Multisource Delay Estimation: Nonstationary Signals

Isabel M. G. Lourtie, *Member, IEEE*, and José M. F. Moura, *Senior Member, IEEE*

Abstract—**This paper studies delay estimation in a very general framework: multiple sources, correlated noises, nonstationary random signals, and varying delays. We derive the optimal maximum likelihood (ML) delay estimator, analyze its performance via the Cramer–Rao inequality, and test by experimentation with synthetic data. The present time-varying delay estimator extends to the nonstationary/multisource environment the estimators of Ng and Bar-Shalom, Kirlin and Dewey, and Knapp and Carter. However, as it will be apparent, our receiver significantly departs from the correlator structures of these authors.**

I. Introduction

TIME delay estimation is an important problem in underwater acoustics. In positioning systems, delay estimation is a first stage that feeds into subsequent processing blocks. The postprocessing blocks use the delay estimates as features from which location parameters and the dynamics of targets are then recovered. In underwater acoustics, most of the work reported on time delay estimation assumes a single source configuration and stationary signal processes observed over a long observation time interval. These are not always realistic assumptions. Noisy targets emanate acoustic signals that are observed at an array of hydrophones. Several targets may be present in the vicinity of the one being located or tracked. At the array of sensors we receive then the superposition of all the radiated signals. Over time, the received signal cannot be assumed to be stationary. The transmission medium, the ocean, shows a complex behavior exhibiting simultaneously temporal and spatial dependence. The radiated signatures are themselves nonstationary. The emitter/receiver relative motions are often not negligible. In this context, a more accurate modeling of the problem is often needed, justifying a multisource configuration with nonstationary signals. However, as we will show, this leads to a rather more complex receiver than the one traditionally used.

For a narrow-band (NB) random signal $x(t)$ with complex signal representation

$$\tilde{x}_b(t)e^{j2\pi f_c t} \tag{1}$$

the delayed replica $x(t - \tau)$ has a complex envelope which is approximately given by

$$\tilde{x}_b(t)e^{-j2\pi f_c \tau}. \tag{2}$$

Manuscript received May 6, 1988; revised June 25, 1990.
I. M. G. Lourtie is with CAPS, Complexo, Instituto Superior Técnico, P-1096 Lisboa Cedex, Portugal.
J. M. F. Moura is with LASIP, Department of Electrical and Computer Engineering, Carnegie-Mellon University, Pittsburgh, PA 15213.
IEEE Log Number 9042718.

In (2), we used the NB approximation

$$\tilde{x}_b(t - \tau) \simeq \tilde{x}_b(t). \tag{3}$$

On the other hand, for a stationary random signal with a large time bandwidth product BT, working with the Fourier representation, we have the Fourier transform of the delayed replica $x(t - \tau)$ as

$$X(f)e^{-j2\pi f\tau} \tag{4}$$

where $X(f)$ is the Fourier transform of $x(t)$. Equations (2) and (4) show that under assumptions of either NB signals or stationary signals with large BT, the delay dependence factors out of the structure of the signal. This factorization is a key point determining the structure of the present day delay estimators, see, for example, Knapp and Carter [3] who showed that the optimal maximum likelihood (ML) delay estimator for stationary signals with large BT is essentially a cross correlator.

We consider in this paper the problem of delay estimation for nonstationary random signals with arbitrary time bandwidth product BT. For these signals, the factorization of the type of equation (2) or (4) is no longer possible. We derive the structure of the ML delay estimator which turns out to be remarkably different from that of a cross correlator. Lacking stationarity, the Fourier transform techniques that underly the design of these cross correlators are no longer useful. Rather, the maximization step of the ML delay estimator is preceded by a recursive filter. This filter estimates the random signal conditioned on knowledge of the values of the delays. It generalizes the Kalman–Bucy filter to the case of estimation of waveforms with delays.

In deriving the ML estimator, we consider a very general setup for the problem: i) nonstationary correlated multiple source signals, ii) time-varying delays (relative or absolute) parameterized by a finite number of unknown quantities, iii) nonstationary, possibly mutually correlated, observation noises, iv) arbitrary (short or long) observation time interval. In doing so, we extend to the nonstationary, arbitrary BT case, the works of Knapp and Carter [3], [4], and of Ng and Bar-Shalom [1], who studied the problem of multisource environment, and of Kirlin and Dewey [2], that assumed a single source with spatially correlated measurement noise.

When the sources are moving, besides delay, we need to estimate the Doppler shift. We derive the joint optimal ML delay/Doppler estimator. This extends to the more general class of nonstationary signals with arbitrary BT the works of Knapp and Carter [5], and of Kirlin, Moore,

1053-587X/91/0500-1033$01.00

and Kubichek [6]. The latter work develops a suboptimal strategy to deal with signals where the nonstationarity is due to time-varying delays. This is achieved by sequentially processing the data into continuous blocks of length T seconds. Within each block the delay is assumed to be time invariant, but allowed to change over successive blocks. In [6], a linear Kalman filter is used as a postprocessor that filters the sequence of delay measurements. These delay estimates result from processing the most recent received data block by means of a delay estimator such as the one described by Hassab and Boucher [7]. For narrow-band signals, and when the delays are themselves sample functions of a random process Bucy, Moura, Mallinckrodt [8], Leitão and Moura [9], Bethel and Rahikka [10] have developed Bayes law type receivers to estimate the delay waveform. Moura and Baggeroer [11] have used the estimators of [9] with real data underwater signals that propagated under the Arctic ice crust.

In Section II, we formulate and model the delay estimation problem. A multisource environment is considered. The signals are sample functions of Gaussian nonstationary stochastic processes, described via a state space model, and the delays are continuous real functions. The maximum likelihood (ML) delay estimator for the above context is derived in Section III. The ML delay estimator maximizes, over the unknown delay parameters, the log-likelihood function (LLF). When the source signals $y_l(t)$ are not deterministic known functions, but rather sample functions of nonstationary random processes, the construction of the likelihood function involves the minimum mean-square error estimate (MMSE) $\hat{y}_l(t - D_l^s)$ of the delayed source signals $y_l(t - D_l^s)$ given the observation set, and conditioned on the values of the delays D_l^s. This is the first step of the ML receiver. We develop the causal MMSE recursive filter for $\hat{y}_l(t - D_l^s)$ which exhibits the structure of a Kalman–Bucy filter for signals with delays. These MMSE estimates $\hat{y}_l(t - D_l^s)$, which are functions of the unknown delays D_l^s's, are used in the LLF whose maxima yield the delay estimates. In this sense, the ML-delay receiver here studied generalizes to unknown nonstationary random signals, the FSK-type receiver of radar contexts. To evaluate the estimator performance, in Section IV we develop analytically the time evolution of the Cramer–Rao bound. Under the general framework considered herein, the Fisher information matrix requires the gradient with respect to the delays of the estimate of the signal process. Our parameterization in terms of the GKBF, promptly provides algorithmically the means to compute these gradients. To gain insight into the general ML delay estimator developed herein, in Section V we analyze its asymptotic structure under stationary signals, long observation time interval (SLOT approximation), and time-invariant delay assumptions. For the SLOT case, we provide a frequency interpretation showing that the new delay estimator recovers the structure of a generalized cross correlator, [1]–[4]. Assuming a multisource environment with spatially uncorrelated observation noise and SLOT conditions we obtain the processor of Ng and Bar-Shalom [1], while considering a single source configuration we recover either the estimator of Kirlin and Dewey [2], for spatially correlated measurement noise, or the generalized cross correlator of Knapp and Carter [3], [4], when the observation noise is uncorrelated among sensors. In Section VI, we carry out a simulation study that illustrates the capabilities of the general ML processor and compare it with the classical generalized cross-correlator structure. Finally, Section VII presents the paper's conclusions.

Notation: The general setup of the problem involves a detailed notation that is presented here. The following symbols are used throughout:

i) $\otimes$: Kronecker product [12] which, for example, for $X, Y \in \mathcal{R}^2 \times \mathcal{R}^2$, is defined by

$$X \otimes Y = \begin{bmatrix} x_{11} Y & x_{12} Y \\ x_{21} Y & x_{22} Y \end{bmatrix} \in \mathcal{R}^4 \times \mathcal{R}^4. \tag{5}$$

ii) diag $\{\bullet\}$: diagonal matrix of block-matrix elements, e.g.,

$$\text{diag}\,\{A_l\} = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_L \end{bmatrix}, \qquad l = 1, 2, \cdots, L \tag{6}$$

where A_l are matrices of appropriate dimensions.

iii) I_S: identity matrix of order $S \times S$.

iv) $\mathbf{1}_S = [1 \cdots 1]^T$: one vector of order S.

v) $\nabla_{tr_1r_2}(*)$, $\nabla_{tr}(*)$, and $\nabla_t(*)$: differential operators defined by

$$\nabla_{tr_1r_2}(*) = \gamma(t) \left[\frac{\partial}{\partial t}(*)\right] \gamma(t) + \beta(t, r_1) \left[\frac{\partial}{\partial r_1}(*)\right] \gamma(t) + \gamma(t) \left[\frac{\partial}{\partial r_2}(*)\right] \beta(t, r_2) \tag{7}$$

$$\nabla_{tr}(*) = \gamma(t) \left[\frac{\partial}{\partial t}(*)\right] + \beta(t, r) \left[\frac{\partial}{\partial r}(*)\right] \tag{8}$$

$$\nabla_t(*) = \frac{d}{dt}(*) \tag{9}$$

where t, r_1, r_2 are scalars representing time and delay variables, γ and β are diagonal matrices.

vi) $\langle *, \bullet \rangle_R$: weighted inner product, i.e.,

$$\langle *, \bullet \rangle_R = *^T R^{-1} \bullet \tag{10}$$

where R is a covariance matrix.

vii) $\text{Tr}_N(\bullet)$: matricial trace operator, defined as follows: given a block square matrix

$$X = [X_{ij}] \in \mathcal{R}^{MN} \times \mathcal{R}^{MN} \tag{11}$$

with the blocks $X_{ij} \in \mathcal{R}^N \times \mathcal{R}^N$, $\text{Tr}_N(X)$ is the $M \times M$ matrix

$$\text{Tr}_N(X) = [\text{tr } X_{ij}] \in \mathcal{R}^M \times \mathcal{R}^M. \quad (12)$$

viii) On occasion, a short notation is used for functions, e.g.,

$$P_{tr_1r_2} = P(t, r_1, r_2). \quad (13)$$

When no ambiguity arises, arguments may also be omitted.

