does not admit a simple interpretation). Finally, it follows from (71), (34), and (35) that the λ_i and ψ_i are the eigenvalues and eigenfunctions associated to the process z. These observations can be summarized by saying that *the quadratic postprocessor is a LLP for the output y of the UB in Fig. 4.*

Some additional comments can be made regarding the configuration of Fig. 6. First, even if this configuration is causal, its insertion into Fig. 4 does not result in a "globally" causal LLP because the whitening filter and the UB are in general noncausal, with processing delays of the order of the correlation times of the kernels $L(t, u)$ (1) and $Q(t, u)$ (56). Second, whenever the kernel $K(t, u)$ is state-space representable (or more generally separable), a direct realization of h_1 in terms of initial value differential equations is possible [14].

D. Alternative LLP Configurations

The LLP configurations in Figs. 4–7 were based on the factorization Properties 1–3 and on (44). Alternative configurations based on Properties 1′, 2′, and 3′ and on (53) can be derived in an obvious way. Here, we briefly discuss one of these alternative configurations because of its connection to Schweppe's work [9].

Using (11′) and (49) and making the appropriate colored noise modifications, the LLP configuration shown in Fig. 8 is easily derived. This configuration has a very interesting property which follows from the orthogonality principle, namely: the output of the filter G_1 is a noncausal linear MMSE estimate of Pa from x_c. The particular decomposition of G_1 in Fig. 8 provides some insights into how such an estimate is obtained. Indeed, observe that the output of the operator W^{-1} is given by $Pa + W^{-1}LQn_c$. Hence, in the absence of noise, Pa can be recovered exactly at the output of W^{-1}. This is because $W^{-1}L^*Q^{1/2}$ is the generalized inverse of $Q^{1/2}L$ in (58) [17, p. 45]. Hence, MMSE estimation of Pa from x_c can be visualized in three steps: prewhitening, generalized inversion, and optimum filtering with $(I_1 + R_a W)^{-1} R_a W$. The configuration of Fig. 8 generalizes the so-called "decoupled-beam data processor" obtained by Schweppe [9] in the study of least squares array processors for pure time delay propagation models. In particular, the operator W^{-1} extends the concept of a decoupling matrix introduced in [9].

E. Postprocessor Approximations

A low signal-to-noise ratio (SNR) situation occurs whenever $\|K\| < 1$, where $\|K\|$, the norm of K, is defined by $\|K\| = \sup\{\|K\alpha\|_1 : \alpha \in S_1, \|\alpha\|_1 = 1\}$. In

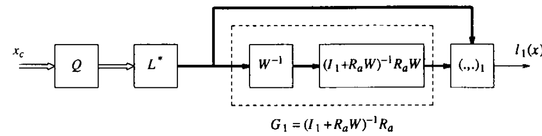


Fig. 8. Generalized decoupled-beam processor.

such a case, the Neumann series [15, p. 86] $(I_1 + K)^{-1} = I_1 - K + K^2 - \cdots$ converges, and its substitution into (38) yields

$$H_1 = K - K^2 + \cdots. \tag{74}$$

Retaining only the first term in the above series and substituting the resulting approximation $H_1 \approx K$ in Fig. 4, we obtain (after reconfiguration) the LLP shown in Fig. 9. Except for $Q = R_{n_c}^{-1}$, no other operator inversion is required with this configuration. When the source signals $a_i(t)$ are uncorrelated, the processor R_a in Fig. 9 decouples into N parallel processors, one for each output of the operator L^*. This property of the LLP was observed by Wax and Kailath [7] in the study of the maximum likelihood estimator of source locations in a stationary environment.

At the other extreme, a high SNR situation occurs whenever K is invertible (on $N^\perp$) and $\|K^{-1}\| < 1$. In this case, we find that

$$H_1 = I_1 - K^{-1} + \cdots. \tag{75}$$

Making the substitution $H_1 \approx I_1$ in the quadratic postprocessor of Fig. 4 reveals that at high SNR, the LLP simply computes the energy contained in the N outputs of the UB (another interesting feature of the UB). An equivalent space-time LLP configuration for high SNR is shown in Fig. 10. This configuration is robust in the sense that it is independent of the statistics of the source signal $a(t)$. Finally, it is interesting to note the similarity existing between the low and high SNR LLP configurations in Figs. 9 and 10, and the generalized decoupled-beam processor in Fig. 8.

F. Practical Considerations

Thus far, we have been concerned with the study of the properties and structure of the space-time LLP for a fixed source-array configuration. We have assumed that the number of sources N, the channel impulse response $L(t, u; \theta_l)$, and the signal and noise autocorrelation functions $R_a(t, u; \theta_a)$ and $R_n(t, u)$ were known. In practical applications, however, such detailed *a priori* knowledge may not be available and this ultimately affects the way in which the LLP is realized.

At first, the most fundamental difficulty seems to come from the requirement that the number of sources N be known in the signal model of Section II-A. However, the difficulty is only apparent. For instance, consider the problem of localizing two point sources in the array far field. In this case, a space-time LLP based on a single source model can be used with near optimal performances

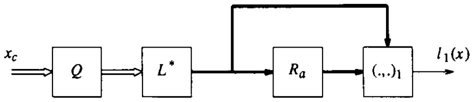


Fig. 9. Low SNR space-time LLP configuration ($\|K\| \ll 1$).

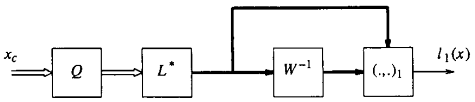


Fig. 10. High SNR space-time LLP configuration ($\|K^{-1}\| \ll 1$).

as long as the angular separation between the sources is sufficiently large. Although performance will deteriorate if the angular separation decreases below a certain level, it can be partially corrected by using a LLP based on $N = 2$ instead of $N = 1$ [8]. Thus the number of sources N used in the signal model need not be the same as the actual number of sources present in the environment; N should rather be determined from resolution considerations.

