Bearings-Only Tracking for Maneuvering Sources

JEAN-PIERRE LE CADRE, Member, IEEE

OLIVIER TRÉMOIS
IRISA/CNRS
France

Classical bearings-only target-motion analysis (TMA) is restricted to sources with constant motion parameters (usually position and velocity). However, most interesting sources have maneuvering abilities, thus degrading the performance of classical TMA. In the passive sonar context a long-time source-observer encounter is realistic, so the source maneuver possibilities may be important in regard to the source and array baseline. This advocates for the consideration and modeling of the whole source trajectory including source maneuver uncertainty. With that aim, a convenient framework is the hidden Markov model (HMM). A basic idea consists of a two-levels discretization of the state-space. The probabilities of position transition are deduced from the probabilities of velocity transitions which, themselves, are directly related to the source maneuvering capability.

The source state sequence estimation is achieved by means of classical dynamic programming (DP). This approach does not require any prior information relative to the source maneuvers. However, the probabilistic nature of the source trajectory confers a major role to the optimization of the observer maneuvers. This problem is then solved by using the general framework of the Markov decision process (MDP).

Manuscript received December 26, 1995; revised October 20, 1996.

IEEE Log No. T-AES/34/1/00187.

This work was supported by DCN/Ing/Sud (Direction Const. Navales), France.

Authors' current addresses: J.-P. Le Cadre, IRISA/CNRS, Campus de Beaulieu, 35042 Rennes Cedex, France, e-mail: lecadre@irisa.fr; O. Trémois, Thomson-Marconi Sonars, Route de Ste-Anne du Portzic, 29601 Brest Cedex, France, e-mail: OLIVIER.O.T.tremois@tms.thomson.fr.

I. INTRODUCTION

The basic problem of target-motion analysis (TMA) is to estimate the trajectory of an object (or source) from noise-corrupted sensor data. However, for numerous practical applications and especially for long-time scenarios, the source is maneuvering. Tracking methods for maneuvering sources is a classical theme and a huge amount of literature is devoted to this subject [9]. The fields of applications are as varied as radar systems [11], infrared systems [3], and sonar [21].

One of the first attempts to estimate the trajectory of a maneuvering target was presented by Singer [28] where an exponentially correlated model of moving target maneuver was considered. More precisely, the target acceleration was modeled as a stochastic process having an exponential correlation function. However, the suggested Kalman filter was matched to the averaged maneuver so there was no adaptativity with respect to the changes in maneuver. Various authors have tried to improve tracking performance by using time-varying filters (see for instance [23]).

Targets and especially manned targets most of the time have a constant speed vector. At random moments, they perform sudden maneuvers and after resume with uniform motion. The algorithms which estimate trajectory of such targets can be broadly divided into three classes.

The first one is the fixed model approach, where the general approach is to provide the best trajectory smoothing possible, trying to adjust to maneuvers within the constraints of a unique model. The limitations and drawbacks of such methods are evident.

In order to remedy the drawbacks of the unique model, switching model algorithms have been developed. They consist of one Kalman filter tuned to each motion model, and they try to detect model change and activate the correct Kalman filter. It often seems reasonable to consider the statistical properties of the innovation [30] to detect a maneuver occurence. It is also possible to consider general statistical methods for the detection of model changes [5].

The third approach is that of Multiple Models algorithms where all the Kalman filters are active at all times and the process of model change is Markovian with known transition probabilities. The optimal solution to the maneuvering target tracking then requires the consideration of an exponentially growing number of histories. So suboptimal solutions like GPB 1 or GPB 2 have been proposed [12]. For the IMM approach [6] the estimated state consists of the estimated states of the various elementary models, themselves estimated by linear combinations of the previous estimates. These (suboptimal) algorithms do not require the problematic step of maneuver detection since they consider a progressive change of the model.

According to this brief panorama, it is obvious that tracking maneuvering sources has motivated important and fruitful efforts. However, it is worth noting that these efforts are mainly focused on radar system applications [11]. The context of sonar systems is rather different since it is frequently a passive system whose observations (basically the estimated bearings) depend non-linearly on the state. Furthermore, considering a long-time source-observer encounter, the source maneuver possibilities are quite numerous and diverse. A similar remark can be made for the detection of the source maneuvers which requires a suitable estimation of the source state, itself needing a sufficient signal-to-noise ratio and, overall, a sufficient time between consecutive source maneuvers. All these considerations advocate for the consideration and modeling of the whole source trajectory including source maneuver uncertainty.

With that aim, a convenient framework is the hidden Markov model (HMM), widely used in other contexts like speech processing [26], frequency line-tracking [33] and recently in infrared detection and tracking of dim moving sources [4]. A first attempt at using HMM tools in the sonar system area seems to be the works of Martinerie, et al. [21, 22] where the observations were constituted of the measurements of three active buoys. Applications to passive sonar systems appeared in [18, 31, 32]. In order to apply HMM methods to the bearings-only tracking (BOT) context, a basic idea consists of a two-levels discretization of the state-space. The probabilities of position transitions are deduced from the probabilities of velocity transitions. These transitions correspond themselves to the maneuverability constraints inherent to the source. It is thus possible to avoid a too-precise source maneuver modeling to the benefit of a rather coarse stochastic modeling.

General Markovian models of source trajectory are considered in Section II. More precisely, discretization of Markovian modeling are investigated from which the (discretized) probabilities of transition are deduced. A general two-level transition process is thus developed, especially relevant to target-motion analysis (TMA). The observation model is presented in Section III. In the BOT context, the source state is only partially observed through noisy nonlinear measurements. Given a measurement sequence, the estimation problem consists in finding the sequence of states which maximizes a conditional probability density function (pdf). This is achieved by means of classical dynamic programming (DP) algorithm (Viterbi) [10, 16], described in Section IV. This approach is an elegant solution to the maneuvering target tracking problem since it does not require any prior information on the source maneuvers.

Even if the state sequence estimation appears as a direct application of DP, a major problem (specific to BOT) comes from the optimization of the observer trajectory. The problem is then immersed in the general framework of Markov decision process (MDP [2, 13]). In the MDP context, the problem is to determine the control policy optimizing an averaged cost functional. However, our problem differs from the classical one by the nature of the cost functional which is defined as matrix space. The choice of the matrix functional is very critical. Actually, it is required that the matrix functional satisfies a monotonicity property which considerably reduces the possibilities for its choice.

The problem has been considered both from the complete and partial information point of views. For the complete information approach, the sequence of source states is assumed to be known. The aim of this analysis is mainly to provide a catalogue of optimized observer trajectory and to investigate the theoretical problems occuring with the application of the DP principle (see Section V). A general framework based on the calculation of Fisher Information Matrices (FIM) of increasing dimensionality is developed. These matrices are the FIM matrices associated with the estimation of the source state sequence of a randomly maneuvering source. The major contribution is this general framework, the MDP problem itself is solved by means of the classical DP principle.

Practically, the source state is only partially observed through nonlinear measurements. A natural approach is then that of partially observable Markov decision process (POMDP) for which a general framework is presented (Section VI). More precisely, the absence of knowledge of the state leads to replace it by the information vector. The components of this information vector are the state probabilities conditionned on the history of measurements and controls. Actually, it may be shown that this information vector is itself a Markov process. Using updating formulas, the DP algorithm may be rewritten as an extremization of a scalar product. This is the algorithm of Smallwood and Sondik, which is very briefly presented in Section V. For a summary of the related optimization problems and the practical implementations, we refer to [15, 20].