II. Problem Formulation

A. Statement of the Problem

We now formulate the *multiple source* delay estimation problem, when the processor array has $S \geq 1$ sensors, and the receiving signal at each sensor is a superposition of $L \geq 1$ sources. Specifically, let $\{z(t, s) \in \mathcal{R}, t \geq 0\}$ be a sample function of the observed process at sensor s, modeled by

$$z(t, s) = \sum_{l=1}^{L} \alpha_l^s(t) y_l(t - D_l^s(t)) + v(t, s),$$
$$t \in \boldsymbol{T}, \quad s \in \boldsymbol{S}, \quad l \in \boldsymbol{L} \quad (14)$$

where $t \in \boldsymbol{T} = [0, T] \subseteq \mathcal{R}^+$ is the time variable, $s \in \boldsymbol{S} = \{1, 2, \cdots, S\}$ is the sensor index, and $l \in \boldsymbol{L} = \{1, 2, \cdots, L\}$ is the source index. Functions $D_l^s(t)$ and $\alpha_l^s(t)$ represent the *time-varying* delay and the attenuation associated to each pair (s, l). The observation noise $\{v(t, s) \in \mathcal{R}, t \in \boldsymbol{T}\}$ is a nonstationary zero mean Gaussian temporally white *spatially correlated* noise process with cross-covariance $E\{v(t, s)v(\sigma, m)\} = R_{sm}(t)\delta(t - \sigma)$, $(s, m \in \boldsymbol{S})$. The signal $\{y_l(t) \in \mathcal{R}, t \geq t_0\}$ emanating from source $l \in \boldsymbol{L}$ is a stochastic *nonstationary* process assumed to be uncorrelated with the noise $\{v(t, s)\}$. Throughout the paper, the signals are assumed scalar. Generalization to vector signals is trivial.

At each sensor of the receiving array, the observed signa' is the superposition of delayed (travel time) and attenuated replicas of the emitted signals. Inherent to the problem description is a propagation effect introduced through the travel time delays $(D_l^s(t))$ and the attenuation factors $(\alpha_l^s(t))$. These are functions of both the transmission channel impulse response, and of the sources and sensors geometry and relative motion.

Equation (14) describes the problem. It encompasses a *multisource* configuration with a single direct acoustic path from each source $l \in \boldsymbol{L}$ to each sensor $s \in \boldsymbol{S}$. It also includes the single source *multipath* geometry if we take a single signal $y(t) = y_l(t)$ propagating through multiple paths $l \in \boldsymbol{L}$ from the source to each sensor $s \in \boldsymbol{S}$ of the array. Fig. 1 shows the multisource configuration for two sources and an array of three hydrophones.

In our formulation, the time-varying delay function $D_l^s(t)$, $\forall t \in \boldsymbol{T}$, is assumed to satisfy the following conditions:

i) it is a continuous function of t;
ii) it has first derivative with respect to t;

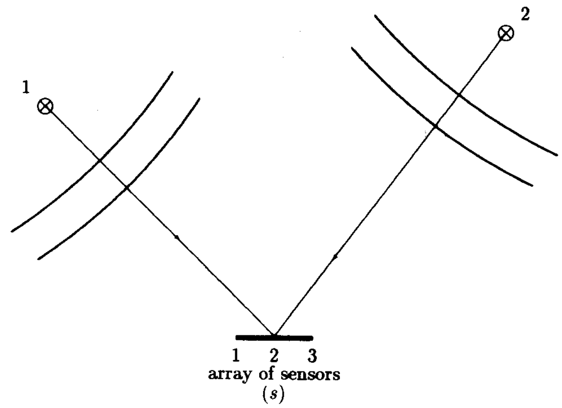


Fig. 1. Multisource, single direct acoustic path geometry.

iii) it is described by a deterministic time-varying function which is parameterized by a finite number of unknown parameters, i.e.,

$$D_l^s(t) = \tau_l(t, \theta_l) \quad (15)$$

where θ_l is the vector of unknown parameters.

Under i) above, for bounded $\boldsymbol{T}$, we can define

$$D_{\min} \leq \min_{t \in \boldsymbol{T}, s \in \boldsymbol{S}, l \in \boldsymbol{L}} D_l^s(t) \quad (16)$$

and

$$D_{\max} \geq \max_{t \in \boldsymbol{T}, s \in \boldsymbol{S}, l \in \boldsymbol{L}} D_l^s(t). \quad (17)$$

The values $D_{\min}$ and $D_{\max}$ are, respectively, lower and upper bounds to all the time-varying delay functions $D_l^s(t)$. In the sequel, we need the technical assumption

$$t_0 \leq -D_{\max} \quad (18)$$

where t_0 represents a time reference with respect to which all sources start radiating the signals $y_l(t)$, $l \in \boldsymbol{L}$. Equation (18) guarantees that the process $y_l(t - D_l^s(t))$ in (14) exists for every $t \in \boldsymbol{T}$, that is, all signal sources are present during the entire observation time interval $\boldsymbol{T}$. For a linear array, and depending upon the case under study, absolute or relative delay, the parameter $D_{\min}$ is defined accordingly. For an absolute delay, since

$$\forall t \in \boldsymbol{T}, \quad s \in \boldsymbol{S}, \quad l \in \boldsymbol{L}, \quad D_l^s(t) \geq 0$$

$D_{\min}$ has the trivial value

$$D_{\min} = 0.$$

For the relative delay case, i.e., when the differential delays $D_l^s(t)$ can be negative, $D_{\min}$ is a function of the intersensors separation and/or of the maximum range of the time-varying delays. If there are no constraints on the source location, the relative delay D_l^s is minimum or maximum for the endfire configuration.

The work presented in this paper does not assume any particular parameterization $\tau_l(t, \theta_l)$ of the delay functions. It remains valid for a general parameterization. A discussion of parameterizations for a planar source/receiver geometry, with a linear array shape and a target following a linear uniform motion can be found in Moura and Baggeroer [13], and Moura [14], [15]. The ML estimation of the time-varying delays $D_l^s(t)$ is accomplished after the ML estimates of the parameters are performed.

To model general nonstationary source signatures, the signals $y_l(t)$ are described as sample functions of Gauss Markov processes, which are the output of a linear system driven by noise. In formal terms, the source signature is described by

$$\frac{d}{dt} x_l(t) = A_l(t)x_l(t) + B_l(t)u_l(t), \qquad t \geq t_0,\ l \in L, \quad (19)$$

$$y_l(t) = C_l(t)x_l(t). \quad (20)$$

The initial condition $x_l(t_0) \in \Re^{n_l}$ is a Gaussian random vector with mean value $\bar{x}_l(t_0) \in \Re^{n_l}$ and cross-covariance matrix $E\{[x_l(t_0) - \bar{x}_l(t_0)][x_k(t_0) - \bar{x}_k(t_0)]^T\} = \Sigma_{lk}(t_0, t_0)$ $(l, k \in L)$. The dynamics disturbance $\{u_l(t) \in \Re^{m_l}, t \geq t_0\}$ is a Gaussian white noise vector process, independent of the observation noise $\{v(t, s)\}$ and of the random initial condition $x_k(t_0)(k \in L)$, with cross-covariance matrix $E\{u_l(t)u_k^T(\sigma)\} = Q_{lk}(t)\delta(t - \sigma)$. For $l \neq k$, the noises $u_l(t)$ and $u_k(t)$ may be statistically correlated. In this case, the signatures $y_l(t)$ and $y_k(t)$, emitted by sources l and k, are correlated. If the sources are mutually uncorrelated, we simply take

$$Q_{lk}(t) = 0 \quad \text{and} \quad \Sigma_{lk}(t_0, t_0) = 0, \qquad \forall l \neq k. \quad (21)$$

Depending on the system matrices $A_l(t)$, $B_l(t)$, and $C_l(t)$, (19), (20) model *nonstationary* narrow- or broad-band processes. When the system matrices are time invariant, the matrix A is asymptotically stable, and the observation interval is large, model (19), (20) leads then to signals which are asymptotically stationary. Equations (19), (20) represent a paradigm which is alternative to Fourier representation of signals. The latter describes signals as linear superposition of complex exponentials, and in the context of delay estimation leads, under a large BT assumption, to batch cross-correlation techniques. Equations (19), (20) represent the signals as output of linear systems driven by noise. They are well-suited representations to handle nonstationary environments and lead to estimators that emphasize time recursiveness. Equations (14) and (19), (20) are to be interpreted in a formal sense. A rigorous writing requires the Ito calculus formulation, e.g., see [16].

B. Modeling

As referred to in Section I, the log-likelihood function involves the minimum mean-square error (MMSE) estimate of the received signals. Because at different sensors, these correspond to delayed replicas of sample functions of nonstationary random processes, we need to develop filtering structures for nonstationary stochastic processes with delays. The complexity of the problem is brought about not so much by the nonstationarity but more so by the necessity of filtering with delay of random signals. We achieve this by reformulating the problem into the context of linear filtering of signals described by partial differential equations (distributed parameter systems). As we will see, the present problem corresponds, however, to a very specific context within this very general topic.

To be able to obtain a compact description we work with vector and matrix notations. For that, define the extended processes that collect the snapshots over the array of sensors:

i) Received signal vector:

$$Z(t) = [z(t, 1) \vdots z(t, 2) \vdots \cdots \vdots z(t, S)]^T. \quad (22)$$

ii) Measurement noise vector:

$$V(t) = [v(t, 1) \vdots v(t, 2) \vdots \cdots \vdots v(t, S)]^T. \quad (23)$$

iii) Delay line vector $D_l(t)$ for source l:

$$D_l(t) = [D_l^1(t) \vdots D_l^2(t) \vdots \cdots \vdots D_l^S(t)]. \quad (24)$$

iv) The $L \times S$ delay matrix:

$$D(t) = [D_1^T(t) \vdots D_2^T(t) \vdots \cdots \vdots D_L^T(t)]^T. \quad (25)$$

v) State vector:

$$X(t, r) = [X^{1^T}(t, r) \vdots X^{2^T}(t, r) \vdots X^{S^T}(t, r)]^T. \quad (26)$$

vi) State vector for sensor s:

$$X^s(t, r) = [x_1^T(\xi_1^s(t, r)) \vdots x_2^T(\xi_2^s(t, r)) \vdots \cdots \vdots x_L^T(\xi_L^s(t, r))]^T. \quad (27)$$

vii) Equivalent auxiliary time variable:

$$\xi_l^s(t, r) = t - D_{\min} - r(D_l^s(t) - D_{\min}). \quad (28)$$

viii) Input noise vector:

$$U(t) = [u_1^T(t - D_{\min}) \vdots u_2^T(t - D_{\min}) \vdots \cdots \vdots u_L^T(t - D_{\min})]^T. \quad (29)$$

The parameter $r \in [0, 1]$ in (26)–(28) is useful in describing the evolution in the time interval $[t - D_l^s(t), t - D_{\min}]$ of the state vectors $x_l (l \in L)$ that model the emitted signals y_l (see (19), (20)). It introduces gradually and implicitly the nonstationarity imposed on the signals by the time-varying delays. In (28), $\xi_l^s(t, r)$ is an auxiliary time variable that represents, for fixed observation time t, the time evolution of the emitted signal when propagating along path (l, s) from source l to sensor s. For example, if $D_l^s(t)$ is an absolute delay, we have

$$\xi_l^s(t, r) = t - rD_l^s(t). \quad (30)$$

In the above context, at fixed time t, along path (l, s) from source l to sensor s the emitted signal is at source l

$$y_l(t) = y_l(\xi_l^s(t, 0)) \tag{31}$$

while at the receiver is

$$y_l(t - D_l^s(t)) = y_l(\xi_l^s(t, 1)). \tag{32}$$

In between, $y_l(\xi_l^s(t, r))$ is a nonobservable signal at a fictitious sensor located a distance $rd_l^s(t)$ apart from the emitter (Fig. 2). When $r = 1$, $X(t, 1)$ collects the state vectors $x_l(t - D_l^s(t))$ used in modeling the signals received at each sensor of the array.