Once the number of sources N has been determined, the choice of $L(t, u; \theta_l)$ follows from a study of the various sound propagation mechanisms taking place. In the absence of more extended knowledge, one can simply assume pure time delay propagation. However, if the transmission process is known to involve more complicated effects such as dispersion or multipath, these can be included in L.

A lack of *a priori* knowledge about the signal and noise statistics is common in applications of array processing and most techniques developed to overcome this difficulty in stationary environments can in principle be extended to more general nonstationary signal models. For example, consider the configuration in Fig. 8 and suppose that detailed knowledge of R_a is unavailable. Since the output of G_1 is an estimate of a (assume $P = I_1$ for simplicity), this output can be used to form an estimate of R_a, say $\hat{R}_a$. Then, upon replacement of R_a by $\hat{R}_a$ in Fig. 8, we obtain an "approximate" implementation of G_1.

V. Examples

In order to illustrate the theory presented in this paper, three examples of its application to signal models involving time-varying delays are now presented.

A. Stuller's Results

Stuller [11] studied the problem of maximum likelihood estimation of time-varying delay for a single source signal monitored in the presence of white noise at two different locations. Using the principle of reversibility [1, p. 289], he was able to obtain closed form expressions for the elements of the 2×2 matrix kernel $H_2(t, u)$ in (12), in terms of the (unknown) solution of a scalar integral equation. In this example, we present an alternative derivation of this result based on the factorization properties of Section III. Besides being considerably simpler, our approach provides additional insights.

The signal model considered in [11] falls into the general description of Section II-A, with the following specifications:

a) $N = 1, \quad M = 2.$

b) $$[La](t) = \begin{bmatrix} a(t) \\ ca(t - d(t)) \end{bmatrix}. \tag{76}$$

c) $$R_n(t, u) = N_0 \delta(t - u) I_{2\times 2}. \tag{77}$$

In the above expressions, $c \neq 0$ is an attenuation coefficient, $d(t)$ is a time-varying delay function, and N_0 is the noise power level in the sensor outputs. We assume that $0 \le d(t) \le T_f - T_i$ and $d'(t) < 1$, where the prime denotes derivative with respect to time. The last condition ensures that the function $f(t) = t - d(t)$ is strictly increasing and therefore invertible. The corresponding inverse, i.e., $f^{-1}(t)$, is denoted by $\beta(t)$.

We begin by constructing the mathematical quantities needed to apply the factorization properties. The proper choice for the interval of integration $I = [t_i, t_f]$ is given by $t_i = f(T_i)$ and $t_f = T_f$. With this choice, the null space of L reduces to $N = \{0\}$ and, as a result, the transmission signal subspace $N^\perp$ coincides with the source signal space S_1, i.e., $N^\perp = S_1$. To evaluate L^*, first observe that

$$L(t, u) = \begin{bmatrix} \delta(t - u) \\ c\delta(f(t) - u) \end{bmatrix}. \tag{78}$$

Substituting (78) into (22), we find that

$$[L^*\xi](t) = \begin{cases} c\beta'(t)\xi_2(\beta(t)), & t \in A_1 \\ \xi_1(t) + c\beta'(t)\xi_2(\beta(t)), & t \in A_2 \\ \xi_1(t), & t \in A_3 \end{cases} \tag{79}$$

where $\xi = [\xi_1, \xi_2]^T$ is an arbitrary element in the observation space S_2, $A_1 = (f(T_i), T_i)$, $A_2 = (T_i, f(T_f))$ and $A_3 = (f(T_f), T_f)$. Using (76), (77), and (79), it is easy to show that

$$[W^\mu \alpha](t) = \{N_0^{-1}\rho(t)\}^\mu \alpha(t), \qquad t \in I \tag{80}$$

where $\mu \in \{\pm\frac{1}{2}, \pm 1\}$, $\alpha \in S_1$ and

$$\rho(t) \equiv \begin{cases} c^2\beta'(t), & t \in A_1 \\ 1 + c^2\beta'(t), & t \in A_2 \\ 1, & t \in A_3. \end{cases} \tag{81}$$

We note from (79) and (81) that $[L^*\xi](t)$ and $\rho(t)$ are discontinuous at $t = T_i$ and $t = f(T_f)$. The discontinuities, or *edge effects*, result from the fact that the observation interval is finite.

The operators U and U^* are given by

$$[U\alpha](t) = \begin{bmatrix} \{\rho(t)\}^{-1/2}\alpha(t) \\ c\{\rho(f(t))\}^{-1/2}\alpha(f(t)) \end{bmatrix} \tag{82}$$

$$[U^*\xi](t) = \{\rho(t)\}^{-1/2}[L^*\xi](t). \tag{83}$$

The interpretation of U^* as a beamforming operation is consistent with (83) and (79). Indeed, choosing temporarily $c = 1$ and $d(t) = d_0$ (constant delay), we find that over the interval A_2, which is the dominant interval when $T_f - T_i >> d(t)$, $[U^*\xi](t) = \{\xi_1(t) + \xi_2(t + d_0)\}/\sqrt{2}$. This result corresponds to a simple delay and sum beamforming operation on $\xi(t)$. In the general case, the operation $\xi_2(t) \to \beta'(t)\xi_2(\beta(t))$ in (79) can be thought of as a compensation scheme that counteracts the effects of the delay $d(t)$.