II. BEARINGS-ONLY TMA FOR MANEUVERING SOURCES: A GENERAL FRAMEWORK

The source located at the coordinates (r_{xs}, r_{ys}) moves with a constant velocity (v_{xs}, v_{ys}). The state vectors of the source and the observer are

$$\mathbf{x}_s \triangleq [r_{xs}, r_{ys}, v_{xs}, v_{ys}]^T,$$

$$\mathbf{x}_o \triangleq [r_{xo}, r_{yo}, v_{xo}, v_{yo}]^T.$$

In terms of the relative state vector $\mathbf{x}$, defined by

$$\mathbf{x} = \mathbf{x}_s - \mathbf{x}_o \triangleq [r_x, r_y, v_x, v_y]^T$$

the discrete-time state equation takes the following form:

$$\mathbf{x}_{k+1} = \mathbf{\Phi}(t_k, t_{k+1})\mathbf{x}_k + \mathbf{u}_k + \mathbf{G}\boldsymbol{\omega}_k$$

where

$$\mathbf{\Phi}(t_k, t_{k+1}) = \begin{bmatrix} \mathbf{Id}_2 & \alpha\mathbf{Id}_2 \\ \mathbf{0} & \mathbf{Id}_2 \end{bmatrix}, \quad \text{where}$$
$$\alpha = (t_{k+1} - t_k) = \Delta T, \quad \text{and} \tag{1}$$
$$\mathbf{Id}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In the above definition, the vector $\mathbf{u}_k$ accounts for the effects of the observer accelerations, while ω_k corresponds to a temporally (vectorial) white noise modeling the proper source motion. We thus have

$$\begin{aligned} \mathrm{Cov}\{\omega_k, \omega_{k'}\} &= \delta_{kk'}\mathbf{Q}_k \\ \mathrm{Cov}\{\omega_k, \mathbf{x}_0\} &= \mathbf{0}. \end{aligned} \tag{2}$$

From (2) we see that the system[1] (1) has the Markov property, so that:

$$p(\mathbf{x}_{k+1} \mid \mathbf{x}_k, \ldots, \mathbf{x}_0) = p(\mathbf{x}_{k+1} \mid \mathbf{x}_k). \tag{3}$$

In the BOT context, the measurements are the estimated source bearings $\hat{\theta}_k$ (defined relatively to the North axis):

$$y_k \triangleq \hat{\theta}_k = \theta_k + \nu_k = f(\mathbf{x}_k) + \nu_k$$

with

$$\begin{aligned} \theta_k &= \mathrm{Arctan}\left(\frac{r_x(t_k)}{r_y(t_k)}\right) = f(\mathbf{x}_k), \\ \nu_k &\sim \mathcal{N}(0, \sigma_k^2). \end{aligned} \tag{4}$$

In (4) σ_k^2 represents the variance of the measurement noise and is given by the Woodward's formula[2] [8]. Furthermore, the following independence properties are quite realistic:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{x}_0, \nu_k\} &= \mathbf{0} \\ \mathrm{Cov}\{\omega_k, \nu_k\} &= \mathbf{0}. \end{aligned} \tag{5}$$

In view of (5) the following property is satisfied

$$p(y_k \mid \mathbf{x}_k, \ldots, \mathbf{x}_0, y_{k-1}, \ldots, y_0) = p(y_k \mid \mathbf{x}_k). \tag{6}$$

Now, it remains to specify the matrices $\mathbf{Q}$ and $\mathbf{G}$. Actually these matrices are deduced from the modeling of the source acceleration and various models are considered later.

Let us stress upon the nonlinearity of the measurement functional $f(\mathbf{x}_k)$ so that linear algorithms (Kalman filters) are inappropriate. This leads us to consider the use of DP (Viterbi algorithm) for estimating the sequence of states modeling the source trajectory [13, 16].

This type of algorithm requires a discretization of the state space. Obviously there is a compromise between the discretization size and the computation cost. Each discretized value of the state will be a four-dimensional cell (2D for the position and 2D for the velocity vector). If the source state is in a cell $\mathbf{q}_i$, then its discretized state will be defined by $\mathbf{q}_i$.

We now consider the calculation of the matrices $\mathbf{Q}$ and $\mathbf{G}$ for various modeling of the source accelerations. Classical calculations yield[3] [13]:

$$p(\mathbf{x}_{k+m} \mid \mathbf{x}_k, \mathbf{u}_{k+m-1}, \ldots, \mathbf{u}_k) \sim \mathcal{N}\left(\mathbf{A}^m\mathbf{x}_k + \sum_{j=0}^{m-1}\mathbf{A}^{m-j-1}\mathbf{u}_{k+j}, \mathbf{\Sigma}_{k+m|k}\right)$$

with

$$\begin{aligned} \mathbf{A} &\triangleq \mathbf{\Phi}(t_k, t_{k+1}) \\ \mathbf{\Sigma}_{k+m|k} &\triangleq \mathrm{Cov}\left\{\sum_{j=0}^{m-1}\mathbf{A}^{m-j-1}\boldsymbol{\omega}_{k+j}\right\} \\ &= \sum_{j=0}^{m-1}\mathbf{A}^j\mathbf{G}\mathbf{Q}\mathbf{G}^T\mathbf{A}^{jT}. \end{aligned} \tag{7}$$

Note that in (7) the matrix $\mathbf{Q}_k$ is assumed to be constant. The simpler model is

$$\left| \begin{aligned} v_{x,k+1} &= v_{x,k} + \Delta T\omega_{x,k} \\ v_{y,k+1} &= v_{y,k} + \Delta T\omega_{y,k}. \end{aligned} \right.$$

We thus have

$$\left| \begin{aligned} r_{x,k+1} &= r_{x,k} + \Delta T v_{x,k} + \frac{\Delta T^2}{2}\omega_{x,k} \\ r_{y,k+1} &= r_{y,k} + \Delta T v_{y,k} + \frac{\Delta T^2}{2}\omega_{y,k}. \end{aligned} \right. \tag{8}$$

Assuming $\omega_{x,k}$ and $\omega_{y,k}$ to be independent, the x and y motions may be separately modeled. We then have [19]:

$$\mathbf{x}_k = (r_{x,k}, v_{x,k})^T,$$
$$\mathbf{A} = \begin{pmatrix} 1 & \Delta T \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{G} = \begin{pmatrix} \frac{\Delta T^2}{2} \\ \Delta T \end{pmatrix},$$
$$\mathbf{Q} = \sigma^2$$

and therefore

$$\mathbf{G}\mathbf{Q}\mathbf{G}^T = \sigma^2\mathbf{G}\mathbf{G}^T. \tag{9}$$

[1] δ: Kronecker symbol; p: probability.

[2] $\mathcal{N}$: normal density.

[3] $\mathbf{\Sigma}_{k+m|k}$ is the covariance matrix of the resulting noise at time $k+m$ given the noise from time k to $k+m$.

Elementary calculations then yield[4]

$$p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) \sim \delta\left(v_{x,k+1} + v_{x,k} - \frac{2}{\Delta T}(r_{x,k+1} - r_{x,k})\right) \times \mathcal{N}\left(r_{x,k} + \Delta T v_{x,k}, \frac{\Delta T^4}{4}\sigma^2\right)$$

and recalling the following result (direct consequence of the Bayes rule):

$$\begin{aligned} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) &= p(r_{x,k+1} \mid r_{x,k}, v_{x,k}) \\ &\quad \times p(v_{x,k+1} \mid r_{x,k+1}, r_{x,k}, v_{x,k}) \qquad (10) \\ &= p(v_{x,k+1} \mid r_{x,k}, v_{x,k}) \\ &\quad \times p(r_{x,k+1} \mid v_{x,k+1}, v_{x,k}, r_{x,k}), \\ &= p(v_{x,k+1} \mid v_{x,k}) \\ &\quad \times p(r_{x,k+1} \mid v_{x,k+1}, v_{x,k}, r_{x,k}) \qquad (11) \end{aligned}$$

we finally obtain from direct identification (with (10)):

$$\left| \begin{aligned} & p(r_{x,k+1} \mid \mathbf{x}_k) = \mathcal{N}\left(r_{x,k} + \Delta T v_{x,k}, \frac{\Delta T^4}{4}\sigma^2\right) \\ & p(v_{x,k+1} \mid r_{x,k+1}, \mathbf{x}_k) = \delta\left(v_{x,k+1} + v_{x,k} - \frac{2}{\Delta T}(r_{x,k+1} - r_{x,k})\right). \end{aligned} \right. \qquad (12)$$

The transition probabilities relative to velocities appear rather restrictive. So we now consider a more general model of source motion, more precisely [25]:

$$\left| \begin{aligned} \dot{r}_x &= v_x \\ \dot{v}_x &= \dot{\omega}_x \end{aligned} \right. \qquad (13)$$

where $\dot{\omega}_x$ is a white Gaussian noise (ω is a Wiener–Levy process), and the symbol "˙" denotes the differentiation (versus time).

Discretizing (13) we obtain

$$v_{x,k+1} = v_{x,k} + \xi_{x,k}$$

with

$$\xi_k = \omega_x(t_{k+1}) - \omega_x(t_k), \qquad \text{w.g.n.} \quad \mathcal{N}(0, \Delta T \sigma^2). \qquad (14)$$

Integrating (14) the position equation is obtained, yielding

$$\begin{aligned} r_{x,k+1} &= r_{x,k} + \int_{t_k}^{t_{k+1}} [v_{x,k} + (\omega_x(t) - \omega_x(t_k))]\, dt \\ &= r_{x,k} + \Delta T v_{x,k} + \int_{t_k}^{t_{k+1}} (\omega_x(t) - \omega_x(t_k))\, dt \\ &= r_{x,k} + \Delta T v_{x,k} + \int_{t_k}^{t_{k+1}} (t_{k+1} - t)\, d\omega_x(t). \qquad (15) \end{aligned}$$

[4] δ is the Kronecker symbol.