Under definitions (22)–(29), the multisource delay problem is now compactly described. The vector process $\{Z(t) \in \mathbb{R}^S, t \in \boldsymbol{T}\}$ observed at the receiving array is

$$Z(t) = Y(t) + V(t) \tag{33}$$

where the observation noise $\{V(t) \in \mathbb{R}^S, t \in \boldsymbol{T}\}$ is a nonstationary Gaussian white noise vector process with covariance matrix $E\{V(t)V^T(\sigma)\} = R(t)\delta(t - \sigma)$. When the noise is spatially uncorrelated, $R(t)$ is diagonal. The received signal $\{Y(t) \in \mathbb{R}^S, t \in \boldsymbol{T}\}$

$$Y(t) = C(t)X(t, 1) \tag{34}$$

is a stochastic nonstationary vector process, uncorrelated with the noise $\{V(t)\}$. It is interpreted as the output of a linear distributed parameter system. From (26), (27) and (19), the state vector $X(t, r)$ is described by the partial differential equation

$$\nabla_{tr}X(t, r) = 0, \qquad r \in [0, 1], \qquad t \geq 0 \tag{35}$$

where $\nabla_{tr}(\cdot)$ is the differential operator introduced in (8), with

$$\gamma(t) = \text{diag}\ \{\gamma^s(t)\} \tag{36}$$

where

$$\gamma^s(t) = \text{diag}\ \{[D_l^s(t) - D_{\min}]I_{n_l}\} \tag{37}$$

and

$$\beta(t, r) = \text{diag}\ \{\beta^s(t, r)\} \tag{38}$$

where

$$\beta^s(t, r) = \text{diag}\ \{[1 - r\nabla_t D_l^s(t)]I_{n_l}\}. \tag{39}$$

The boundary condition for (35) is at $r = 0$

$$\nabla_t X(t, 0) = A(t)X(t, 0) + B(t)U(t). \tag{40}$$

Matrices A, B, and C in model (34)–(40) are given by

$$A(t) = I_S \otimes \text{diag}\ \{A_l(t - D_{\min})\} \tag{41}$$

$$B(t) = 1_S \otimes \text{diag}\ \{B_l(t - D_{\min})\} \tag{42}$$

and

$$C(t) = \text{diag}\ \{C^s(t)\} \tag{43}$$

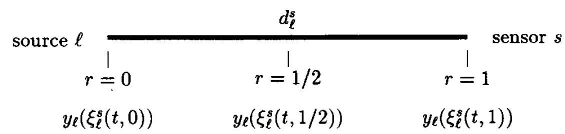


Fig. 2. Path (l, s) between source l and sensor s at a fixed time t.

with

$$C^s(t) = [\alpha_1^s(\xi_1^s)C_1(\xi_1^s) \ \vdots \ \alpha_2^s(\xi_2^s)C_2(\xi_2^s) \ \vdots \ \cdots \ \vdots \ \alpha_L^s(\xi_L^s)C_L(\xi_L^s)] \tag{44}$$

where ξ_l^s, defined by (28), is taken at $r = 1$. Expressions (34)–(44) model the received signal $Y(t)$. Although not explicitly specified, the signal $Y(t)$ and the state vector $X(t, r)$ are functions of the delay matrix $D(t)$.

The partial differential equation (35) describes the propagation of the state vectors $x_l(\xi_l^s)$ of the radiated signals along each propagation path (l, s). The boundary equation (40) describes the signals at the source locations. They correspond to time $t - D_{\min}(r = 0)$, rather than $t = 0$. This is a useful gimmick that allows us to consider relative delays. In this case, $D_{\min} < 0$ and the boundary condition is advanced relative to the signal received at the reference sensor. This procedure represents a time-invariant translation which does not modify the emitted signal characterization.

III. Time-Varying Delay Estimation

A. Estimation Criterion: Maximum Likelihood

The observations along the time interval $[0, t]$ and array sensor set S are collected by

$$Z^t = \{Z(\sigma): 0 \leq \sigma \leq t\}. \tag{45}$$

We assume that the time delays, if time varying, are described by deterministic functions whose structure is specified up to a finite number of unknown parameters collected in the vector θ. Let the time-varying delay realization be

$$D^t(\theta) = \{D(\sigma) = \tau(\sigma, \theta): 0 \leq \sigma \leq t\}. \tag{46}$$

The vector θ is estimated by maximum likelihood (ML) techniques [17], [18], i.e., the ML-estimate $\hat{\theta}(t)$ of θ given the observation set Z^t is

$$\hat{\theta}(t) = \arg \max_{\theta \in \Theta} J(t; D^t(\theta)) \tag{47}$$

where $J(t; D^t)$ is the log-likelihood function (LLF). The maximization operation in (47) is carried out on the compact domain $\Theta \subseteq \mathbb{R}^p$ where the vector θ is assumed to lie.

B. Log-Likelihood Function (LLF)

In this subsection, we present the LLF when the time variable t is continuous. To get insight into the structure of the LLF, we derive it in Appendix A for the discrete time case. For continuous time t the LLF structure is established in Appendix B. The evaluation of the LLF involves the minimum mean-square error (MMSE) causal estimate of the received signal $Y(t)$, conditioned on both the observation set Z^t and the time-varying delay realization $D^t(\theta)$.

The log-likelihood function (LLF) is given by (B.13) of Appendix B

$$J(t; D^t) = \int_0^t [\langle \hat{Y}_\sigma, Z_\sigma \rangle_R - \tfrac{1}{2} \langle \hat{Y}_\sigma, \hat{Y}_\sigma \rangle_R - \tfrac{1}{2} \operatorname{tr}(R^{-1} P_Y)] \, d\sigma \tag{48}$$

where $\hat{Y}(\sigma)$ is the minimum mean-square error (MMSE) estimate of the signal process $Y(\sigma)$, and $P_Y(\sigma)$ is the signal error estimation covariance matrix. Both $\hat{Y}(\sigma)$ and $P_Y(\sigma)$ are functions of the time delay realization $D^\sigma(\theta)$. The first and second terms on the RHS of expression (48) depend on the observation process: the first term computes the cross correlation, normalized by the observation noise covariance, between the observation process $Z(t)$ and the MMSE signal estimate $\hat{Y}(t)$; the second term represents the energy of the MMSE signal estimate. The third term on the RHS of (48) is a correction term. It depends only on the signal estimate error covariance $P_Y(t)$. The next subsection shows that $P_Y(t)$ does not depend on the observation process, and thus can be precomputed.

To conclude, we emphasize that the ML delay estimate $\hat{D}(t)$, which is obtained by maximizing the LLF given by (48), involves the MMSE estimate of the received signal $\hat{Y}(t)$. If the emitted signals were deterministic, that is, the received signal $Y(t)$ was known except for the finite dimensional vector of parameters θ that describes the delay realization $D^t(\theta)$, then it would be trivial to show that

$$\hat{Y}(t) = Y(t) \tag{49}$$

and

$$P_Y(t) = 0 \tag{50}$$

leading to the well-known expression for the LLF for a deterministic signal (see [17, ch.4])

$$J(t, D^t) = \int_0^t [\langle Y_\sigma, Z_\sigma \rangle_R - \tfrac{1}{2} \langle Y_\sigma, Y_\sigma \rangle_R] \, d\sigma. \tag{51}$$

In the general case of random $Y(t)$, we describe in the next subsection how to construct the MMSE estimate $\hat{Y}(t)$ of $Y(t)$ conditioned on the value of the delay.

C. Generalized Kalman–Bucy Filter (GKBF)

In the previous subsection, the LLF was expressed in terms of the minimum mean-square error (MMSE) estimate $\hat{Y}(t)$ of $Y(t)$. Because the signals are delayed sample functions of linear random processes, the MMSE estimates are obtained by generalization of the Kalman–Bucy filter to the distributed parameter context of Section II.

The optimal filtering theory of Kalman and Bucy [19] was first extended by Kwakernaak [20] to include linear systems with multiple time-invariant delays in both the dynamics and the observations. Yu *et al.* [21] generalized Kwakernaak's results to nonlinear distributed parameter systems. The structure of [21] is too general. The generalized Kalman–Bucy filter (GKBF) that we need for the delay estimator is obtained by deriving the causal MMSE filter when delays (possibly time varying) are present only in the observations. The details of the derivation are lengthy. They are available in [18].

Under the general linear framework stated in Section II, the signal MMSE estimate $\hat{Y}(t)$ and the error covariance matrix $P_Y(t)$, appearing in (48), are given by

$$\hat{Y}(t) = C(t)\hat{X}(t, 1) \tag{52}$$

and

$$P_Y(t) = C(t)P(t, 1, 1)C^T(t) \tag{53}$$

where

$$\hat{X}(t, r) = E\{X_{tr} | Z^t, D^t\} \tag{54}$$

and $P(t, r_1, r_2)$ is the error covariance conditioned on D^t

$$P(t, r_1, r_2) = E\{[X_{tr_1} - \hat{X}_{tr_1}][X_{tr_2} - \hat{X}_{tr_2}]^T | D^t\} \tag{55}$$

with $t \geq 0$, $r_1, r_2 \in [0, 1]$. Denoting the filter gain by

$$K(t, r) = P_{tr_1} C_t^T R_t^{-1} \tag{56}$$

the state vector MMSE estimate $\hat{X}(t, r)$ is given by the partial differential equation

$$\nabla_{tr} \hat{X}_{tr} = \gamma_t K_{tr} \tilde{Z}_t, \qquad t \in \boldsymbol{T}, \qquad r \in [0, 1] \tag{57}$$

with the boundary condition

$$\nabla_t \hat{X}_{t0} = A_t \hat{X}_{t0} + K_{t0} \tilde{Z}_t \tag{58}$$

and the initial condition

$$\hat{X}_{0r} = \overline{X}_{0r} \tag{59}$$

where

$$\tilde{Z}(t) = Z(t) - \hat{Y}(t) \tag{60}$$

is the innovations process, and matrix $\gamma(t)$ is defined in (36), (37).

The state estimate error covariance matrix $P(t, r_1, r_2)$ is given by the partial differential equation

$$\nabla_{tr_1r_2} P_{tr_1r_2} = -\gamma_t K_{tr_1} R_t K_{tr_2}^T \gamma_t, \qquad t \in \boldsymbol{T}, \qquad r_1, r_2 \in [0, 1] \tag{61}$$

with the boundary conditions

$$\nabla_{tr} P_{tr0} = \gamma_t P_{tr0} A_t^T - \gamma_t K_{tr} R_t K_{t0}^T \tag{62}$$

$$P_{t0r} = P_{tr0}^T \tag{63}$$

$$\nabla_t P_{t00} = A_t P_{t00} + P_{t00} A_t^T + B_t Q_t B_t^T - K_{t0} R_t K_{t0}^T \tag{64}$$

and the initial condition

$$P_{0r_1r_2} = \Sigma_{0r_1r_2} \tag{65}$$

where the matrix $\Sigma(0, r_1, r_2)$ stands for the cross covariance of the state vectors $X(t, r_1)$ and $X(t, r_2)$.