The kernel of the operator K is given by

$$K(t, u) = N_0^{-1}\{\rho(t)\}^{1/2}R_a(t, u)\{\rho(u)\}^{1/2}. \quad (84)$$

It equals the autocorrelation (7') of the process $z(t) = [W^{1/2}a](t) = \{N_0^{-1}\rho(t)\}^{1/2}a(t)$. If the process $a(t)$ can be described by a linear state equation, so is $z(t)$ and therefore, a standard Kalman-Bucy filter can be used to implement the postprocessor of Fig. 4. This approach provides an alternative to a recent technique [12] which, for the example under consideration, requires the use of a generalized Kalman-Bucy filter applied directly to the sensor outputs. Using postbeamforming Kalman-Bucy filtering rather than generalized Kalman-Bucy filtering of the sensor outputs appears to be particularly attractive in the case of $M >> 1$.

We now proceed with the factorization of the eigenfunctions ϕ_i and the Fredholm resolvent H_2. Considering the simplicity of the resulting expressions, we prefer to use Properties 1' and 2' rather than Properties 1 and 2. Using (46)-(48) (with L replaced by $N_0^{-1/2}L$), we obtain

$$\phi_i(t) = N_0^{-1/2}\begin{bmatrix}\eta_i(t) \\ c\eta_i(t - d(t))\end{bmatrix}, \qquad t \in J \quad (85)$$

$$N_0^{-1}\int_I R_a(t, u)\eta_i(u)\rho(u)\,du = \lambda_i\eta_i(t), \qquad t \in I \quad (86)$$

$$N_0^{-1}\int_I \eta_i(t)\eta_j(t)\rho(t)\,dt = \delta_{ij}. \quad (87)$$

Using (49) and (50) (again with L replaced by $N_0^{-1/2}L$), we obtain

$$H_2(t, u) = N_0^{-1}\begin{bmatrix}G_1(t, u) & cG_1(t, f(u)) \\ cG_1(f(t), u) & c^2G_1(f(t), f(u))\end{bmatrix} \qquad t, u \in J \quad (88)$$

$$G_1(t, u) + N_0^{-1}\int_I G_1(t, v)R_a(v, u)\rho(v)\,dv = R_a(t, u), \qquad t, u \in I. \quad (89)$$

It can be verified that (88) and (89) are equivalent to [11, eq. (3.44), (3.49), (3.50), (3.52), and (3.53)]. However, (88) and (89) give more insight into the structure of $H_2(t, u)$ because of their simplicity and their direct relation with the operator L. Specialization of (51), (52), and (54) to this example is trivial and we shall not present the results.

Finally, we note that the possibility of factoring the eigenfunctions $\phi_i(t)$ as in (85)-(87) and the existence of the canonical LLP configuration of Fig. 4 seem to have been overlooked in [11].

B. Generalization of Knapp and Carter's Processor

Knapp and Carter [10] derived the maximum likelihood processor for estimating a linearly varying time delay between two noisy versions of a common source signal. They considered stationary signal and noise processes and assumed that the observation time was large in comparison to the delay and to the signal and noise correlation times. In this example, we extend their results to the case of M sensors, nonlinearly varying time delays and complex, frequency-dependent attenuation.

The signal model under consideration is specified as follows:

a) M arbitrary, $N = 1$.

b) The ith component of the $M \times 1$ impulse response $L(t, u)$ is given by

$$L_i(t, u) = (2\pi)^{-1}\int_{-\infty}^{\infty} C_i(\omega)\, e^{j\omega[t - d_i(t) - u]}\,d\omega \quad (90)$$

where $C_i(\omega) = C_i^*(-\omega)$ is a frequency dependent attenuation coefficient, and $d_i(t)$ is a slowly varying time delay function satisfying the condition $|d_i'(t)| << 1$.

c) The processes $a(t)$ and $n(t)$ are stationary with power spectral densities $A(\omega)$ and $N(\omega)$, respectively.

d) The noise components at different sensors are uncorrelated, that is, $N(\omega) = \text{diag}[N_1(\omega), \cdots, N_M(\omega)]$.

e) $T \equiv T_f - T_i >> d_{\max}, \tau_a, \tau_n, \tau_c$, where $d_{\max}$ is the maximum value of $|d_i(t)|$, τ_a and τ_n are the correlation times of $a(t)$ and $n(t)$, and τ_c is the largest correlation time of the impulse responses associated with the $C_i(\omega)$.

To determine the components of the LLP, we shall require different approximations. They can be justified because the errors introduced are negligible (and generally overshadowed by a lack of *a priori* knowledge). Besides, many of these approximations also occur in the derivations of the conventional "optimum" array processors [4]-[8].

Because $T >> d_{\max}, \tau_c$, we set $I = J = [T_i, T_f]$. Substituting (90) into (22) and defining $\xi(u) \equiv 0$ for $u \notin J$, we find that

$$[L^*\xi](t) = (2\pi)^{-1}\int_{-\infty}^{\infty} d\omega\left\{\sum_{i=1}^{M} C_i^*(\omega)\int_{-\infty}^{\infty} du\, \beta_i'(u)\xi_i(\beta_i(u))e^{-j\omega u}\right\}e^{j\omega t}, \qquad t \in I \quad (91)$$

where $\beta_i(t)$ is the inverse function of $f_i(t) = t - d_i(t)$. The kernel $W(t, u)$ is given by

$$W(t, u) = \int_J d\tau \int_J d\mu\, L^T(\tau, t)Q(\tau, \mu)L(\mu, u) \qquad t, u \in I. \quad (92)$$