Finally, the following discretized model has been obtained

$$\left| \begin{aligned} r_{x,k+1} &= r_{x,k} + \Delta T v_{x,k} + \eta_{x,k} \\ v_{x,k+1} &= v_{x,k} + \xi_{x,k} \end{aligned} \right.$$

where

$$\left| \begin{aligned} & \xi_{x,k} \sim \text{w.g.n. } \mathcal{N}(0, \Delta T \sigma^2) \\ & \eta_{x,k} = \int_{t_k}^{t_{k+1}} (t_{k+1} - t)\, d\omega_x(t). \end{aligned} \right. \qquad (16)$$

Using (16) the matrix $\mathbf{Q}$ (eq. (2)) is easily calculated (see Appendix A):

$$\mathbf{Q} \triangleq \mathrm{Cov}\begin{pmatrix} \eta_{x,k} \\ \xi_{x,k} \end{pmatrix} = \sigma^2 \begin{pmatrix} \frac{\Delta T^3}{3} & \frac{\Delta T^2}{2} \\ \frac{\Delta T^2}{2} & \Delta T \end{pmatrix}. \qquad (17)$$

The transition probabilities are then deduced from (16) and (17) by integrating on the corresponding cells the density of the Gaussian vector $(r_{x,k}, v_{x,k})^T = \mathbf{x}_k^T$. Writing (15) in a vectorial form:

$$\mathbf{x}_{k+1} = \mathbf{M}\mathbf{x}_k + \mathbf{w}_k$$

where

$$\mathbf{M} = \begin{pmatrix} 1 & \Delta T \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{w}_k = \begin{pmatrix} \eta_{x,k} \\ \xi_{x,k} \end{pmatrix}, \qquad \mathbf{x}_k = \begin{pmatrix} r_{x,k} \\ v_{x,k} \end{pmatrix} \qquad (18)$$

direct calculations give the transition probabilities (see Appendix B):

$$p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) = p(r_{x,k+1} \mid \mathbf{x}_k, v_{x,k+1}) p(v_{x,k+1} \mid \mathbf{x}_k)$$

where

$$\begin{aligned} & p(r_{x,k+1} \mid \mathbf{x}_k, v_{x,k+1}) \sim \mathcal{N}\left(r_{x,k} + \frac{\Delta T}{2}(v_{x,k} + v_{x,k+1}), \sigma^2 \frac{\Delta T^3}{12}\right) \\ & p(v_{x,k+1} \mid \mathbf{x}_k) \sim \mathcal{N}(v_{x,k}, \Delta T \sigma^2). \end{aligned} \qquad (19)$$

Therefore, fixing a single parameter (σ^2) we are able to calculate the transition matrix of the Markov chain defined by (18) where the vector $\mathbf{x}_k$ is associated with the center of a cell. The transition probabilities are then calculated by integrating the pdf defined by (19) on the volume cells $\mathbf{q}_i$ and $\mathbf{q}_j$:

$$\Pr\{\mathbf{x}_{k+1} \in \mathbf{q}_j \mid \mathbf{x}_k \in \mathbf{q}_i\} = \frac{\int_{\mathbf{q}_j}\int_{\mathbf{q}_i} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) p(\mathbf{x}_k)\, d\mathbf{x}_k\, d\mathbf{x}_{k+1}}{\int_{\mathbf{q}_i} p(\mathbf{x}_k)\, d\mathbf{x}_k}. \qquad (20)$$

For the sake of computation time (Viterbi algorithm), transition probabilities under a certain (relative) threshold are considered null. The transition matrix then becomes essentially sparse. This calculation of the transition matrix is illustrated by Fig. 1, which depicts the set of transitions from the origin. Each vertex of the grids, corresponds to

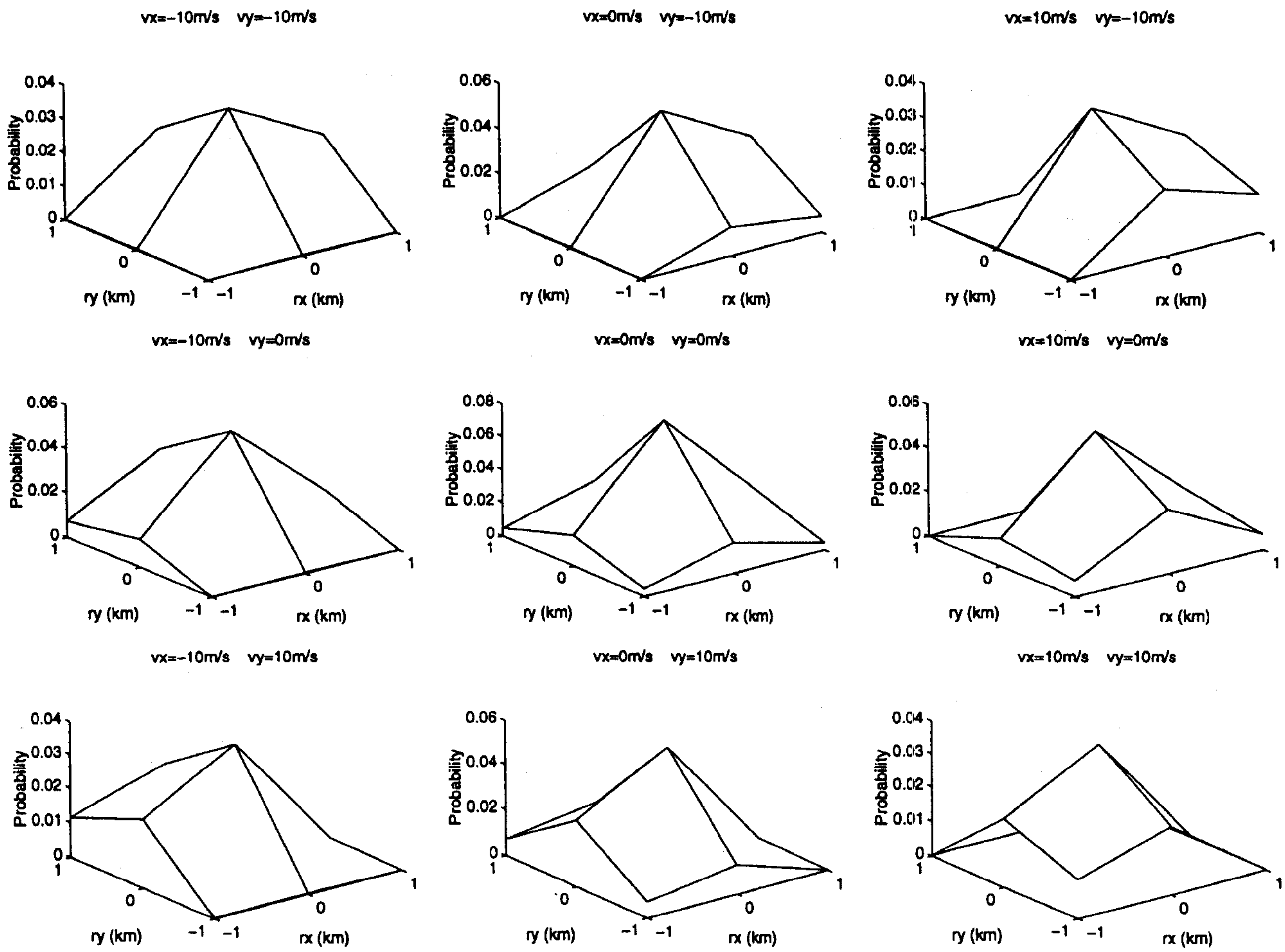


Fig. 1. Transition matrix from origin. Each graph corresponds to position transition for a fixed speed.

a cell. For this example, the width of the position cells is 1 km, the width of the speed cells is 10 m/s and the sampling time is 100 s. The variance of the Wiener–Levy process has been set to 1. This leads to

$$\mathbf{Q} = \begin{pmatrix} \frac{10^6}{3} & \frac{10^4}{2} \\ \frac{10^4}{2} & 10^2 \end{pmatrix}.$$

Under certain conditions (e.g., cell size) the pdf $p(\mathbf{x}_k)$ may be considered as approximately constant on the cell $\mathbf{q}_i$ so that the transition probability becomes

$$\Pr\{\mathbf{x}_{k+1} \in \mathbf{q}_j \mid \mathbf{x}_k \in \mathbf{q}_i\} = \frac{\int_{\mathbf{q}_j}\int_{\mathbf{q}_i} p(\mathbf{x}_{k+1} \mid \mathbf{x}_k)\, d\mathbf{x}_k\, d\mathbf{x}_{k+1}}{\mathrm{Vol}(\mathbf{q}_i)}. \tag{21}$$

($\mathrm{Vol}(\mathbf{q}_i)$ is the volume of the cell $\mathbf{q}_i$.)