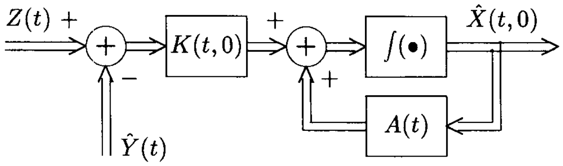


Fig. 3. Generalized Kalman-Bucy filter: boundary condition.

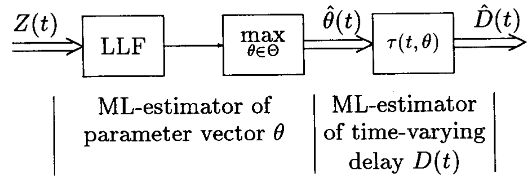


Fig. 4. Time-varying delay estimator.

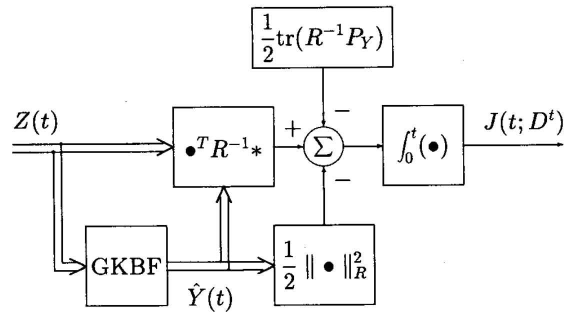


Fig. 5. Log-likelihood function (LLF) block.

Equations (52)–(65) describe the generalized Kalman-Bucy filter (GKBF), which extends the Kalman-Bucy filter [19] to systems with multiple time-varying delays in the observation process. The GKBF is a recursive system that reproduces the dynamics structure of the signal, (35) and (40), driven by the innovations process.

The received signal $Y(t)$ contains different delayed replicas of the emitted signals. The boundary condition (58) computes the estimates of the state vectors x_l, $\forall l \in \boldsymbol{L}$, describing each radiating source, at time instant $t - D_{\min}$. In the above context, the boundary condition (58), Fig. 3, performs a prediction operation in the sense that it gives estimates of the state vectors x_l at a time instant

$$t - D_{\min} \geq t - D_l^s(t), \quad t \in \boldsymbol{T}, \qquad l \in \boldsymbol{L}, \qquad s \in \boldsymbol{S}. \tag{66}$$

On the other hand, the propagation equation (57) evaluates the estimates of the state vectors x_l in the time interval $[t - D_l^s(t), t - D_{\min}]$, that is, it carries out a smoothing operation over the boundary state estimate.

The filter gain requires the solution of the coupled partial differential equations (61)–(64). These represent the extension to filtering with time-varying delays in the observation process of the Riccati equation of the Kalman-Bucy theory [19]. Because of the linearity assumption on the signal process model, the covariance matrix $P(t, r_1, r_2)$ does not depend on the observation process being precomputable. As mentioned, the interested reader is referred to [18] for the complete derivation.

D. Estimator Structure

As referred to before, $\hat{Y}(t)$ depends on the time delay realization $D^t(\theta)$. Based on the same set of data Z^t, distinct delay realizations $D^t(\theta)$ lead to different signal estimates $\hat{Y}(t)$. The ML processor chooses for the delay estimate the delay realization that is most likely to have caused the particular set of observations Z^t, that is, the one that accomplishes a better matching between signal modeling and data. The ML time-varying delay estimator developed herein is conceptually equivalent to two blocks (Fig. 4). It first computes the ML estimate of the parameter vector $\hat{\theta}(t)$. The LLF block generates the log-likelihood function (LLF—expression (48), Fig. 5). It constructs the MMSE estimate of the received signal, conditioned on both the observation set Z^t and the time-varying delay realization $D^t(\theta)$, via the generalized Kalman-Bucy filter. The maximization of the LLF is carried out on the compact domain Θ where the parameter vector θ that describes $D(t)$ is assumed to lie. Thereafter, the time-varying delay estimate $\hat{D}(t)$ is constructed based on both the parameter vector estimate $\hat{\theta}(t)$, and the assumed delay function $\tau(t, \theta)$.

When the second and third terms on the LLF (expression (48)) are weakly dependent on the delay matrix D, for example, when the problem geometry assumes a single stationary source configuration, and the observation noise is uncorrelated among sensors, we have

$$\hat{D}(t) \simeq \arg \max_{\theta \in \Theta} \int_0^t \langle Z(\sigma), \hat{Y}(\sigma) \rangle_R \, d\sigma. \tag{67}$$

IV. Cramer-Rao Bound

A. Introduction

In this section we establish the Cramer-Rao bound (CRB) for the parameter vector estimate $\hat{\theta}(t)$. It is well known that the ML method provides asymptotically optimal Gaussian estimates which are consistent (asymptotically unbiased) and whose variance achieves the CRB (efficiency) [17]. For any unbiased estimator, the CRB is given by the inverse of the Fisher information matrix $\Upsilon(Z^t, \theta^a)$ [17], i.e.,

$$E\{[\hat{\theta}(t) - \theta^a][\hat{\theta}(t) - \theta^a]^T\} \geq \Upsilon^{-1}(Z^t, \theta^a) \tag{68}$$

where $\theta^a \in \mathbb{R}^\nu$ is the actual value of the parameter vector. The $\nu \times \nu$ Fisher information matrix is

$$\Upsilon(Z^t, \theta^a) = -E\left\{ \left. \frac{\partial^2 J(t; \theta)}{\partial\theta\partial\theta^T} \right|_{\theta^a} \right\} \tag{69}$$

and $J(t; \theta)$ is the log-likelihood function (LLF) given by (48).

B. Fisher Information Matrix

Denote by

$$\nabla_a \hat{Y}(t) = \left.\frac{\partial \hat{Y}(t)}{\partial \theta}\right|_{\theta^a} \tag{70}$$

the $S\nu$ dimensional gradient of the MMSE signal estimate $\hat{Y}(t)$ with respect to θ, taken at $\theta = \theta^a$. By expanding in a Taylor series $\hat{Y}(t)$ about the actual parameter vector θ^a, noting that $\hat{Y}(t)$ satisfies the orthogonal projection lemma, expression (70) can be written after lengthy algebraic manipulations (which again are omitted here but can be found in [18])

$$\Upsilon(Z^t, \theta^a) = \int_0^t \mathrm{Tr}_S \{[I_\nu \otimes R_\sigma^{-1}][P^a_{\nabla Y}(\sigma) + \overline{\nabla_a \hat{Y}_\sigma}\ \overline{\nabla_a \hat{Y}_\sigma^T}]\}\, d\sigma \tag{71}$$

where $\otimes$ is the Kronecker product introduced in (5)

$$\overline{\nabla_a \hat{Y}(t)} = E\{\nabla_a \hat{Y}(t)\} \tag{72}$$

and

$$P^a_{\nabla Y}(t) = E\{[\nabla_a \hat{Y}(t) - \overline{\nabla_a \hat{Y}(t)}][\nabla_a \hat{Y}(t) - \overline{\nabla_a \hat{Y}(t)}]^T\} \tag{73}$$

represent, respectively, the mean value and the covariance matrix of the gradient stochastic process $\nabla_a \hat{Y}(t)$.

Fig. 6 shows the evaluation in block diagram of the Fisher information matrix. The block GM generates the first- and the second-order moments of the gradient vector process $\nabla_a \hat{Y}(t)$. To obtain the gradient process $\nabla_a \hat{Y}(t)$ two approaches can be considered: i) Numerical solution: $\nabla_a \hat{Y}(t)$ is computed numerically from $\hat{Y}(t)$; ii) Analytical solution: obtain, from the GKBF model, a dynamical equation for the gradient

$$\nabla_a \hat{X}(t, r) = \left.\frac{\partial \hat{X}(t, r)}{\partial \theta}\right|_{\theta^a}. \tag{74}$$

Then describe the gradient $\nabla_a \hat{Y}(t)$ by means of the state vector $X^a(t, r)$, its estimate $\hat{X}^a(t, r)$, and the gradient $\nabla_a \hat{X}(t, r)$. The superscript a means that the process at which it is applied is taken at the actual parameter vector θ^a. For the analytical approach, it is shown in [18] that $\nabla_a \hat{Y}(t)$ is modeled as the output of a dynamic distributed parameter system with the Gaussian white noise inputs $\{U(t)\}$ and $\{V(t)\}$.

V. SLOT Approximation: Time-Invariant Delay, Stationary Processes, Long Observation Time Interval

A. Introduction

In Sections III and IV, we presented the structure of the ML delay estimator, of the GKBF used to construct the causal MMSE filters for random signals with delays, and derive expressions for the Cramer–Rao bounds for the delays. Here, we show that the ML estimator extends to the nonstationary multisource contexts, the more often en-

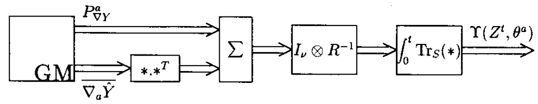


Fig. 6. Fisher information matrix.

countered cross-correlator techniques used in delay estimation when the signals are assumed stationary. We do this by carrying out the asymptotic analysis of the ML estimator of Section III under stationary long observation time interval (SLOT approximation), and time-invariant delay assumptions. A further technical condition needed is that the signal processes are completely controllable and observable [19].

In the above context, the covariance matrix $P(t, r_1, r_2)$ (61)–(64) is asymptotically time invariant, i.e.,

$$\lim_{t \to \infty} P(t, r_1, r_2) = P_\infty(r_1, r_2). \tag{75}$$

The steady-state error covariance matrix $P_\infty(r_1, r_2)$ gives rise to an asymptotically time-invariant generalized Kalman–Bucy filter (IGKBF).

We proceed now by factorizing this asymptotic Riccati equation, showing that the IGKBF corresponds to a generalized Wiener filter, and then interpreting this as the usual cross-correlator structures arising in the delay estimators of, for example, [1]–[3].

B. The Generalized Wiener Filter (GWF)

Denote by $\hat{Y}_I(\omega)$ and $Z_I(\omega)$ the integrated Fourier transforms of the vectorial processes $\hat{Y}(t)$ and $Z(t)$, [17], [22], [23]

$$\hat{Y}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega t}\, d\hat{Y}_I(\omega) \tag{76}$$

and accordingly for $Z(t)$. The IGKBF equivalent frequency domain representation is (see Appendix C)

$$d\hat{Y}_I(\omega) = F_W(\omega)\, dZ_I(\omega) \tag{77}$$

with

$$F_W(\omega) = I - H(\omega, D) \tag{78}$$

where

$$H^{-1}(\omega, D) = I_S + Ce^{-j\omega\gamma}\left[(j\omega I - A)^{-1}K_0 + \int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma\right]. \tag{79}$$

The transfer matrix $F_W(\omega)$ of the IGKBF corresponds to a generalized Wiener filter extended to systems which have delays in the observation process. This is a linear, causal system that, for a stationary signal process $Y(t)$, minimizes the mean-square error between $Y(t)$ and its estimate $\hat{Y}(t)$. We can show (see [18] for details) that the generalized Wiener filter (GWF) (77)–(79) is conceptually equivalent to two cascade filters (Fig. 7). The first

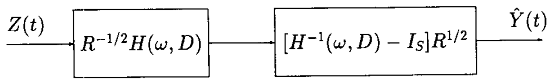


Fig. 7. Generalized Wiener filter.

one, with transfer matrix $R^{-1/2}H(\omega, D)$, is a realizable, minimum phase, whitening filter. The second one, with transfer matrix $[H^{-1}(\omega, D) - I_S]R^{1/2}$, is a realizable filter that gives at its output the signal steady-state MMSE estimate $\hat{Y}(t)$.