Under the assumption e) of long observation time, we have

$$Q(t, u) \approx \text{diag}[Q_1(t-u), \cdots, Q_M(t-u)] \quad (93)$$

$$Q_i(\tau) = (2\pi)^{-1} \int_{-\infty}^{\infty} d\omega \, \{N_i(\omega)\}^{-1} e^{j\omega\tau}. \quad (94)$$

Moreover, we can replace the limits of integration in (92) by $\pm\infty$. After a straightforward calculation, we find

$$W(t, u) \approx (2\pi)^{-2} \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \sum_{i=1}^{M} \cdot C_i^*(\omega_1) C_i(\omega_2) e^{j(\omega_1 t - \omega_2 u)} \cdot \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} d\mu \, \beta_i'(\tau)\beta_i'(\mu) Q_i(\beta_i(\tau) - \beta_i(\mu)) e^{-j(\omega_1\tau - \omega_2\mu)}. \quad (95)$$

Since $\beta_i(t) = t + d_i(\beta_i(t))$ and $|d_i'(t)| << 1$, we have

$$\beta_i(t) \approx t + d_i(t)$$
$$\beta_i'(t) \approx 1$$
$$\beta_i(t) - \beta_i(u) \approx t - u \quad (96)$$

with relative errors of the order of $|d_i'(t)|$. An interesting interpretation of these approximations in terms of source velocity and noise bandwidth can be found in [20, p. 72]. Using (96) in (91) and (95) and making the necessary manipulations, we obtain

$$[L^*\xi](t) \approx (2\pi)^{-1} \int_{-\infty}^{\infty} d\omega \left\{ \sum_{i=1}^{M} C_i^*(\omega) \int_{-\infty}^{\infty} \cdot du \, \xi_i(u + d_i(u)) e^{-j\omega u} \right\} e^{j\omega t} \quad (97)$$

$$W(t, u) \approx (2\pi)^{-1} \int_{-\infty}^{\infty} d\omega \, D(\omega) e^{j\omega(t-u)} \quad (98)$$

$$D(\omega) = \sum_{i=1}^{M} \frac{|C_i(\omega)|^2}{N_i(\omega)}. \quad (99)$$

Observe that $W(t, u)$ is (approximately) time invariant, a remarkable feature of the particular problem considered here. Since $R_a(t, u)$ is also time invariant and the observation interval is long, this implies that $[W^\mu](t, u)(\mu = \pm\frac{1}{2}, -1)$, $K(t, u)$, $H_1(t, u)$ and $G_1(t, u)$ are all time invariant with Fourier transforms given by

$$W^\mu \leftrightarrow \{D(\omega)\}^\mu \quad (100)$$

$$K \leftrightarrow D(\omega)A(\omega) \quad (101)$$

$$H_1 \leftrightarrow \frac{D(\omega)A(\omega)}{1 + D(\omega)A(\omega)} \quad (102)$$

$$G_1 \leftrightarrow \frac{A(\omega)}{1 + D(\omega)A(\omega)}. \quad (103)$$

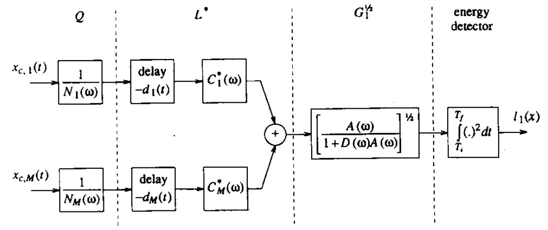


Fig. 11. LLP configuration for example of Section V-B.

The resulting LLP configuration is shown in Fig. 11. When $d_i(t) = d_i$ and $C_i(\omega) = 1$, i.e., constant delays and perfect propagation, this configuration reduces to that of [5].

Finally, a closed-form expression can be obtained for the bias term l_2 (7). Indeed, since $K(t, u)$ is time invariant and the observation interval is long, its eigenvalues λ_i are simply given by the spectrum (101) evaluated at $\omega = \omega_i$, where $\omega_i = 2\pi i/T$, $i = 0, \pm 1, \cdots$ [1, p. 207]. That is

$$\lambda_i \approx D(\omega_i)A(\omega_i), \qquad i = 0, \pm 1, \cdots. \quad (104)$$

Using (104) in (7) and replacing the infinite summation by an integral, we obtain

$$l_2 \approx \frac{T}{2\pi} \int_{-\infty}^{\infty} \ln[1 + D(\omega)A(\omega)] \, d\omega. \quad (105)$$

Note that l_2 does not depend on the delay functions $d_i(t)$.

The LLP derived in this example can be used for maximum likelihood estimation of the delays $d_i(t)$. In this respect, it is important to know how uncertainties in the signal power spectrum $A(\omega)$ will affect the accuracy of the resulting estimates. This problem is addressed in [20], [21] in the case $C_i(\omega) = 1$. There, a lack of *a priori* knowledge concerning spectral parameters θ_a was shown not to affect the minimum variance attainable in estimating delay parameters. To achieve this minimum, however, it may be necessary to maximize the LLF over a larger set of parameters including θ_a.

C. An Example Involving Two Sources

Consider the following signal model:

a) $M \geq 2$, $N = 2$.

b)
$$[La](t) = \begin{bmatrix} a_1(t - d_{11}(t)) + a_2(t - d_{12}(t)) \\ \vdots \\ a_1(t - d_{M1}(t)) + a_2(t - d_{M2}(t)) \end{bmatrix} \quad (106)$$

where $|d_{ij}'(t)| << 1$.

c) The sources are uncorrelated, i.e., $R_a(t, u) = \text{diag}[R_1(t, u), R_2(t, u)]$.

d) The noise component is white, i.e., $R_n(t, u) = \delta(t - u)I_{M \times M}$.

e) $T >> d_{\max}, \tau_a$.