As the transitions on x and y are independent, we obtain using Bayes rule:

$$\Pr\{\mathbf{x}_{k+1} \mid \mathbf{x}_k\} = \Pr\{\mathbf{r}_{k+1} \mid \mathbf{r}_k, \mathbf{v}_k\}\Pr\{\mathbf{v}_{k+1} \mid \mathbf{v}_k, \mathbf{r}_{k+1}, \mathbf{r}_k\}. \tag{22}$$

In order to simplify this model, one can consider that the transitions probabilities relative to position and velocity are independent, so that (22) yields

$$\Pr\{\mathbf{x}_{k+1} \mid \mathbf{x}_k\} = \Pr\{\mathbf{r}_{k+1} \mid \mathbf{r}_k, \mathbf{v}_k\}\Pr\{\mathbf{v}_{k+1} \mid \mathbf{v}_k\}. \tag{23}$$

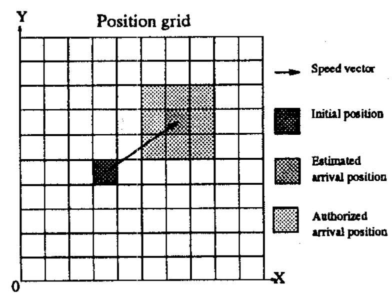


Fig. 2. Position transition process in two-level point of view.

In that way, we have a two-level transition process which is depicted in Figs. 2 and 3. The transition probability on the position which is centered on the estimated position ($\mathbf{r}_{t+1} = \mathbf{r}_t + \mathbf{v}_t$), and, independently, the velocity is authorized to change smoothly.

III. OBSERVATION MODEL

In the passive sonar (BOT) context, the observations are estimated bearings (denoted y_k). The

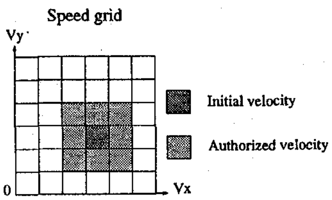


Fig. 3. Velocity transition process in two-level point of view.

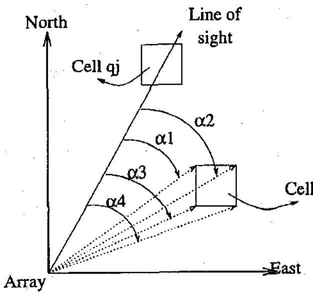


Fig. 4. Different types of cells in observation process.

conditional pdf of the observations is given by

$$\Pr\{y_k \mid \mathbf{x}_k \in \mathbf{q}_i\} = \frac{\int_{\mathbf{q}_i} p(y_k \mid \mathbf{x}_k) p(\mathbf{x}_k)\, d\mathbf{x}_k}{\int_{\mathbf{q}_i} p(\mathbf{x}_k)\, d\mathbf{x}_k}$$

or more simply, if we assume that $p(\mathbf{x}_k)$ is constant on $\mathbf{q}_i$:

$$\Pr\{y_k \mid \mathbf{x}_k \in \mathbf{q}_i\} = \frac{\int_{\mathbf{q}_i} p(y_k \mid \mathbf{x}_k)\, d\mathbf{x}_k}{\mathrm{Vol}(\mathbf{q}_i)}. \tag{24}$$

For the sake of computation time and memory, this calculation may be replaced by a (slightly) less rigorous one:

$$\Pr\{y_k \mid \mathbf{x}_k \in \mathbf{q}_i\} = \frac{\Pr\{\mathbf{x}_k \in \mathbf{q}_i \mid y_k\}\Pr\{y_k\}}{\Pr\{\mathbf{x}_k \in \mathbf{q}_i\}}. \tag{25}$$

In the absence of any prior information about the probabilities $\Pr\{y_k\}$ and $\Pr\{\mathbf{x}_k \in \mathbf{q}_i\}$, it is reasonable to consider them as constant, so that

$$\Pr\{y_k \mid \mathbf{x}_k \in \mathbf{q}_i\} \sim \Pr\{\mathbf{x}_k \in \mathbf{q}_i \mid y_k\}. \tag{26}$$

The density $\Pr\{\mathbf{x}_k \in \mathbf{q}_i \mid y_k\}$ itself is calculated by considering the relative positions (Cartesian coordinates) of the cell $\mathbf{q}_i$ on the one hand and of the line of sight issued from the array center and associated with y_k on the other hand.

Fig. 4 depicts one way to compute $\Pr\{\mathbf{x}_k \in \mathbf{q}_i \mid y_k\}$ taking the maximum of the pdf in the cell. Cell $\mathbf{q}_j$ is intersected by the line of sight, so its coefficient will be set to 1. Cell $\mathbf{q}_i$ is not intersected by the line of sight, so its coefficient will be set to the maximum of the exponential part of the pdf which is found on the corner which has the lowest bearing deviation from the line of sight (the first one in the example). Its probability is set to $\exp[-\frac{1}{2}(\alpha_1/\sigma_\theta)^2]$. After that, all the coefficients are multiplied by a constant so as to have the sum equal to one.

In that case the discrimination between the cells is done only on the position space. If multiple bearings is used to compute a global observation, different speed for a same position will yield different probabilities.

IV. ESTIMATION OF SOURCE TRAJECTORY

We assume that the source trajectory may be modeled by a Markov chain with known transition probabilities. The observations are the estimated bearings y_k ($y_k = \hat{\theta}_k$). Note that y_k is a scalar nonlinear function of the 4-dimensionnal state $\mathbf{x}_k$. The estimation problem we deal with is posed as follows.

Larson–Peschon (LP) problem [16]: Given a measurement sequence $\mathcal{Y}_k = \{y_1, y_2, \ldots, y_k\}$ find the sequence $\mathcal{X}_k$ of states that maximizes the conditional pdf:

$$p(\mathcal{X}_k \mid \mathcal{Y}_k). \tag{27}$$

A sequence of states $\mathcal{X}_k$ is called a trajectory. The trajectory corresponding to a solution of the LP problem is denoted

$$\hat{\mathcal{X}}_{k|k} \triangleq \{\hat{\mathbf{x}}_{0|k}, \hat{\mathbf{x}}_{1|k}, \ldots, \hat{\mathbf{x}}_{k|k}\}.$$

We thus have

$$p(\mathcal{X}_{k|k} \mid \mathcal{Y}_k) = \max_{\mathcal{X}_k} p(\mathcal{X}_k \mid \mathcal{Y}_k). \tag{28}$$

Using Bayes rule, we obtain directly

$$p(\mathcal{X}_{k+1} \mid \mathcal{Y}_{k+1}) = \frac{p(y_{k+1} \mid \mathcal{X}_{k+1}, \mathcal{Y}_k) p(\mathcal{X}_{k+1} \mid \mathcal{Y}_k)}{p(y_{k+1} \mid \mathcal{Y}_k)}. \tag{29}$$

Invoking the Markov hypothesis, we have

$$p(y_{k+1} \mid \mathcal{X}_{k+1}, \mathcal{Y}_k) = p(y_{k+1} \mid \mathbf{x}_{k+1}). \tag{30}$$

The next step consists in simplifying $p(\mathcal{X}_{k+1} \mid \mathcal{Y}_k)$ thanks to the Bayes rule and the Markov property:

$$\begin{aligned} p(\mathcal{X}_k, \mathbf{x}_{k+1} \mid \mathcal{Y}_k) &= p(\mathbf{x}_{k+1} \mid \mathcal{X}_k, \mathcal{Y}_k) \frac{p(\mathcal{X}_k, \mathcal{Y}_k)}{p(\mathcal{Y}_k)} \\ &= p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) p(\mathcal{X}_k \mid \mathcal{Y}_k). \end{aligned} \tag{31}$$