C. Log-Likelihood Function Asymptotic Analysis

To be able to compare the general ML processor developed in this paper with the solutions available in the literature, e.g., [1]–[3], we need to express the log-likelihood function (LLF) (48), and the corresponding time delay processor structure, in the frequency domain. These are obtained by working with truncated versions of the processes to a bounded interval $\boldsymbol{T}$. Under the hypothesis of stationarity, assuming $\boldsymbol{T}$ is large compared to the correlation times of the processes plus the maximum delay magnitude $|D|_{\max}$ (SLOT approximation), the estimate of the (truncated version) of the received signal $\hat{Y}_T(t)$ is the output given by the GWF when the input is the (truncated) observed process $Z_T(t)$. Processes $\hat{Y}_T(t)$ and $Z_T(t)$ being limited to T seconds are finite energy processes, which then admit Fourier transforms, $\hat{Y}_T(\omega) = d\hat{Y}_{T_l}(\omega)/d\omega$ and $Z_T(\omega) = dZ_{T_l}(\omega)/d\omega$ [22]. From expression (77), it follows

$$\hat{Y}_T(\omega) = F_W(\omega) Z_T(\omega). \tag{80}$$

As referred to before, the LLF has two parts

$$J(T; D) = J_o(T; D) + J_c(T; D). \tag{81}$$

The first one, $J_o(T; D)$, first and second terms on the RHS of expression (48), depends on the observation process $Z_T(t)$; the second part, $J_c(T; D)$, is the third term on the RHS of (48). From (48), under the SLOT approximation

$$J_c(T; D) = -\tfrac{1}{2} T \operatorname{tr}[R^{-1} P_{Y_\infty}] \tag{82}$$

where P_{Y_∞} is the steady-state signal estimate error covariance matrix, while the first term of the log-likelihood function is

$$J_o(T; D) = \int_{-\infty}^{+\infty} [\langle Z_T(t), \hat{Y}_T(t)\rangle_R - \tfrac{1}{2}\langle \hat{Y}_T(t), \hat{Y}_T(t)\rangle_R]\, dt \tag{83}$$

where the integration over $]-\infty, +\infty[$ is valid given the time limitation imposed to the processes $\hat{Y}_T(t)$ and $Z_T(t)$.

Application of Parseval's theorem to the term $J_o(T; D)$ gives

$$J_o(T; D) = \int_{-\infty}^{+\infty} Z_T^T(-\omega)[R^{-1} - G_{ZZ}^{-1}(\omega)] Z_T(\omega) \frac{d\omega}{4\pi} \tag{84}$$

where, it is shown in Appendix D, the power spectral density of the observation process is

$$G_{ZZ}(\omega) = H^{-1}(\omega, D) R H^{-T}(-\omega, D). \tag{85}$$

The term $J_o(T; D)$, expression (84), is a generalized correlation function that computes the autocorrelation of the prefiltered observation process.

Until now the signals were assumed to be mutually correlated, as well as the observation noise processes. In the next subsection, we are going to consider the simpler case of mutually uncorrelated signals.

D. Mutually Uncorrelated Signals

Assume that the emitted signals $y_l(t)$, $l \in \boldsymbol{L}$, are mutually uncorrelated. In other words, the covariance matrix Q is diagonal. The power spectral density of the observation process is then

$$G_{ZZ}(\omega) = R + \sum_{l=1}^{L} \mathcal{V}_l(\omega) G_l(\omega) \mathcal{V}_l^T(-\omega) \tag{86}$$

where

$$\mathcal{V}_l(\omega) = [\alpha_l^1 e^{-j\omega D_l^1} \,\vdots\, \alpha_l^2 e^{-j\omega D_l^2} \,\vdots\, \cdots \,\vdots\, \alpha_l^S e^{-j\omega D_l^S}]^T \tag{87}$$

is the steering vector, and $G_l(\omega)$ is the power spectral density of the signal $y_l(t)$ radiated by the source l, $l = 1, \cdots, L$.

An ML estimate of the $L \times (S - 1)$ differential delay matrix

$$\Delta D = \begin{bmatrix} \Delta D_1^1 & \Delta D_1^2 & \cdots & \Delta D_1^{S-1} \\ \Delta D_2^1 & \Delta D_2^2 & \cdots & \Delta D_2^{S-1} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta D_L^1 & \Delta D_L^2 & \cdots & \Delta D_L^{S-1} \end{bmatrix}, \tag{88}$$

where the relative delays are

$$\Delta D_l^{s-1} = D_l^s - D_l^{s-1}, \tag{89}$$

is obtained by solving the root equation

$$\nabla J(T; D) = \frac{\partial J(T; D)}{\partial(\Delta D)} = 0 \tag{90}$$

where ∇ is the gradient operator. By differentiation of $J(T; D)$ (expressions (81), (82) and (84)) with respect to ΔD_l^s, $l \in \boldsymbol{L}$, $s \in \boldsymbol{S}$, assuming the signal estimation error is small, we get approximately

$$\frac{\partial J(T; D)}{\partial \Delta D_l^s} \cong \int_0^{+\infty} j\omega\{\Gamma_l^{-1} Z_T^{*T} \mathcal{G}_l^{-1} V_l \Lambda_l^s V_l^{*T} \cdot \mathcal{G}_l^{-1} Z_T - T b_l^s\} \frac{d\omega}{2\pi} \tag{91}$$

where the superscript * stands for complex conjugate

$$V_l(\omega) = \operatorname{diag}\{\alpha_l^s e^{-j\omega D_l^s}\} \tag{92}$$

$$\Lambda_l^s(\omega) = \Phi_s - b_l^s \mathbf{1}_S \mathbf{1}_S^T \tag{93}$$

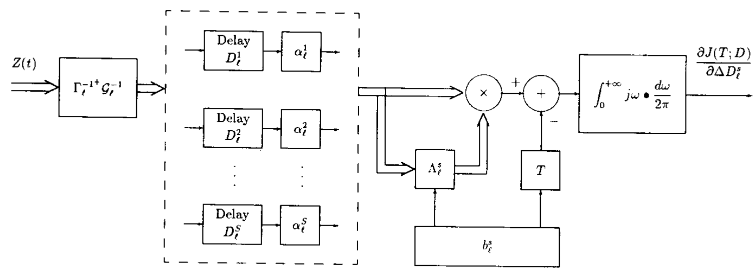


Fig. 8. Generation of $\partial J(T; D)/\partial \Delta D_l^s$: expression (91).

$$\Phi_s = \sum_{m=1}^{S} [-\mathbf{1}_S e_m^T + e_m \mathbf{1}_S^T] \tag{94}$$

$$b_l^s(\omega) = \Gamma_l^{-1} \operatorname{tr} [\mathcal{G}_l^{-1} V_l \Phi_s V_l^{*T}] \tag{95}$$

$$\Gamma_l(\omega) = \mathcal{V}_l^{*T} \mathcal{G}_l^{-1} \mathcal{V}_l + G_l^{-1} \tag{96}$$

$$\mathcal{G}_l(\omega) = R + \sum_{k=1, k \neq l}^{L} \mathcal{V}_k G_k \mathcal{V}_k^{*T} \tag{97}$$

$$e_m = [0 \cdots 0 \;\; \underbrace{1}_{\text{column } m} \;\; 0 \cdots 0]^T \in \mathbb{R}^S \tag{98}$$

and $\mathbf{1}_S$ is the one vector introduced in Section I.

The delay processor assembles $L(S - 1)$ parallel processing channels that perform the derivative $\partial J(T; D)/\partial \Delta D_l^s$. Fig. 8 shows the path (l, s) processing channel block diagram. The causal filter $\Gamma_l^{-1}(\omega)^+ \mathcal{G}_l^{-1}(\omega)$, where $\Gamma_l^{-1}(\omega)^+$ contains all poles and zeros of $\Gamma_l^{-1}(\omega)$ that lie on the left-half s plane (stable, minimum phase), is a function of the delays regarding the first l sources. The term $b_l^s(\omega)$ is a correction term due to either the presence of a number of sources L greater than 1 and/or the spatially correlated observation noise.

For a multisource configuration, under the SLOT approximation with spatially uncorrelated observation noise, i.e., diagonal covariance matrix R, Ng and Bar-Shalom [1] developed a ML processor to estimate the delay matrix ΔD. These authors modeled the observation noises and the signals as zero mean band-limited Gaussian processes. For a white noise assumption it is straightforward to show that the structure of Ng and Bar-Shalom (see [1, eq. (9), (10)], and the processor stated above (see (91)–(98)) are equivalent. The general ML processor presented in this paper (48), (52)–(65) generalizes the estimator of Ng and Bar-Shalom to *nonstationary* mutually correlated signal processes with possibly mutually spatially correlated observation noises.

Assuming a single source configuration ($L = \{1\}$), i.e., taking $l = 1$ in (88)–(98), the ML differential delay estimate is given by the solution of the system of $S - 1$ equations (90), (91), with

$$\mathcal{G}(\omega) = R. \tag{99}$$

The delay receiver involves $S - 1$ parallel processing channels. Each channel block diagram is obtained from the one in Fig. 8, with $\mathcal{G}(\omega)$ defined in (99). For this case, the prefilter $\Gamma^{-1}(\omega)^+ R^{-1}$, although dependent on the delay vector D, becomes a causal system. The term $b^s(\omega)$ is the correction regarding the nonzero observation noise cross covariance. This structure is now equivalent to that of Kirlin and Dewey (see [2, eq. (6)]). A similar argument also recovers the generalized cross correlator of Knapp and Carter [3], when we specialize the structure of (90), (91) to a stationary single source geometry under a spatially uncorrelated observation noise assumption.

VI. Simulation Results

To illustrate the general ML-processor behavior we present an experimental study based on synthetic data. In [18], as well as [24]–[26], several other examples have been described, where we consider the performance of the novel delay estimator when the source signal is nonstationary, in presence of a directional interference, for a two sources configuration, Doppler estimation, and joint estimation of both delay and signal spectrum. Here we do not repeat these experiments that show that the ML estimator developed in this paper achieves asymptotically the Cramer–Rao bound. Rather, we present a comparison with the classical cross-correlator structure when the signals are not stationary. This is, of course, a situation where the classical estimator does not strictly apply.

To implement the delay processor, the parameter domain Θ is restricted to a grid of allowable values. The ML estimator reduces to a bank of parallel processing blocks, each one computing the log-likelihood function (LLF) tuned to each feasible solution $\theta \in \Theta$. This scheme converges uniformly to the optimal algorithm when the mesh of the discretization grid goes to zero.