As before, define $f_{ij}(t) = t - d_{ij}(t)$ and denote the corresponding inverse functions by $\beta_{ij}(t)$. Invoking the condition e) and making approximations similar to (96), we obtain

$$[L^*\xi](t) = \begin{bmatrix} \sum_{i=1}^{M} \xi_i(t + d_{i1}(t)) \\ \sum_{i=1}^{M} \xi_i(t + d_{i2}(t)) \end{bmatrix} \quad (107)$$

$$W_{pq}(t, u) = \sum_{i=1}^{M} \delta(f_{iq}(\beta_{ip}(t)) - u) \quad (108)$$

where $W_{pq}(t, u)$ $(p, q = 1, 2)$ are the elements of the 2×2 matrix kernel $W(t, u)$. In this example, the difficulty comes from the off-diagonal terms in $W(t, u)$. To understand the nature of this difficulty, consider the operator WR_a which occurs in the factorization (49), (50) and in the LLP configuration of Fig. 8. After a straightforward computation, the following expression can be obtained for the matrix kernel of this operator:

$$\begin{bmatrix} MR_1(t, u) & \sum_{i=1}^{M} R_2(t + d_{i1}(t) - d_{i2}(t), u) \\ \sum_{i=1}^{M} R_1(t + d_{i2}(t) - d_{i1}(t), u) & MR_2(t, u) \end{bmatrix} \quad (109)$$

Because of the nonzero off-diagonal terms in the above expression, the postbeamforming processor $G_1 = R_a(I_1 + WR_a)^{-1}$ in Fig. 8 will not, in general, decouple into two parallel subprocessors (unless $\| WR_a \| << 1$). Although a complete analysis of this problem is beyond the scope of this paper, at least one particular case of practical interest exists where the off-diagonal elements in (109) can be neglected. Indeed, suppose that the following two conditions are satisfied:

a') $M >> 1$.

f) $|d_{i1}(t) - d_{i2}(t) - d_{j1}(t) + d_{j2}(t)| > 2\tau_a$, $i \neq j$.

Then, for any given values of t and u, only one term in the summations over i in (109) can give a significant contribution. When $M >> 1$, the off-diagonal elements in (109) can therefore be neglected. As a result, the processor G_1 in Fig. 8 decouples into 2 parallel sub-subprocessors and the outputs (or beams) of L^* can be processed independently of each others. This decoupling property of G_1 provides a very fundamental explanation of the well-known fact that sources having sufficient angular separation do not interfere considerably.

VI. Conclusions

In this paper, we have studied the structural properties of the space-time LLP under very general time-varying conditions. In particular, we have shown that the M-dimensional integral equations specifying the LLF, where M is the number of sensors, can be transformed into equivalent N-dimensional integral equations, where N is the number of sources in the model. This transformation made it possible to decompose the LLP into a cascade of three specialized subprocessors (Fig. 4): a whitening filter, a unitary beamformer (UB) and a quadratic postprocessor (QPP).

These results provide physical insight into the structure of the nonstationary space-time LLP. More specifically, they point to the fundamental (and distinct) roles played by the UB and the QPP. The UB is the first processing stage in computing the LLF for the prewhitened sensor outputs. It transforms its M-component input into a N-component output by applying a unitary transformation that maximizes the array gain. This transformation is independent of the source signal statistics and is entirely characterized by the transmission process and the noise statistics. The postprocessor finally computes the LLF by performing a quadratic (power-like) operation on the output of the UB. In fact, the postprocessor is itself a LLP for the output of the UB. This space-time processing dichotomy, i.e., unitary beamforming followed by quadratic postprocessing, has many practical implications. First, since the UB is independent of θ_a, its N-component output is a sufficient statistic for the estimation of θ_a. That is, the output of the UB contains all the information necessary to estimate θ_a when θ_l is fixed. Second, when $N << M$, the UB drastically reduces the number of time functions that need to be quadratically processed, therefore resulting in computational simplifications. Finally, since the postprocessor is a LLP, standard techniques are available for its implementation.

Besides unifying many earlier results, the present study provides a useful framework for tackling many practical, as yet unsolved problems in nonstationary array processing. For example, in [22], we have used the factorization properties of Section III to derive the maximum likelihood estimator of time delay for short observation intervals (this problem is nonstationary in nature because of edge effects). This resulted in a novel processor with improved performance when compared to conventional time delay estimators.

Appendix I

Proof of Property 1:

a) Using (32) and (28), we have

$$R_s\phi_i = UKU^*U\psi_i = UK\psi_i = \lambda_i U\psi_i = \lambda_i\phi_i \quad (A1)$$

and using (30), we have

$$(\phi_i, \phi_j)_2 = (U\psi_i, U\psi_j)_2 = (\psi_i, \psi_j)_1 = \delta_{ij}. \quad (A2)$$

b) Let $\xi \in \boldsymbol{R} \cap \{\phi_i\}^\perp$. We must show that $\xi = 0$ [15, p. 33]. Because $\xi \in \boldsymbol{R}$, (29) implies that $\xi = U\alpha$, with $\alpha = U^*\xi \in N^\perp$. Because $\xi \in \{\phi_i\}^\perp$, we have

$$0 = (\xi, \phi_i)_2 = (U\alpha, U\psi_i)_2 = (\alpha, \psi_i)_1 \quad \text{(A3)}$$

for all indices i. Since $\{\psi_i\}$ is complete in $N^\perp$ by assumption, (A3) implies $\alpha = 0$ and, consequently, $\xi = U\alpha = 0$.

c) Recall that the null space of L^* is $\boldsymbol{R}^\perp$. Hence, if $\xi \in \boldsymbol{R}^\perp$, $R_s\xi = LR_aL^*\xi = 0$.