Combining (29) to (31), the LP following recursion is obtained

$$p(\mathcal{X}_{k+1} \mid \mathcal{Y}_{k+1}) = \left(\frac{p(y_{k+1} \mid \mathbf{x}_{k+1}) p(\mathbf{x}_{k+1} \mid \mathbf{x}_k)}{p(y_{k+1} \mid y_k)} \right) p(\mathcal{X}_k \mid \mathcal{Y}_k) \tag{32}$$

and the DP recursion is directly deduced:

$$I^*(\mathbf{x}_{k+1}) = \max_{\mathbf{x}_k} \{ p(y_{k+1} \mid \mathbf{x}_{k+1}) p(\mathbf{x}_{k+1} \mid \mathbf{x}_k) I^*(\mathbf{x}_k) \}$$

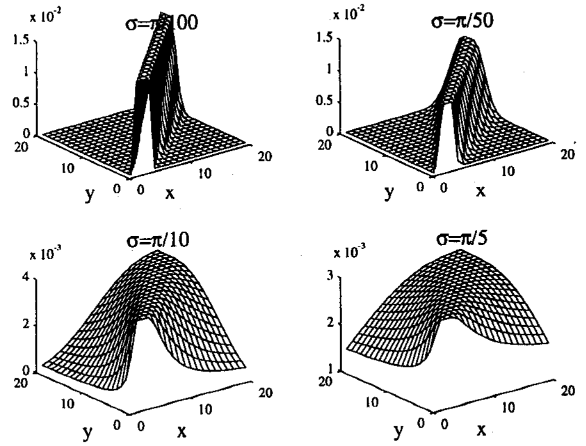


Fig. 5. Probabilities in position space for various variances of estimated bearings.

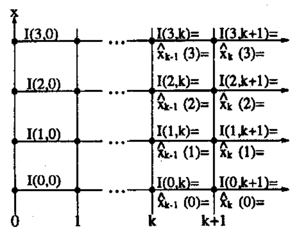


Fig. 6. Computational procedure to implement DP recursion.

$$\hat{\mathbf{x}}_k(\mathbf{x}_{k+1}) = \underset{\mathbf{x}_k}{\operatorname{argmax}}\{p(y_{k+1} \mid \mathbf{x}_{k+1})p(\mathbf{x}_{k+1} \mid \mathbf{x}_k)I^*(\mathbf{x}_k)\}$$

where

$$I^*(\mathbf{x}_k) \sim \max_{\mathcal{X}_k} p(\mathcal{X}_k \mid \mathcal{Y}_k). \tag{33}$$

The state maximizing $I^*(\mathbf{x}_k)$ is denoted $\hat{\mathbf{x}}_{k|k}$, the rest of the trajectory is found by tracing backward a stage at a time, formally (eq. (33)):

$$\hat{\mathbf{x}}_{i|k} = \hat{\mathbf{x}}_i(\hat{\mathbf{x}}_{i+1|k})$$

the procedure being initialized by $\hat{\mathbf{x}}_{k|k}$. It is illustrated in Figs. 5 and 6.

A particular difficulty comes from the discretization. Actually, for computation time saving, it is necessary to choose a rather coarse grid but the occupation time of a cell then becomes very inferior to the discretization time. The DP algorithm then performs poorly. A first remedy consists of increasing the discretization time and therefore reducing the number of observations. This approach presents evident drawbacks. So, a natural way to overcome these difficulties consists in considering a vectorial observation formed of adjacent observations. In the simpler modeling, the conditional pdfs of the observations are calculated by replacing the components of the observation vector by their centroid.

The LP recursion requires that the transition probabilities $p(\mathbf{x}_{k+1} \mid \mathbf{x}_k)$ adequately describe the reality. However, the algorithm appears quite tolerant with respect to (wrt) the parameters of the Markov model. This seems essentially due to the generality of the two-levels Markov model.

V. OPTIMIZATION OF OBSERVER MOTION

Since the performance of the source trajectory estimation is tightly related to the observer motion, this step is instrumental. This problem is immersed in the MDP framework.

The general aim of the MDP is to determine a sequence of decisions, generally denoted d (the observer maneuver in our context), which maximizes a criterion related to the states of the Markov chain. If the process is in state i at time k and if a control d is chosen, then a two step process is launched [13, 27]:

1) a cost $c(i,d)$ is incurred,

2) the next state of the system is chosen according to the transition $p_{ij}(d)$.

If we let $\mathbf{x}_k$ denote the state of the process at time t_k and d_k the decision chosen at t_k, the assumption (2) is equivalent to state that

$$\Pr\{\mathbf{x}_{k+1} = j \mid \mathbf{x}_0, d_0, \ldots, \mathbf{x}_k = i, d_k = d\} = p_{ij}(d). \tag{34}$$

Thus both elementary costs $c(i,d)$ and the transition probabilities $p_{ij}(d)$ are functions only of the last state and of the subsequent control. It is easily shown that if a stationary policy π is employed, the sequence of states $\{\mathbf{x}_k,\ k = 0,1,\ldots\}$ forms a Markov chain with transition probabilities $p_{ij}(d) \triangleq p_{ij}(\pi(i))$, giving thus the denomination MDP to the process.

We denote $d_k = g_k(\mathbf{x}_k)$ the control process associated with policy π. The problem is to find an optimal control process d_k which minimizes a cost functional. Classicaly, this cost functional is defined from elementary cost functionals $c_k(\mathbf{x}_k, d_k)$.

This functional represents the cost to pay if, at time t_k, the state of the process is $\mathbf{x}_k$ and the control is d_k. On a finite horizon T, the following cost functional is considered:

$$\begin{aligned} J(g) &= \mathbb{E}(g)c_T(\mathbf{x}_T) + \sum_{k=0}^{T-1} c_k(\mathbf{x}_k, g(\mathbf{x}_k)) \\ &= \sum_{\mathbf{x}\in\mathcal{S}} \Pr\{\mathbf{x}_T = \mathbf{x}\}c_T(\mathbf{x}) \\ &\quad + \sum_{k=0}^{T-1}\sum_{\mathbf{x}\in\mathcal{S}} \Pr\{\mathbf{x}_k = \mathbf{x}\}c_k(\mathbf{x}, g(\mathbf{x})) \end{aligned} \tag{35}$$

where $\mathcal{S}$ is the finite set of states.

The MDP problem consists of finding the control g^* which minimizes (35).

Practically, (35) is solved by a computationally efficient method based on the optimality principle (due to Bellman [14]). This principle simply asserts that if $\pi^* = \{u_0^*, u_1^*, \ldots, u_{T-1}^*\}$ is an optimal control law then the truncated control law $\{u_i^*, \ldots, u_{T-1}^*\}$ is optimal for the ith truncated problem.

Using the optimality principle, the MDP problem may be solved by the following algorithm (called DP algorithm):

$$J_k^*(i) = \min_{d \in \mathcal{C}} \sum_{j \in \mathcal{S}} [c_{ij}(d) + J_{k+1}^*(j)] P_{ij}(d). \qquad (36)$$

The optimal control law is obtained as the solution of

$$g^*(i) = \operatorname*{argmin}_{d \in \mathcal{C}} \sum_{j \in \mathcal{S}} [c_{ij}(d) + J_{k+1}^*(j)] P_{ij}(d) \qquad (37)$$

initialized by $J^*(T,i) = 0$ $(i \in \mathcal{S})$.

This is the classical approach of DP. However our problem, which consists in optimizing the observer trajectory, may differ from it by the nature of the cost functional. It is a functional of the FIM, indeed the FIM is a convenient measure of estimability.

Various choices of the FIM functional have been considered in the literature. So, to be more precise, the cost functional (35) must be replaced[5] by

$$J(g) = H\left[\mathbb{E}_g\left\{\sum_{k=0}^{T-1} \mathbf{F}_k(\mathbf{x}_k, \mathbf{x}_{k+1}, g(\mathbf{x}_k)) + \mathbf{F}_T(\mathbf{x}_T)\right\}\right] \qquad (38)$$

where H is a matrix functional, and $\mathbf{F}_k(\mathbf{x}_k, \mathbf{x}_{k+1}, g(\mathbf{x}_k))$ is the instantaneous FIM associated with the transition from state $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ and the control $g(\mathbf{x}_k)$.

The choice of H is critical since the control policy depends on it. Actually, it can be shown [17] that a necessary condition for applying the optimality principle (37) is that H satisfies the following condition.

Monotonicity Property. Let $\mathbf{A}$ and $\mathbf{B}$ (in $\mathcal{H}_n$)[6] be two matrices, then H ($\mathcal{H}_n \to \mathbb{R}$, differentiable) owns the DP property if the following monotonicity property holds.