A Monte Carlo experiment is analyzed next. We carry out the comparison study of our processor with that of Knapp and Carter [3]. For both structures, we evaluate their statistical mean-square error performance, and compare them to that predicted by the Cramer–Rao theory of Section IV.

The array of hydrophones has 2 sensors. The observations are modeled by (14), with

$$\alpha^s = 1, \qquad D^s = (s - 1)D^a, \qquad s = 1, 2 \tag{100}$$

where the actual delay value is $D^a = 0.5$ ms. The delay processor does not know the actual delay D^a, but bounds

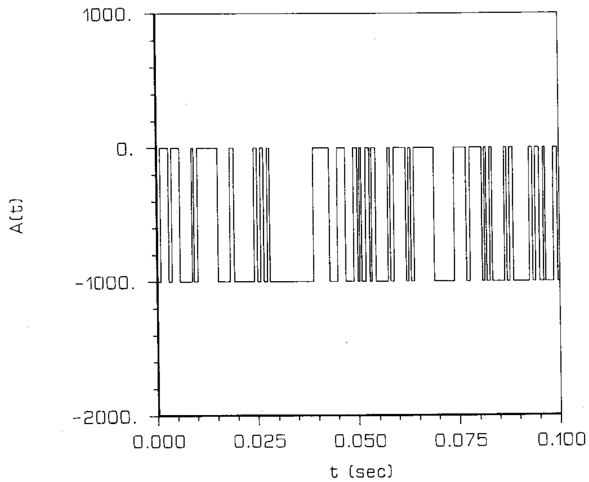


Fig. 9. Time evolution of $A(t)$.

the *a priori* region of uncertainty to

$$D \in [-2.5, +2.5] \quad \text{ms.} \tag{101}$$

This constraint is established from the prior knowledge concerning the problem geometry. The observation noise is spatially and temporally white with spectral height

$$R(s) = 0.001, \quad s = 1, 2. \tag{102}$$

The single signal process $y(t)$ is a zero mean nonstationary process modeled by the linear dynamical system (19), (20) with $A(t)$ shown in Fig. 9, and

$$B = 10, \quad C = 1, \quad Q = 20, \quad \text{cov } x_0 = 1. \tag{103}$$

The time dependence on $A(t)$ may model, for example, a fast maneuvering target subject to sudden changes in its trajectory.

Taking a discretization step of 0.25 ms for the delay domain, both the ML processor developed herein and the generalized cross correlator (GCC) of Knapp and Carter [3] are implemented through a bank of 21 filters working in parallel. We compare both estimators for different observation time interval durations T, where we take

$$T = 5n \quad \text{ms}, \quad n = 1, 2, \cdots, 20. \tag{104}$$

For each value of T, we average the results of 150 Monte Carlo simulations. The time evolution of both the sample mean (SM) and the mean-square error (MSE) of the delay estimate are plotted in Figs. 10 and 11, respectively. From Fig. 10 we see that, for small t, the GCC suffers a larger bias on the delay estimate than the ML processor. Furthermore, as t increases, Fig. 10 shows that the ML estimate is asymptotically unbiased while the GCC retains a residual bias. Fig. 11 shows that the mean-square error (MSE) of the GCC is much larger than the MSE of the ML estimator developed in this paper. In particular, as t increases, Fig. 11 shows that the ML estimate is asymptotically efficient, i.e., its MSE converges to the Cramer–Rao bound (CRB).

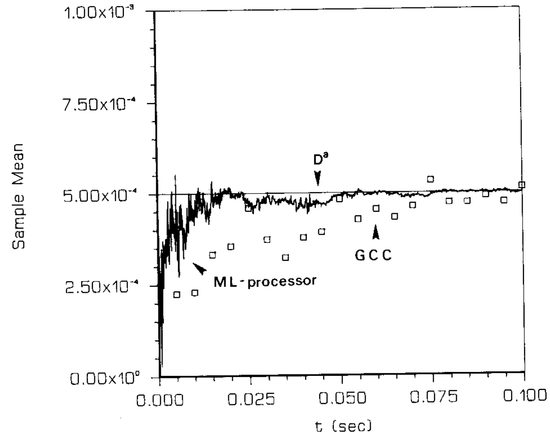


Fig. 10. Delay estimate sample mean.

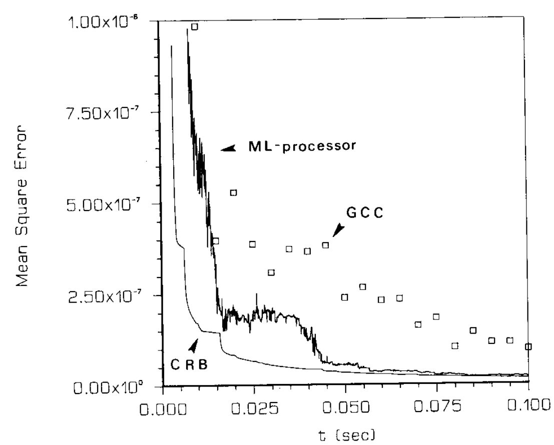


Fig. 11. Delay estimate mean-square error.

VII. Conclusion

The paper reports on maximum likelihood time-varying delay estimation with stochastic nonstationary signals. The framework includes multiple sensors with possibly spatially correlated noise, for either multisources in a single path, or a single source in a multipath environment.

The maximum likelihood (ML) delay estimator maximizes the log-likelihood function (LLF). This function involves estimates of each of the source signals received at each array sensor. To obtain these estimates, we reformulated the nonstationary multisource/multisensor delay problem in the framework of a distributed parameter system model. The signal estimation was now constructed as the causal minimum mean-square error (MMSE) filter for the distributed parameter signal model.

The paper presented the structure of the ML-delay estimator. Due to the nonstationarity assumed, the estimator emphasizes time domain techniques. We provided an interpretation of the delay estimator by studying it under SLOT conditions. We showed that under stationary and long observation intervals, working in the frequency domain, the Riccati equation associated with our generalized Kalman–Bucy filter (GKBF) can be factored, leading to the transfer function of a corresponding generalized Wiener filter (GWF). Using this GWF into the LLF recovers the more common cross-correlator delay estimators of [1]–[3]. The statistical analysis of the delay estimator is carried out in terms of the Cramer–Rao bound. The paper presents an experimental study based on synthetic data that illustrates the behavior of the ML-delay estimator and compares it with the classical generalized cross-correlator configuration.

Appendix A

In this Appendix we derive the LLF for discrete time t. Two steps are considered: in the first one, the model discretization is carried out, while in the second one, the LLF is established.

1. Model Discretization

Define in $[0, t]$ the partition

$$0, \tau, 2\tau, \cdots, N\tau = t \qquad (A.1)$$

with τ small and positive. The discretization of the observation set (45) leads to the observation record

$$Z^t = \{Z(k\tau): k = 0, 1, \cdots, N\} \qquad (A.2)$$

modeled by the algebraic equation (from (33))

$$Z(k\tau) = Y(k\tau) + V(k\tau), \qquad k = 0, 1, \cdots, N \qquad (A.3)$$

where $\{V(k\tau), k = 0, 1, \cdots, N\}$ is a zero mean Gaussian white noise vector sequence, with normalized covariance matrix $R(k\tau)/\tau$.

To the observation record (A.2) corresponds the signal sequence

$$Y^t = \{Y(k\tau): k = 0, 1, \cdots, N\}. \qquad (A.4)$$

Denote the corresponding delay sequence as

$$D^t = \{D(k\tau): k = 0, 1, \cdots, N\}. \qquad (A.5)$$

2. Establishment of the LLF

The log-likelihood function (LLF) is

$$J(t, D^t) = \ln p(Z^t|D^t) \qquad (A.6)$$

with

$$\begin{aligned} p(Z^t|D^t) &= E_{Y^t}[p(Z^t|Y^t; D^t)|D^t\} \\ &= E_{Y^t}[p(Z(t)|Y^t, Z^{t-\tau}, D^t)p(Z^{t-\tau}|Y^t, D^t)|D^t\} \end{aligned} \qquad (A.7)$$

where E_{Y^t} denotes the expectation operation over all paths of the signal sequence Y^t.

From (A.3)

$$p(Z(t)|Y^t, Z^{t-\tau}, D^t) = p(Z(t)|Y(t), D(t)). \qquad (A.8)$$

From (A.7), noting that

$$E_{Y^{t-\tau}}\{p(Z^{t-\tau}|Y^t, D^t)|D^t\} = p(Z^{t-\tau}|Y(t), D^t) \qquad (A.9)$$

we have

$$p(Z^t|D^t) = E_{Y(t)}\{p(Z(t)|Y(t), D(t))p(Z^{t-\tau}|Y(t), D^t)|D^t\} \qquad (A.10)$$

where $E_{Y(t)}$ denotes the expectation operation over all paths of the random variable $Y(t)$. Since

$$p(Z^{t-\tau}|D^t) = p(Z^{t-\tau}|D^{t-\tau}) \qquad (A.11)$$

from Bayes law [22], we can write

$$p(Z^{t-\tau}|Y(t), D^t) = \frac{p(Y(t)|Z^{t-\tau}, D^t)p(Z^{t-\tau}|D^{t-\tau})}{p(Y(t)|D^t)}. \qquad (A.12)$$

Substitution of (A.12) in (A.10) leads to

$$p(Z^t|D^t) = E_{Y(t)}\{p(Z(t)|Y(t), D(t))|Z^{t-\tau}, D^t\}p(Z^{t-\tau}|D^{t-\tau}). \qquad (A.13)$$

From (A.3), we have

$$p(Z(t)|Y(t), D(t)) = \mathfrak{N}\left(Z(t) - Y(t), \frac{R(t)}{\tau}\right) \qquad (A.14)$$

where $\mathfrak{N}(.,.)$ represents the normalized Gauss function. On the other hand, noting that $Y(t)$ is Gaussian

$$p(Y(t)|Z^{t-\tau}, D^t) = \mathfrak{N}(Y(t) - \hat{Y}(t), P_Y(t)) \qquad (A.15)$$

where

$$\hat{Y}(t) = E\{Y(t)|Z^{t-\tau}, D^t\} \qquad (A.16)$$

is the minimum mean-square error (MMSE) estimate of the received signal, and

$$P_Y(t) = E\{[Y(t) - \hat{Y}(t)][Y(t) - \hat{Y}(t)]^T|Z^{t-\tau}, D^t\} \qquad (A.17)$$

is the signal error estimation covariance matrix.