Proof of Property 2: It can be easily verified that

$$(I_2 + UKU^*)^{-1} = I_2 - UK(I_1 + U^*UK)^{-1}U^*. \quad \text{(A4)}$$

This identity can be regarded as an extension of the conventional matrix inversion lemma to linear operators. Because the range of K is included in $N^\perp$, it follows from (28) that $U^*UK = K$. Hence (A4) reduces to

$$(I_2 + UKU^*)^{-1} = I_2 - UK(I_1 + K)^{-1}U^*. \quad \text{(A5)}$$

Making the substitution $R_s = UKU^*$ in (12′) and using (A5), we obtain

$$H_2 = UKU^*(I_2 + UKU^*)^{-1} = UK(I_1 + K)^{-1}U^*. \quad \text{(A6)}$$

Proof of Property 3: From (7) and (13), we have

$$\int_J \mathrm{Tr}\{h_2(t, t)\}\, dt = \sum_{i=1}^{\infty} \ln(1 + \lambda_i) \quad \text{(A7)}$$

where λ_i are the eigenvalues of R_s. Because (43) linking h_1 to K is identical to (14) linking h_2 to R_s, we can write directly

$$\int_J \mathrm{Tr}\{h_1(t, t)\}\, dt = \sum_{i=1}^{\infty} \ln(1 + \gamma_i) \quad \text{(A8)}$$

where γ_i are the eigenvalues of K. Since, by Property 1, R_s and K have the same eigenvalues, (42) follows from (A7) and (A8).

Appendix II

Proof of Property 3′: Let τ be a parameter with value in $I = [t_i, t_f]$ and consider the eigenvalue problem

$$\int_{t_i}^{\tau} du\, R_a(t, u) \int_{t_i}^{\tau} dv\, W(u, v)\eta_i(v; \tau) = \lambda_i(\tau)\eta_i(t; \tau), \quad t_i \le t \le \tau \quad \text{(A9)}$$

$$\int_{t_i}^{\tau} ds\, \eta_i^T(s; \tau) \int_{t_i}^{\tau} dt\, W(s, t)\eta_j(t; \tau) = \delta_{ij} \quad \text{(A10)}$$

where $W(t, u)$ is the integral kernel associated with the operator W. For $\tau = t_f$, (A9) and (A10) are equivalent to (46) and (47) (we assume $N = \{0\}$ and therefore $P = I_1$ in (46)). Hence, $\lambda_i(t_f) = \lambda_i$. Applying

$$\int_{t_i}^{\tau} ds\, \eta_i^T(s; \tau) \int_{t_i}^{\tau} dt\, W(s, t) \quad \text{(A11)}$$

on both sides of (A9) and using (A10) to simplify the result, we find

$$\lambda_i(\tau) = \int_{t_i}^{\tau}\int_{t_i}^{\tau} dt\, du\, \zeta_i^T(t; \tau)R_a(t, u)\zeta_i(u; \tau) \quad \text{(A12)}$$

where

$$\zeta_i(u; \tau) \equiv \int_{t_i}^{\tau} dv\, W(u, v)\,\eta_i(v; \tau). \quad \text{(A13)}$$

Hence, $\lambda_i(t_i) = 0$. Because of the particular boundary values of $\lambda_i(\tau)$, we can write the bias term l_2 (7) in the form

$$l_2 = \sum_{i=1}^{\infty}\left\{\ln(1 + \lambda_i(t_f)) - \ln(1 + \lambda_i(t_i))\right\} = \sum_{i=1}^{\infty}\int_{t_i}^{t_f} \frac{d\lambda_i(\tau)/d\tau}{1 + \lambda_i(\tau)}\, d\tau. \quad \text{(A14)}$$

To obtain a simple expression for $d\lambda_i(\tau)/d\tau$, we proceed as in the derivation of [1, eq. (3-163)]. To begin, observe that

$$\int_{t_i}^{\tau} du\, R_a(t, u)\,\zeta_i(u; \tau) = \lambda_i(\tau)\eta_i(t; \tau). \quad \text{(A15)}$$

This result follows directly from (A9) and (A13). Differentiating (A12) with respect to τ and using (A15) to simplify the result, we find

$$\frac{d\lambda_i(\tau)}{d\tau} = 2\lambda_i(\tau)\left\{\eta_i^T(\tau; \tau)\zeta_i(\tau; \tau) + \int_{t_i}^{\tau} dt\, \eta_i^T(t; \tau)\frac{d\zeta_i(t; \tau)}{d\tau}\right\}. \quad \text{(A16)}$$

Using (A13), it is not difficult to verify that the integral in (A16) is equal to

$$\frac{1}{2}\frac{d}{d\tau}\left\{\int_{t_i}^{\tau} dt\, \eta_i^T(t; \tau)\zeta_i(t; \tau)\right\} \quad \text{(A17)}$$

which is equal to zero in light of (A10) and (A13). Hence

$$\frac{d\lambda_i(\tau)}{d\tau} = 2\lambda_i(\tau)\eta_i^T(\tau; \tau)\zeta_i(\tau; \tau). \quad \text{(A18)}$$

With the help of (A18), (A14) can be written as

$$l_2 = 2\int_{t_i}^{t_f} \mathrm{Tr}\{F(\tau, \tau; \tau)\}\, d\tau \quad \text{(A19)}$$

where $F(t, u; \tau) \equiv 0$ for $t > \tau$ or $u > \tau$, and

$$F(t, u; \tau) \equiv \sum_{i=1}^{\infty} \frac{\lambda_i(\tau)}{1 + \lambda_i(\tau)}\zeta_i(t; \tau)\eta_i^T(u; \tau) \quad \text{(A20)}$$

for $t_i \le t, u \le \tau$.