Let $\mathbf{A}$ and $\mathbf{B}$ in $\mathcal{H}_n$ s.t. $H(\mathbf{A}) > H(\mathbf{B})$ then whatever the matrix $\mathbf{C}$ in $\mathcal{H}_n : H(\mathbf{A}+\mathbf{C}) > H(\mathbf{B}+\mathbf{C})$.

It has been shown that the functionals H having the monotonicity property may be written as

$$H(\mathbf{A}) = g(\mathrm{Tr}(\mathbf{AR}))$$

where g is a monotone real function, Tr is the matrix trace and $\mathbf{R}$ is a fixed matrix.

It remains now to consider more precisely the structure of the FIM. The partial derivatives of the exact bearing θ_k wrt $\mathbf{x}_0$ the state vector, are easily obtained (rectilinear and uniform motion), yielding the following gradient vector $\mathbf{g}_k$:

$$\mathbf{g}_k = \left(\frac{\cos\theta_k}{r_k}, -\frac{\sin\theta_k}{r_k}, \frac{k\cos\theta_k}{r_k}, -\frac{k\sin\theta_k}{r_k}\right)^T \qquad (39)$$

$$(\mathbf{x}_0 = (r_x(0), r_y(0), v_x, v_y)^T).$$

Under the rectilinear motion assumption, the elementary FIM $\mathbf{F}_m$ is directly deduced from (39), i.e.,

$$\mathbf{F}_n = \sum_{k=1}^{n} \mathbf{g}_k \mathbf{g}_k^T = \begin{pmatrix} \sum_{k=1}^{n} \boldsymbol{\Omega}_k & \sum_{k=1}^{n} k\Delta T \boldsymbol{\Omega}_k \\ \sum_{k=1}^{n} k\Delta T \boldsymbol{\Omega}_k & \sum_{k=1}^{n} k^2 \Delta T^2 \boldsymbol{\Omega}_k \end{pmatrix}$$

with

$$\boldsymbol{\Omega}_k = \frac{1}{r_k^2 \sigma_{\theta_k}^2} \begin{pmatrix} \cos^2\theta_k & -\sin\theta_k\cos\theta_k \\ -\sin\theta_k\cos\theta_k & \sin^2\theta_k \end{pmatrix}. \qquad (40)$$

This structure of the FIM is quite remarkable and may be easily extended to the case of a maneuvering source. For that purpose, the source trajectory is modeled as a multileg one which is quite coherent with our discrete modeling of the source trajectory. Consequently, for this approach, the dimension of the state vector is enlarged for each new (from the reference time) leg since it includes now the initial position of the source as well as its various velocity vectors (associated with the successive legs).

To be more precise, consider a source trajectory formed of n legs (all of the same length), each one corresponding to J bearings, the complete FIM (denoted $\mathbf{F}_{1,nJ}$) takes the following form:

$$\begin{aligned} \mathbf{F}_{1,nJ} &= \mathbf{F}_{1,J} + \mathbf{F}_{J+1,2J} + \cdots + \mathbf{F}_{(n-1)J,nJ} \\ &= \sum_{k=1}^{J} [\mathbf{d}_{0,1+n}(k)\mathbf{d}'_{0,1+n}(k)] \otimes \boldsymbol{\Omega}_k \\ &\quad + \sum_{k=J+1}^{2J} [\mathbf{d}_{1,1+n}(k)\mathbf{d}'_{1,1+n}(k)] \otimes \boldsymbol{\Omega}_k + \cdots \\ &\quad + \sum_{k=(n-1)J+1}^{nJ} [\mathbf{d}_{n-1,1+n}(k)\mathbf{d}'_{n-1,1+n}(k)] \otimes \boldsymbol{\Omega}_k \\ &= \sum_{m=1}^{n} \sum_{k=(m-1)J+1}^{mJ} [\mathbf{d}_{m-1,1+n}(k)\mathbf{d}'_{m-1,1+n}(k)] \otimes \boldsymbol{\Omega}_k \end{aligned}$$

with

$$\mathbf{d}_{p,q}(k) = \underbrace{(1\ \overbrace{J\Delta T \cdots J\Delta T}^{p}\ (k-pJ)\Delta T\ \overbrace{0\cdots 0}^{q-p-2})'}_{q} \qquad (41)$$

and $\boldsymbol{\Omega}_k$ is given by (40), $\otimes$ is the Kronecker product.

[5] E: expectation.

[6] $\mathcal{H}_n$ is space of Hermitian matrices.

Consider now the calculation of the FIM $\mathbf{F}_k$ with:

$$lJ < k \leq (l+1)J \quad (l \text{ index of the leg})$$

then quite similarly to (41) we have

$$\mathbf{F}_k = [\mathbf{d}_{l-1,l+1}(k)\mathbf{d}_{l-1,l+1}^T(k)] \otimes \mathbf{\Omega}_k \tag{42}$$

and the trace is:

$$\begin{aligned}\mathrm{Tr}(\mathbf{F}_k) &= [\mathbf{d}_{l-1,l+1}^T(k)\mathbf{d}_{l-1,l+1}(k)] \cdot \mathrm{Tr}(\mathbf{\Omega}_k) \\ &= [\mathbf{d}_{l-1,l+1}^T(k)\mathbf{d}_{l-1,l+1}(k)].\end{aligned} \tag{43}$$

For this cost functional, the DP algorithm takes the following form:

Initialization:
For $i = 1,\ldots,N$

$$J_T^*(i) = 0.$$

Recursion:
For $k = T-1,\ldots,0$ and for $i = 1,\ldots,N$ do:

$$J_k^*(i) = \max_{d_k \in D} \sum_j p_{ij}(d_k)(\mathrm{Tr}[\mathbf{F}_k(i,d_k,j)] + J_{k+1}^*(j))$$

$$g_k^*(i) = \operatorname*{argmax}_{d_k \in D} \sum_j p_{ij}(d_k)(\mathrm{Tr}[\mathbf{F}_k(i,d_k,j)] + J_{k+1}^*(j)).$$

$\mathbf{F}_t(i,d_k,j)$ is the FIM associated to the fact that the source comes from the state i to the state j while the observer takes the decision d.

We noticed at the end of Section IV that for computation time problem, the sampling time had to be increased. This technique leaves numerous unprocessed data which can be used to compute the trace of the "instantaneous" FIM. Suppose that only one sample over N_s is used in the estimation process, here are few notations:

δt is the sampling time (of the basic system),
$\Delta T = N_s \delta t$ is the time between two steps of the estimation process (DP algorithm),
J is the Number of DP steps on one leg,
l is the number of the current leg $l = 1,2,\ldots,$
$r(t)$ is the range of the source at time t,
$\sigma_{\theta(t)}$ is the standard deviation of bearing estimation at time t.

With these notations, the trace of the FIM associated to the transition between states i and j under command d_t takes the following form:

$$\mathrm{Tr}[\mathbf{F}_t(i,d_t,j)] = \sum_{\tau=0}^{N_s-1} \frac{1 + (l-1)J^2\Delta T^2 + \left(t + \dfrac{\tau}{N_s} - (l-1)J\right)^2 \Delta T^2}{r^2\left(t + \dfrac{\tau}{N_s}\right)\sigma^2_{\theta(t+(\tau/N_s))}}. \tag{44}$$

The range of the source depends directly on i, j, and d_t. $\sigma^2_{\theta(t+(\tau/N_s))}$ depends not only on command d_t but also on the observer heading and on source bearing. To compute the trace we have to be able to compute the heading of the observer from the system state. That is why a new system state had been chosen, including the relative position of the source, the velocity of the source and the one of the observer. Furthermore, the decisions act directly on the system state via the observer velocity.

However, in practical situations, the source state is not directly observable. The only available information consist of estimated bearings. The MDP problem then becomes far more complicated. A general approach consists in using the general framework of POMDP.

Let us recall now the basics of POMDP. The central process (the Markov chain) $\mathbf{x}_k$ is not directly observable. An observation θ_k (here the bearing) is associated with $\mathbf{x}_k$. Let $\Pi(\mathbf{x}) = \{\pi \in \mathbb{R}^n \mid \pi \geq 0, \sum_{i=1}^N \pi_i = 1\}$ be the set of all the distribution on $\mathbf{x}$ and $H_k = \{\pi(1), d_1, \theta_1, \ldots, d_{k-1}, \theta_{k-1}\}$ ($\theta \in \theta$: observation, $d \in \mathcal{D}$: decision or action, $\mathcal{D}$ finite set) the "history" of the decisions and observations up to time k.