Substitution of (A.14) in (A.13), considering (A.15), leads to

$$p(Z^t|D^t) = \mathfrak{N}\left(Z(t) - \hat{Y}(t), P_Y(t) + \frac{R(t)}{\tau}\right) p(Z^{t-\tau}|D^{t-\tau}). \qquad (A.18)$$

From (A.18), recalling the partition (A.1) considered in the discretization of the time interval $[0, t]$, we obtain

$$\begin{aligned} \ln p(Z^t|D^t) = \sum_{k=0}^{N} \Big\{ &[Z(k\tau) - \hat{Y}(k\tau)]^T[P_Y(k\tau)\tau \\ &+ R(k\tau)]^{-1}[Z(k\tau) - \hat{Y}(k\tau)]\tau \\ &- \frac{1}{2} \ln \Big\{(2\pi)^S \det \left[\frac{R(k\tau)}{\tau}\right] \\ &\cdot \det [I + R^{-1}(k\tau)P_Y(k\tau)\tau]\Big\}\Big\}. \end{aligned} \qquad (A.19)$$

In (A.19), the term $(2\pi)^S \det [R(k\tau)/\tau]$ does not depend on the delay sequence D', its inclusion on the LLF being not necessary. Thus

$$J(t; D') = \sum_{k=0}^{N} \{[Z(k\tau) - \hat{Y}(k\tau)]^T[P_Y(k\tau)\tau + R(k\tau)]^{-1} \cdot [Z(k\tau) - \hat{Y}(k\tau)]\tau - \tfrac{1}{2} \ln \det [I + R^{-1}(k\tau)P_Y(k\tau)\tau]\}. \quad (A.20)$$

Expanding as a Taylor series, around $\tau = 0$, the second term on the RHS of expression (A.20), keeping $k\tau = t_k$ fixed, neglecting the second and higher order terms, leads to the LLF for discrete time

$$J(t; D') = \sum_{k=0}^{N} \{[Z(k\tau) - \hat{Y}(k\tau)]^T[P_Y(k\tau)\tau + R(k\tau)]^{-1} \cdot [Z(k\tau) - \hat{Y}(k\tau)] - \tfrac{1}{2} \operatorname{tr} [R^{-1}(k\tau)P_Y(k\tau)]\}\tau. \quad (A.21)$$

Appendix B

In this Appendix, the log-likelihood function (LLF) for continuous time t is presented.

As referred to in Section II-A, a rigorous description of the processes requires the Ito calculus formulation. Letting formally $Z(t) = d\zeta(t)/dt$ and $V(t) = dW(t)/dt$ in (33), the observed process is modeled by the stochastic differential equation

$$d\zeta(t) = Y(t)\,dt + dW(t) \quad (B.1)$$

where $\{W(t)\}$ is a Wiener vectorial process. Assume that i) the processes $\{Y(t)\}$ and $\{W(t)\}$ are independent; ii) $E\{\int_0^T \langle Y(t), Y(t)\rangle_R\, dt\} < \infty$, where $\langle *, *\rangle_R$ is the weighted inner product operator defined in (10), and $R(t)$ is the covariance matrix of the observation noise.

In the above context, the log-likelihood function (LLF) is nothing but the logarithm of the Radon–Nikodym derivative of the measure induced by the observation process with respect to the Wiener measure. This derivative is (see [27, ch. 7, theorem 7.13, note 3])

$$J(t; D') = \int_0^t \langle \hat{Y}_\sigma, d\zeta_\sigma\rangle_R - \tfrac{1}{2} \int_0^t \langle \hat{Y}_\sigma, \hat{Y}_\sigma\rangle_R\, d\sigma. \quad (B.2)$$

In (B.2), $\hat{Y}(\sigma)$ is the minimum mean-square error (MMSE) estimate of the signal process $Y(\sigma)$ given the observation process and the delay sequence

$$\hat{Y}(\sigma) = E[Y(\sigma) | \zeta_\vartheta, \vartheta \in [0, \sigma], D'(\theta)]. \quad (B.3)$$

The first term on the right-hand side (RHS) of expression (B.2) represents an Ito stochastic integral [16].

When implementing the LLF, one converts the Ito integral into a Stratonovich integral [16]. Denoting the innovations process as

$$d\tilde{\zeta}(t) = d\zeta(t) - \hat{Y}(t)\,dt \quad (B.4)$$

the LLF becomes

$$J(t; D') = \int_0^t \langle \hat{Y}_\sigma, d\tilde{\zeta}_\sigma\rangle_R + \tfrac{1}{2} \int_0^t \langle \hat{Y}_\sigma, \hat{Y}_\sigma\rangle_R\, d\sigma. \quad (B.5)$$

The first term in (B.5), the Ito integral, can now be written [16] as

$$\int_0^t \langle \hat{Y}_\sigma, d\tilde{\zeta}_\sigma\rangle_R = \int_0^t \langle \hat{Y}_\sigma, d\tilde{\zeta}_\sigma\rangle_R - \tfrac{1}{2} (\hat{Y}_t, \tilde{\zeta}_t)_R. \quad (B.6)$$

The integral in the RHS of (B.6) is a Stratonovich integral which is computed using the ordinary rules of calculus. The second term in the RHS of (B.6) is the quadratic variation [16] of the processes $\hat{Y}(t)$ and $\tilde{\zeta}(t)$. Define in $[0, t]$ the partition

$$\tau^n = \{t_0, t_1, \cdots, t_n\} \quad (B.7)$$

where

$$0 = t_0 < t_1 < \cdots < t_n = t. \quad (B.8)$$

The quadratic variation of the processes $\hat{Y}(t)$ and $\tilde{\zeta}(t)$ is defined by [16]

$$(\hat{Y}_t, \tilde{\zeta}_t)_R = \underset{|\tau^n| \to 0}{\text{l.i.m.}} \sum_{k=0}^{n-1} [\hat{Y}(t_{k+1}) - \hat{Y}(t_k)]^T R^{-1} [\tilde{\zeta}(t_{k+1}) - \tilde{\zeta}(t_k)] \quad (B.9)$$

where $\hat{Y}(t)$, the best linear estimate of $Y(t)$ in the mean-square error sense, is

$$\hat{Y}_t = E\{Y_t\} + \int_0^t P_{Y_\sigma} R_\sigma^{-1}\, d\tilde{\zeta}_\sigma. \quad (B.10)$$

Noting that

$$\tilde{\zeta}(t) = \int_0^t d\tilde{\zeta}_\sigma \quad (B.11)$$

and considering (B.10), the quadratic variation becomes [16]

$$(\hat{Y}_t, \tilde{\zeta}_t)_R = \int_0^t \operatorname{tr} [R_\sigma^{-1} P_{Y_\sigma}]\, d\sigma. \quad (B.12)$$

Using formally $Z(t) = d\zeta(t)/dt$ along with (B.6) and (B.12), we obtain the LLF

$$J(t; D') = \int_0^t [\langle \hat{Y}_\sigma, Z_\sigma\rangle_R - \tfrac{1}{2} \langle \hat{Y}_\sigma, \hat{Y}_\sigma\rangle_R - \tfrac{1}{2} \operatorname{tr} (R^{-1}P_Y)]\, d\sigma \quad (B.13)$$

where the integral is computed using the ordinary rules of calculus, and both $\hat{Y}(\sigma)$ and $P_Y(\sigma)$ are functions of the time delay realization $D^\sigma(\theta)$. The third term on the RHS of (B.13) is the correction term regarding the corresponding Ito integral representation.

Appendix C

In this Appendix we derive the frequency domain representation of the time-invariant generalized Kalman–Bucy filter (IGKBF).

Let $\hat{X}_I(\omega, r)$ $(r \in [0, 1])$ and $Z_I(\omega)$ be the integrated Fourier transforms of, respectively, the state vector $\hat{X}(t, r)$ and the observation process $Z(t)$. From (57), (58) with

$$K(t, r) = K(r) = P_\infty(r, 1)C^T R^{-1} \quad \text{(C.1)}$$

the frequency domain representation of the IGKBF is

$$j\omega\gamma\, d\hat{X}_I(\omega, r) + \frac{\partial}{\partial r}[d\hat{X}_I(\omega, r)] = \gamma K(r)[dZ_I(\omega) - C\, d\hat{X}_I(\omega, 1)] \quad \text{(C.2)}$$

with the boundary condition

$$d\hat{X}_I(\omega, 0) = (j\omega I - A)^{-1}K(0)[dZ_I(\omega) - C\, d\hat{X}_I(\omega, 1)]. \quad \text{(C.3)}$$

For $r = 1$ the solution of (C.2) is

$$d\hat{X}_I(\omega, 1) = e^{-j\omega\gamma}\, d\hat{X}_I(\omega, 0) + e^{-j\omega\gamma}\int_0^1 e^{j\omega\gamma\sigma}\gamma K(\sigma) \cdot d\sigma[dZ_I(\omega) - C\, d\hat{X}_I(\omega, 1)]. \quad \text{(C.4)}$$

Substituting (C.3) into (C.4), premultiplying by C, and noting that

$$d\hat{Y}_I(\omega) = C\, d\hat{X}_I(\omega, 1) \quad \text{(C.5)}$$

where $\hat{Y}_I(\omega)$ is the integrated Fourier transform of the signal estimate $\hat{Y}(t)$, we get

$$d\hat{Y}_I(\omega) = [I_S - H(\omega, D)]\, dZ_I(\omega) \quad \text{(C.6)}$$

where

$$H^{-1}(\omega, D) = I_S + Ce^{-j\omega\gamma}\left[(j\omega I - A)^{-1}K(0) + \int_0^1 e^{j\omega\gamma\sigma}\gamma K(\sigma)\, d\sigma\right]. \quad \text{(C.7)}$$

Appendix D

In this Appendix we establish expression (85). The technique carries out a spectral factorization of the steady state (PDE) Riccati equation. To work with a compact notation, the subscript ∞ on the steady state covariance matrix $P_\infty(r_1, r_2)$ will be often omitted.

From (79), we have

$$\begin{aligned} H^{-1}(\omega, D)RH^{-T}(-\omega, D) &= \left\{I_S + Ce^{-j\omega\gamma}\left[(j\omega I - A)^{-1}K_0 + \int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma\right]\right\} R \\ &\quad \cdot \left\{I_S + \left[K_0^T(-j\omega I - A^T)^{-1} + \int_0^1 K_\sigma^T\gamma e^{-j\omega\gamma\sigma}\, d\sigma\right] e^{j\omega\gamma}C^T\right\} \\ &= R + Ce^{-j\omega\gamma}(j\omega I - A)^{-1}K_0R \\ &\quad + Ce^{-j\omega\gamma}\int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma R \\ &\quad + RK_0^T(-j\omega I - A^T)^{-1}e^{j\omega\gamma}C^T \\ &\quad + R\int_0^1 K_\sigma^T\gamma e^{-j\omega\gamma\sigma}\, d\sigma e^{j\omega\gamma}C^T \\ &\quad + Ce^{-j\omega\gamma}\Big[(j\omega I - A)^{-1}K_0RK_0^T(-j\omega I - A^T)^{-1} \\ &\quad + (j\omega I - A)^{-1}K_0R\int_0^1 K_\sigma^T\gamma e^{-j\omega\gamma\sigma}\, d\sigma \\ &\quad + \int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma RK_0^T(-j\omega I - A^T)^{-1} \\ &\quad + \int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma R\int_0^1 K_\sigma^T\gamma e^{-j\omega\gamma\sigma}\, d\sigma\Big] e^{j\omega\gamma}C^T. \end{aligned} \quad \text{(D.1)}$$