We note that in the space of N-vector functions defined over $[t_i, \tau]$ with scalar product $(\,.\,,\,.\,)_w$ (45), $\int du R_a(t, u)\, W(u, v)$ is the kernel of a self-adjoint operator. Hence, [15, theorem 7.5.2] implies that

$$\int_{t_i}^{\tau} du\, R_a(t, u) W(u, v) = \sum_{i=1}^{\infty} \lambda_i(\tau)\eta_i(t; \tau)\zeta_i^T(v; \tau) \quad (A21)$$

for $t_i \leq t, v \leq \tau$. Using the expansion (A20) and (A21), it is not difficult to verify that $F(t, u; \tau)$ satisfies

$$F(t, u; \tau) + \int_{t_i}^{\tau} dt'\, F(t, t'; \tau) \cdot \int_{t_i}^{\tau} du'\, W(t', u') R_a(u', u) = \int_{t_i}^{\tau} du'\, W(t, u') R_a(u', u), \qquad t_i \leq t, u \leq \tau. \quad (A22)$$

Property 3′ follows from (A19) and (A22) if we define $f(t, u) = F(t, u; t)$.

Appendix III

Proof of (53): Let τ be a parameter in $[t_i, t_f]$ and consider the integral equation

$$G(t, u; \tau) + \int_{t_i}^{\tau} dt'\, G(t, t'; \tau) \int_{t_i}^{\tau} \cdot du'\, W(t', u') R_a(u', u) = R_a(t, u), \qquad t_i \leq t, u \leq \tau. \quad (A23)$$

Equation (A23) is identical to that defining the kernel $G_1(t, u) = G_1^T(u, t)$ of the self-adjoint operator G_1 (50), except that in the latter equation, the upper limit of integration is t_f rather than τ. Therefore, like $G_1(t, u)$

$$G(t, u; \tau) = G^T(u, t; \tau), \qquad t_i \leq t, u \leq \tau \quad (A24)$$

and, obviously

$$G(t, u; t_f) = G_1(t, u), \qquad t_i \leq t, u \leq t_f. \quad (A25)$$

For $\tau = t$, (A23) is identical to (54). Consequently

$$G(t, u; t) = g(t, u), \qquad t_i \leq u \leq t. \quad (A26)$$

Let α be an arbitrary function in S_1. Using (A24)-(A26), we have

$$\begin{aligned}[G_1\alpha](t) &= \int_{t_i}^{t_f} du\, G(t, u; t_f)\alpha(u) \\ &= \int_{t_i}^{t} du\, G(t, u; t)\alpha(u) + \int_{t}^{t_f} d\tau \frac{\partial}{\partial\tau} \int_{t_i}^{\tau} du\, G(t, u; \tau)\alpha(u) \\ &= \int_{t_i}^{t} du\, g(t, u)\alpha(u) + \int_{t}^{t_f} du\, g^T(u, t)\alpha(u) + \int_{t}^{t_f} d\tau \int_{t_i}^{\tau} du \frac{\partial}{\partial\tau} G(t, u; \tau)\alpha(u). \end{aligned} \quad (A27)$$

Proceeding as in the derivation of [13, eq. (2-62)], it can be shown that

$$\frac{\partial}{\partial\tau} G(t, u; \tau) = -g^T(\tau, t) \int_{t_i}^{\tau} dv\, W(\tau, v) G(v, u; \tau) - \left\{ \int_{t_i}^{\tau} dv\, W(\tau, v) G(v, t; \tau) \right\}^T g(\tau, u). \quad (A28)$$

But

$$\int_{t_i}^{\tau} dv\, W(t, v) G(v, u; \tau) = F(t, u; \tau) \quad (A29)$$

since the quantity on the left is a solution to (A22) (this follows from (A23)). Therefore

$$\frac{\partial}{\partial\tau} G(t, u; \tau) = -g^T(\tau, t) f(\tau, u) - f^T(\tau, t) g(\tau, u). \quad (A30)$$

To complete the proof, simply substitute this expression back into (A27).

Appendix IV

Proof of (64): By definition, we have

$$\mathrm{SNR}_{\mathrm{in}} = \frac{1}{M} \int_J \mathrm{Tr}\left\{ R_s(t, t) \right\} dt \quad (A31)$$

$$\mathrm{SNR}_{\mathrm{out}} = \frac{1}{N} \int_I \mathrm{Tr}\left\{ [\Omega R_s \Omega^*](t, t) \right\} dt. \quad (A32)$$

According to Mercer expansion [19]

$$R_s(t, u) = \sum_i \lambda_i \phi_i(t) \phi_i^T(u). \quad (A33)$$

Substituting this expansion into (A31) and (A32), we obtain

$$M\,\mathrm{SNR}_{\mathrm{in}} = \sum_i \lambda_i \quad (A34)$$

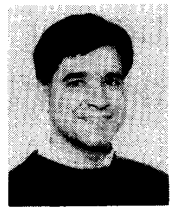
$$N\,\mathrm{SNR}_{\mathrm{out}} = \sum_i \lambda_i \|\Omega\phi_i\|_1^2. \quad (A35)$$