Each vector $\boldsymbol{\pi}$ is a probability distribution on $\mathbf{x}$ (probabilistic interpretation) and is also geometrically represented by a point of the associated unit simplex $\Pi(\mathbf{x})$.

At time k, H_k contains all the information available to the decision maker. The global system evolution is then modeled as follows:

1) Decision, knowing the history of the decisions (controls) and of the observations (i.e., H_k), decision d_k is taken, the system transits from state $\mathbf{x}_k$ to state gx_{k+1} with a given probability of transition p_{ij}^d:

$$p_{ij}^d \triangleq \Pr[\mathbf{x}_{k+1} = j \mid \mathbf{x}_k = i, d_k = d].$$

2) An observation $\theta_k \in \theta$ is received accordingly with the probability $r_{j\theta}^d$:

$$r_{j\theta}^d \triangleq \Pr[\theta_k = \theta \mid \mathbf{x}_{k+1} = j, d_k = d].$$

3) The information vector is updated with two new data: $H_{k+1} = H_k \cup \{d_k, \theta_k\}$.

4) Elementary cost $w_{ij\theta}^d$ is the immediate reward associated with the following event. Under the decision d the state goes from i to j and and observation θ occurs.

The absence of knowledge of the state leads to replace it by the information vector $\boldsymbol{\pi}(k)$ defined by

$$\begin{aligned}\boldsymbol{\pi}(k) &= (\pi_1(k), \ldots, \pi_N(k))^*, \\ \pi_i(k) &\triangleq \Pr[\mathbf{x}_k = i \mid H_k].\end{aligned} \tag{45}$$

Actually it may be shown that $\{\boldsymbol{\pi}(k)\}_k$ is itself a Markov process. This means that $\Pr\{\boldsymbol{\pi}_{k+1} \mid \boldsymbol{\pi}_0,$

$\ldots,\pi_k,d_k\} = \Pr\{\pi_{k+1} \mid \pi_k,d_k\}$. Moreover, for a decision d and an observation θ the following updating equation holds [29]

$$\pi_j(k+1) = \frac{\sum_i \pi_i(k) p^d_{ij} r^d_{j\theta}}{\sum_{i,j} \pi_i(k) p^d_{ij} r^d_{j\theta}} = T(\pi(k) \mid d,\theta). \quad (46)$$

The transformation T updates the π vector conditionally to the new observation θ and the new decision d. In fact, it is straightforwardly obtained by means of the Bayes formula.

With matrix notations, the above formula takes the following form:

$$\boldsymbol{\pi}(k+1) = T(\boldsymbol{\pi}(k) \mid d,\theta) = \frac{1}{\sigma(\boldsymbol{\pi},d,\theta)} \boldsymbol{\pi}^*_k \mathbf{P}^d \mathbf{R}^d(\theta)$$

with

$$\sigma(\boldsymbol{\pi},d,\theta) = \boldsymbol{\pi}^*_k \mathbf{P}^d \mathbf{R}^d(\theta)\mathbf{e}.$$

In the above formula, the matrix $\mathbf{P}^d$ is the transition matrix, whose elements are $\{p^d_{ij}\}$, while the matrix $\mathbf{R}^d(\theta)$ is a diagonal matrix with (j,j) element equal to $r^d_{j\theta}$. Finally, $\mathbf{e}$ is a vector uniquely formed of 1.

The initial POMDP is equivalent to a sequential decision problem with state space $\Pi(\mathbf{x})$ and dynamic equation $\pi_{k+1} = T(\pi_k,d_k,\theta_k)$.

Consider now the computations of the optimal decisions. In that aim, it is legitimate to deal with the maximum expected value ($V_k(\boldsymbol{\pi})$) of the cost function that the system can accrue if the current information vector is $\boldsymbol{\pi}$ and if k iterations remain. Then, expanding over all possible next transitions and observations yields the following recursive equation:

$$V_k(\boldsymbol{\pi}) = \max_{d \in C_k} \left[\sum_{i=1}^{N} \pi_i(k) \sum_{j=1}^{N} p^d_{ij} \sum_{\theta} r^d_{j\theta} \{ w^d_{ij\theta} + V_{k+1}(T(\boldsymbol{\pi} \mid d,\theta)) \} \right]. \quad (47)$$

This equation can be slightly simplified if we define the expected immediate cost $\mathbf{q}$ for state i if the decision d is taken during the next control interval, giving

$$V_k(\boldsymbol{\pi}) = \max_{d \in C_k} \left[\sum_i \pi_i \left[q^d_i + \sum_{j,\theta} p^d_{ij} r_{j\theta} V_{k+1}(T(\boldsymbol{\pi} \mid d,\theta)) \right] \right] \quad (48)$$

with

$$q^d_i = \sum_{j,\theta} p^d_{ij} r^d_{j\theta} w^d_{ij\theta}.$$

Using (46) and (47) the following fundamental result has been obtained by Smallwood and Sondik [29].

The function $V_k(\boldsymbol{\pi})$ is piecewise linear and convex and can be written as

$$V_k(\boldsymbol{\pi}) = \max_{\gamma_k} \langle \boldsymbol{\alpha}^{\gamma_k}(k), \boldsymbol{\pi} \rangle \quad (49)$$

where $\langle \, , \, \rangle$ represents a scalar product and the vectors $\alpha^{\gamma_k}(k)$ are calculated by a recursion derived from (47).

Equation (49) represents the key for practical implementation of the POMDP in our context [31]. The α vectors are calculated by means of the following recursion [29]:

$$\boldsymbol{\alpha}^d(k) = \mathbf{q}^d + \sum_{\theta} \mathbf{P}^d \mathbf{R}^d(\theta) \boldsymbol{\alpha}^{l(\pi,d,\theta)} \quad (50)$$

where $l(\pi,d,\theta)$ is the index of the α vector of $\Gamma_{(k+1)}$ which determines V_{k+1} in the image of $T(\boldsymbol{\pi} \mid d,\theta)$.

The proof of the validity of this recursion is *constructive* [29]. The main idea of the Smallwood and Sondik algorithm consists in starting with an $\boldsymbol{\alpha}$ vector, corresponding to a solution with an initial probability distribution $\boldsymbol{\pi}_0$, to compute its action area looking for the boundaries with contiguous regions until the whole partition of the simplex is done.

To start, one needs to choose an (arbitrary) initial distribution π_0 and to compute the associated α vector, corresponding to the different possible decisions ($\boldsymbol{\alpha}^d(k) = \mathbf{q}^d + \sum_\theta \mathbf{P}^d \mathbf{R}^d(\theta) \boldsymbol{\alpha}^{l(\pi,d,\theta)}$).

Among these $\boldsymbol{\alpha}$ vectors, the one which is "active" on $\boldsymbol{\pi}_0$ is the vector which maximizes the scalar product $\langle \boldsymbol{\pi}_0, \boldsymbol{\alpha}^d(k) \rangle$. We thus have

$$\begin{aligned} V_k(\boldsymbol{\pi}_0) &= \max_d \langle \boldsymbol{\pi}_0, \boldsymbol{\alpha}^d(k) \rangle \\ d^* &= \arg\max_d \langle \boldsymbol{\pi}_0, \boldsymbol{\alpha}^d(k) \rangle \\ \boldsymbol{\alpha}^*(k) &= \boldsymbol{\alpha}^{d^*}(k). \end{aligned} \quad (51)$$

The partitioning of the simplex Π is determined by the two types of constraints:

$$\begin{aligned} &\langle \boldsymbol{\pi}, \mathbf{P}^{d^*} \mathbf{R}^{d^*}(\theta) [\boldsymbol{\alpha}^{l(\pi_0,d^*,\theta)} - \boldsymbol{\alpha}^j] \rangle \\ &\qquad \geq 0 \qquad \forall \quad j \neq l(\boldsymbol{\pi}_0,d^*,\theta), \quad \forall \quad \theta \\ &\langle \boldsymbol{\pi}, \mathbf{P}^{d^*} \mathbf{R}^{d^*}(\theta) [\boldsymbol{\alpha}^{l(\pi_0,d^*,\theta)} - \boldsymbol{\alpha}^j] \rangle \\ &\qquad \geq 0 \qquad \forall \quad d \neq d^*, \quad \forall \quad j \neq l(\boldsymbol{\pi}_0,d,\theta). \end{aligned} \quad (52)$$

The POMDP problem is thus reduced [29] to linear programming. More precisely, one has to deal with the following set of *linear* problems:

$$\begin{aligned} &\forall \quad j = 1,\ldots M \\ &\min_{\pi} \langle \boldsymbol{\pi}, b^j \rangle, \quad \text{(objective function)} \\ &\langle \boldsymbol{\pi}, b^m \rangle \geq 0 \qquad m = 1,\ldots,M \\ &\pi_i \geq 0 \qquad \text{and} \qquad \sum \pi_i = 1. \end{aligned} \quad (53)$$

Various implementations of the basic algorithm exist. Applications to the TMA problem is detailed in [31]. However, it is worth stressing that the computation and, overall, memory requirements rapidly increase with the problem dimensionality thus rendering it very difficult [31].