Each one of the quadratic terms on the RHS of (D.1) can be obtained from the generalized Riccati equation that describes the steady-state covariance matrix $P_\infty(r_1, r_2)$, $r_1, r_2 \in [0, 1]$. From expressions (61)–(64) with $\partial P(t, r_1, r_2)/\partial t = 0$, $P_\infty(r_1, r_2)$ is given by the partial differential equation

$$\nabla_{r_1r_2}P_\infty(r_1, r_2) = -\gamma K(r_1)RK^T(r_2)\gamma \quad \text{(D.2)}$$

with the boundary conditions

$$\nabla_r P_\infty(r, 0) = \gamma P_\infty(r, 0)A^T - \gamma K(r)RK^T(0) \quad \text{(D.3)}$$

$$\nabla_r P_\infty(0, r) = AP_\infty(0, r)\gamma - K(0)RK^T(r)\gamma \quad \text{(D.4)}$$

and

$$AP_\infty(0, 0) + P_\infty(0, 0)A^T + BQB^T - K(0)RK^T(0) = 0 \quad \text{(D.5)}$$

where the differential operators are

$$\nabla_{r_1r_2}(*) = \left[\frac{\partial}{\partial r_1}(*)\right]\gamma + \gamma\left[\frac{\partial}{\partial r_2}(*)\right] \quad \text{(D.6)}$$

and

$$\nabla_r(*) = \frac{d}{dr}(*). \quad \text{(D.7)}$$

Multiplication of (D.2) on the left by $e^{j\omega\gamma r_1}$ and on the right by $e^{-j\omega\gamma r_2}$, and integration on r_1 and r_2 over the interval $[0, 1]$, leads to

$$\begin{aligned} \int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma\, d\sigma R\int_0^1 K_\sigma^T\gamma e^{-j\omega\gamma\sigma}\, d\sigma &= -e^{j\omega\gamma}\int_0^1 P_{1\sigma}\gamma e^{-j\omega\gamma\sigma}\, d\sigma + \int_0^1 P_{0\sigma}\gamma e^{-j\omega\gamma\sigma}\, d\sigma \\ &\quad - \int_0^1 e^{j\omega\gamma\sigma}\gamma P_{\sigma 1}\, d\sigma e^{-j\omega\gamma} + \int_0^1 e^{j\omega\gamma\sigma}\gamma P_{\sigma 0}\, d\sigma. \end{aligned} \quad \text{(D.8)}$$

Noting that

$$e^{j\omega\gamma r}\nabla_r P_{r0} = \nabla_r(e^{j\omega\gamma r}P_{r0}) - j\omega\gamma e^{j\omega\gamma r}P_{r0} \quad \text{(D.9)}$$

multiplication of (D.3) on the left by $e^{j\omega\gamma r}$, and integration on r over the interval [0, 1], yields after manipulation

$$\int_0^1 e^{j\omega\gamma\sigma}\gamma K_\sigma \, d\sigma R K_0^T(-j\omega I - A^T)^{-1}$$
$$= -e^{j\omega\gamma}P_{10}(-j\omega I - A^T)^{-1}$$
$$+ P_{00}(-j\omega I - A^T)^{-1} - \int_0^1 e^{j\omega\gamma\sigma}\gamma P_{\sigma 0}\, d\sigma. \quad \text{(D.10)}$$

Noting that

$$\nabla_r P_{0r}e^{-j\omega\gamma r} = \nabla_r(P_{0r}e^{-j\omega\gamma r}) + P_{0r}e^{-j\omega\gamma r}j\omega\gamma \quad \text{(D.11)}$$

multiplication of (D.4) on the right by $e^{-j\omega\gamma r}$, and integration on r over the interval [0, 1], gives after manipulation

$$(j\omega I - A)^{-1}K_0 R \int_0^1 K_\sigma^T \gamma e^{-j\omega\gamma\sigma}\, d\sigma$$
$$= -(j\omega I - A)^{-1}P_{01}e^{-j\omega\gamma} + (j\omega I - A)^{-1}P_{00}$$
$$- \int_0^1 P_{0\sigma}\gamma e^{-j\omega\gamma\sigma}\, d\sigma. \quad \text{(D.12)}$$

From (D.5), it is straightforward to show that

$$(j\omega I - A)^{-1}K_0 R K_0^T(-j\omega I - A^T)^{-1}$$
$$= -(j\omega I - A)^{-1}P_{00} - P_{00}(-j\omega I - A^T)^{-1}$$
$$+ (j\omega I - A)^{-1}BQB^T(-j\omega I - A^T)^{-1}. \quad \text{(D.13)}$$

Substitution of (D.8), (D.10), (D.12), and (D.13) in (D.1), remarking that the IGKBF gain matrix is

$$K(r) = P_\infty(r, 1)C^T R^{-1} \quad \text{(D.14)}$$

yields after manipulation

$$H^{-1}(\omega, D)RH^{-T}(-\omega, D) = R + G_{YY}(\omega) \quad \text{(D.15)}$$

where

$$G_{YY}(\omega) = Ce^{-j\omega\gamma}(j\omega I - A)^{-1}BQB^T(-j\omega I - A^T)^{-1}e^{j\omega\gamma}C^T \quad \text{(D.16)}$$

is the power spectral density of the received signal $Y(t)$. The RHS of (D.15) represents the power spectral density $G_{ZZ}(\omega)$ of the observed process $Z(t)$, i.e.,

$$H^{-1}(\omega, D)RH^{-T}(-\omega, D) = G_{ZZ}(\omega). \quad \text{(D.17)}$$

References

[1] 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.

[2] R. L. Kirlin and L. A. Dewey, "Optimal delay estimation in a multiple sensor array having spatially correlated noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1387-1396, Dec. 1985.

[3] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 320-327, Aug. 1976.

[4] G. C. Carter, "Time delay estimation," Ph.D. dissertation, University of Connecticut, Storrs, CT, 1976.

[5] 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, no. 6, pp. 1545-1549, June 1977.

[6] R. L. Kirlin, D. F. Moore, and R. F. Kubichek, "Improvement of delay measurements from sonar arrays via sequential state estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 514-519, June 1981.

[7] J. C. Hassab and R. E. Boucher, "Optimum estimation of time delay by a generalized correlator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 373-380, Aug. 1979.

[8] R. S. Bucy, J. M. F. Moura, and A. J. Mallinckrodt, "A Monte Carlo study of absolute phase determination," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 509-520, July 1983.

[9] J. M. N. Leitão and J. M. F. Moura, "Algorithm structures for cyclic phase estimation," presented at GRETSI-81, Nice, France, 1981.

[10] R. E. Bethel and R. G. Rahikka, "An optimum first-order time delay tracker," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 718-725, Nov. 1987.

[11] J. M. F. Moura and A. B. Baggeroer, "Phase unwrapping of signals propagated under the Arctic ice crust: A statistical approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 617-630, May 1988.

[12] A. Graham, *Kronecker Products and Matrix Calculus: with Applications.* Ellis Horwood Ltd., 1981.

[13] J. M. F. Moura and A. B. Baggeroer, "Passive systems theory with narrow-band and linear constraints: Part I—Spatial diversity," *IEEE J. Ocean. Eng.*, vol. OE-3, pp. 5-13, Jan. 1978.

[14] J. M. F. Moura, "Passive systems theory with narrow-band and linear constraints: Part II—Temporal diversity," *IEEE J. Ocean Eng.*, vol. OE-4, pp. 19-30, Jan. 1979.

[15] J. M. F. Moura, "Passive systems theory with narrow-band and linear constraints: Part III—Spatial/temporal diversity," *IEEE J. Ocean. Eng.*, vol. OE-4, pp. 113-119, July 1979.

[16] J. M. F. Moura, "Linear and nonlinear stochastic filtering," in *Signal Processing*, vol. 1, J. L. Lacoume, T. S. Durrani, and R. Storo, Eds. North-Holland, 1987, pp. 205-276.

[17] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I.* New York: Wiley, 1968.

[18] I. M. G. Lourtie, "Estimação Óptima de Tempos de Atraso com Sinais Não-Estacionários," Ph.D. dissertation, Instituto Superior Técnico, Portugal, Oct. 1987; also available as "Optimal estimation of time delays with nonstationary signals," Tech. Rep., LASIP, CMU, June 1988.

[19] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME J. Basic Eng.*, Mar. 1961.

[20] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 169-173, Apr. 1967.

[21] T. K. Yu, J. H. Seinfeld, and W. H. Ray, "Filtering in nonlinear time delay systems," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 324-333, Aug. 1974.

[22] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 1965.

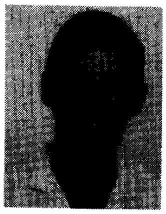
[23] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions.* Englewood Cliffs, NJ: Prentice-Hall, 1962.

[24] I. M. G. Lourtie and J. M. F. Moura, "Time delay determination: Maximum likelihood and Kalman-Bucy type structures," in *Proc. 1985 IEEE Int. Conf. Acoust., Speech, Signal Processing*, Mar. 1985.

[25] I. M. G. Lourtie and J. M. F. Moura, "Optimal estimation of time-varying delay," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Processing*, Apr. 1988.

[26] I. M. G. Lourtie and J. M. F. Moura, "Joint delay and signal deter-

mination," in *Underwater Acoustic Data Processing*, Y. T. Chan, Ed. Kluwer Academic, 1989, pp. 531-535.
[27] R. S. Lipster and A. N. Shiryayev, *Statistics of Random Processes*, vol. I. New York: Springer, 1977.



Isabel M. G. Lourtie (S'85-A'89-M'90) was born in Luanda, Angola, on May 28, 1952. She received the E.E. degree in 1977, and the M.Sc. and Ph.D. degrees in electrical and computer engineering in 1983 and 1988, respectively, all from Instituto Superior Técnico (IST), Lisboa, Portugal.

In 1977, she joined the faculty of the Department of Electrical and Computer Engineering of IST where she has been an Assistant Professor since 1988. She is also a Researcher at Centro de Análise e Processamento de Sinais (CAPS) in Lisboa, Portugal. She has taught courses in systems and signals theory, time series analysis, control systems, network theory and electronics. Her research interests are the applications of signal processing and estimation and system theories, namely, in underwater acoustics. She has presented several papers in this subject area at specialized international meetings.

Dr. Lourtie is a member of the IEEE Signal Processing Society and Ordem dos Engenheiros.

José M. F. Moura (S'71-M'75-SM'90) received the engenheiro electrotécnico degree in 1969 from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and the D.Sc. in electrical engineering and computer science from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1973 and 1975, respectively.

In 1975, he joined the Faculty of IST where he was an Assistant Professor, Professor Agregado (1978), and Professor Catedrático (1979). From 1984 to 1986 he was Genrad Associate Professor of Electrical Engineering and Computer Science (visiting) at M.I.T., also associated with LIDS. Since 1986, he has been a Professor of Electrical and Computer Engineering at Carnegie-Mellon University (CMU), Pittsburgh, PA, where he is also Director of LASIP. He spent the summers of 1978-1981 as a Visiting Scientist with the Department of Aerospace Engineering, University of Southern California (USC), Los Angeles. His research interests lie in statistical signal processing (one and two dimensional), communications, and systems theory. He has had over 80 technical contributions, including invited ones, published in international journals and conference proceedings, and is coeditor of the book *Nonlinear Stochastic Problems* (The Netherlands: Reidel, 1983).

Dr. Moura is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (Sonar and Radar), is a member of several IEEE societies, and of Sigma Xi, AMS, IMS, SIAM, and Ordem dos Engenheiros.