Observe that $\|\Omega\phi_i\|_1 \leq \|\Omega\| \|\phi_i\|_1 = \|\Omega\| = 1$. The last equality follows because $\|\Omega\|^2 = \|\Omega\Omega^*\| = 1$ (for properties of the operator norm, see [17]). Therefore, $N\,\mathrm{SNR}_{\mathrm{out}} \leq M\,\mathrm{SNR}_{\mathrm{in}}$. Finally, when $\Omega = U^*$, the equality holds because $U^*\phi_i = U^*U\psi_i = \psi_i$ and $\|\psi_i\|_1 = 1$.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.
[2] F. Bryn, "Optimum signal processing of three-dimensional arrays operating on Gaussian signals and noise," *J. Acoust. Soc. Amer.*, vol. 34, pp. 289-297, Mar. 1962.
[3] H. Cox, "Optimum arrays and the Schwarz inequality," *J. Acoust. Soc. Amer.*, vol. 45, pp. 228-232, 1969.
[4] W. J. Bangs and P. M. Schultheiss, "Space-time processing for optimal parameter estimation," in *Signal Processing*, J. W. R. Griffiths,

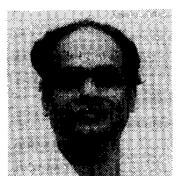
P. L. Stocklin, and C. van Schooneveld, Eds. New York: Academic, 1973, pp. 577-590.
[5] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 608-614, Sept. 1973.
[6] N. L. Owsley and G. R. Swope, "Time delay estimation in a sensor array," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 519-523, June 1981.
[7] M. Wax and T. Kailath, "Optimum localization of multiple sources by passive arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1210-1218, Oct. 1983.
[8] L. C. Ng and Y. Bar-Shalom, "Multisensor multitarget time delay vector estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 669-678, Aug. 1986.
[9] F. C. Schweppe, "Sensor-array data processing for multiple-signal sources," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 294-305, Mar. 1968.
[10] C. H. Knapp and G. C. Carter, "Estimation of time delay in the presence of source or receiver motion," *J. Acoust. Soc. Amer.*, vol. 61, pp. 1545-1549, June 1977.
[11] J. A. Stuller, "Maximum-likelihood estimation of time-varying delay—Part I," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 300-313, Mar. 1987.
[12] I. M. G. Lourtie and J. M. F. Moura, "Optimal estimation of time-varying delay," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* (New York), Apr. 1988, pp. 2622-2625.
[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part III.* New York: Wiley, 1971.
[14] T. Kailath, "Fredholm resolvents, Wiener-Hopf equations, and Riccati differential equations," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 665-672, Nov. 1969.
[15] V. Hutson and J. S. Pym, *Applications of Functional Analysis and Operator Theory.* New York: Academic, 1980.
[16] B. C. Picinbono, "A geometrical interpretation of signal detection and estimation," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 493-497, July 1980.
[17] C. W. Groetsch, *Generalized Inverses of Linear Operators.* New York: Marcel Dekker, 1977.
[18] B. Champagne, "Optimum space-time processing in nonstationary environments," Ph.D. dissertation, Univ. of Toronto, Mar. 1990.
[19] A. V. Balakrishnan, *Applied Functional Analysis.* New York: Springer, 1976.
[20] E. Weinstein and P. M. Schultheiss, "Localization of a moving source using passive array data," Yale University, New Haven, CT, Tech. Rep., Dec. 1978.
[21] P. M. Schultheiss and E. Weinstein, "Estimation of differential Doppler shifts," *J. Acoust. Soc. Amer.*, vol. 66, pp. 1412-1419, Nov. 1979.
[22] B. Champagne, M. Eizenman, and S. Pasupathy, "Exact maximum likelihood time delay estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* (Glasgow, Scotland), May 1989, pp. 2633-2636.



Benoit Champagne (S'87-M'89) was born in Joliette, Quebec, Canada, on January 13, 1961. He received the B.Ing. degree in physics engineering from the Ecole Polytechnique of Montreal in 1983, and the M.Sc. degree in physics from the University of Montreal in 1985. He has recently completed the Ph.D. degree in electrical engineering at the University of Toronto. His current research interests are in the areas of array processing, adaptive filtering, and communication theory.

Moshe Eizenman (M'84) was born in Tel-Aviv, Israel, on December 25, 1952. He received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, Ont., Canada, in 1978, 1980, and 1984, respectively.

Since 1982 he has developed several highly accurate eye-trackers that are currently being integrated into medical and military devices. In 1984 he joined the Institute of Biomedical Engineering at the University of Toronto as an Assistant Professor and founded El-MAR Inc. Since 1984 he has also been an Assistant Professor in the Electrical Engineering and Ophthalmology Departments at the University of Toronto. His current research interests include sonar signal processing, EEG and VEP analysis, control of eye movements, and the development of vision.



Subbarayan Pasupathy (M'73-SM'81) was born in Madras, Tamilnadu, India, on September 21, 1940. He received the B.E. degree in telecommunications from the University of Madras in 1963, the M.Tech. degree in electrical engineering from the Indian Institute of Technology (Madras) in 1966, and the M.Phil. and Ph.D. degrees in engineering and applied science from Yale University in 1970 and 1972, respectively.

He joined the faculty of the University of Toronto in 1973 and became a Professor of Electrical Engineering in 1983. His research interests lie in the areas of communication theory, digital communications, and statistical signal processing. He is a registered Professional Engineer in the province of Ontario. During 1982-1989 he was an Editor for *Data Communications and Modulation* for the IEEE TRANSACTIONS ON COMMUNICATIONS. He has also served as a Technical Associate Editor for the IEEE COMMUNICATIONS MAGAZINE (1979-1982) and as an Associate Editor for the *Canadian Electrical Engineering Journal* (1980-1983). Since 1984, he has been writing a regular column entitled "Light Traffic" for the IEEE COMMUNICATIONS MAGAZINE.