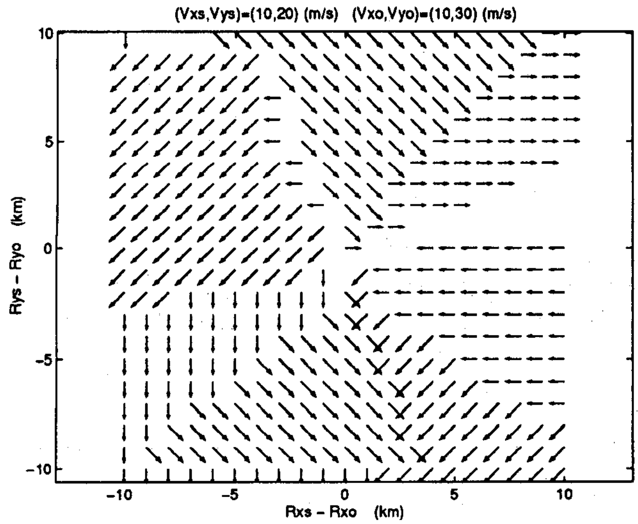


Fig. 7. Array part for $t = 0$.

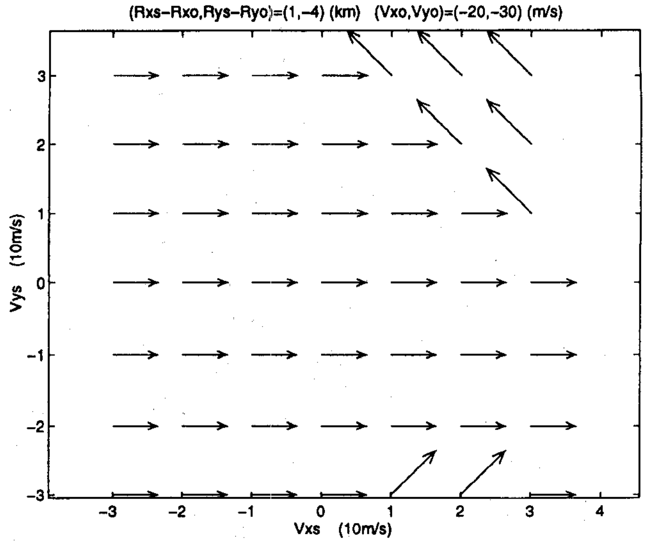


Fig. 8. Array part for $t = 0$.

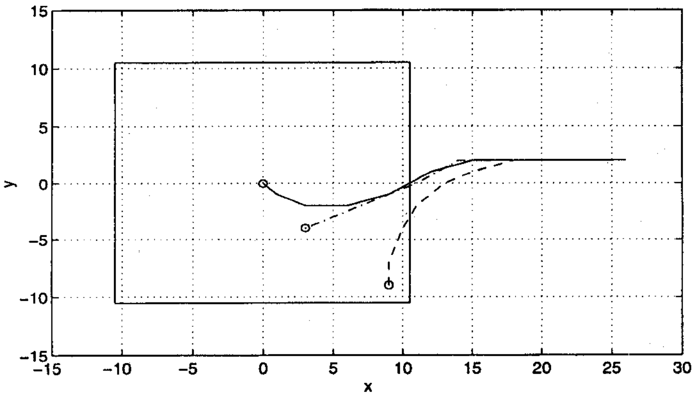


Fig. 9. Source trajectory estimation. Example 1.

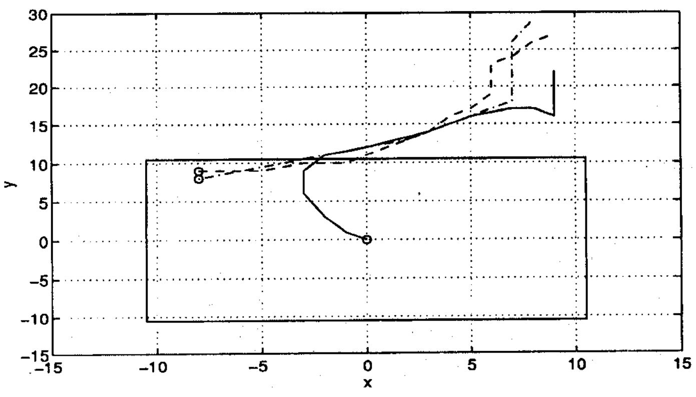


Fig. 10. Source trajectory estimation. Example 3.

VI. NUMERICAL RESULTS

The DP algorithm using the trace of the FIM as the cost function has been programmed. The aim of this algorithm is to compute an array (optimal policy) of optimal decisions for every relative position of the source, for all velocities of the source and of the observer, and finally, for every time on a finite horizon. Figs. 7 and 8 depict two bidimensionnal portions of this array. On each node (system state) an arrow represents the optimal decision, which is the change to apply to the speed vector of the observer if the system is in this state.

Figs. 9 and 10 depict two simulations where the observer uses the exact system state (complete information) to compute the optimal decision given by optimal policy; the source trajectory is then estimated. The solid lines correspond to the observer trajectory, the dashed lines correspond to the real source trajectory and finally the dashdot lines represent the estimation of the source trajectory given by the Viterbi algorithm. A circle corresponds to the beginning of the trajectory. The scale in position is in kilometers. Note that the observer trajectory is always originating at O.

Numerous simulations have been conducted, and two of them are presented here. The main conclusion that can be deduced is that the position is generally estimated with quite a good precision and that the speed estimation is not really crucial for the performance of the algorithm. In general, it can be seen that in the position space, the estimation is better at the end of the simulation than at the beginning.

VII. CONCLUSION

Markovian modeling for maneuvering source trajectory has been considered and their statistical properties (covariance) have been derived. The BOT problem has been considered in this Markovian framework. DP algorithms appear as a feasible and efficient method for solving the BOT problem. A major advantage is that the nonlinearities of the measurements are directly incuded in the algorithm. The exhaustivity of the search procedure constitutes another decisive advantage. Finally, the computation cost is quite acceptable.

The problem of observer trajectory optimization has been considered. Specific problems for application of the general DP principle have been pointed. The general framework of MDP provides a feasible solution for the complete information case. Opposite, the BOT problem with partial information can be theoretically solved by means of POMDP algorithms but basic difficulties emerge due to the number of states and decisions.

APPENDIX A. COVARIANCES

$$\operatorname{Var}(\xi_k) = \Delta T \sigma^2.$$

Recalling the following general result (valid for stochastic integrals):

$$\mathbb{E}\left\{\int_\alpha^\beta f(s)\,dw(s)\int_\alpha^\beta g(s)\,dw(s)\right\} = \sigma^2\int_\alpha^\beta f(s)g(s)\,ds$$

we obtain

$$\begin{aligned}\mathbb{E}&\left\{\left(\int_{t_k}^{t_{k+1}}(t_{k+1}-t)\,dw(t)\right)^2\right\}\\ &= \sigma^2\int_{t_k}^{t_{k+1}}(t_{k+1}-t)^2\,dt = \sigma^2\frac{\Delta T^3}{3}\end{aligned}$$

and similarly

$$\begin{aligned}\mathbb{E}\{\xi_k\eta_k\} &= \mathbb{E}\left\{(w_{k+1}-w_k)\int_{t_k}^{t_{k+1}}(t_{k+1}-t)\,dw(t)\right\}\\ &= \mathbb{E}\left\{\int_{t_k}^{t_{k+1}}\mathbb{1}\,dw(t)\int_{t_k}^{t_{k+1}}(t_{k+1}-t)\,dw(t)\right\}\\ &= \sigma^2\int_{t_k}^{t_{k+1}}(t_{k+1}-t)\,dt = \sigma^2\frac{\Delta T^2}{2}.\end{aligned}$$

APPENDIX B. CONDITIONAL PDF

Since the variables $r_{x,k+1}$ and $v_{x,k+1}$ are jointly Gaussian, the conditional pdf $p(r_{x,k+1} \mid v_{x,k+1})$ is also Gaussian and is given by

$$p(r_{x,k+1} \mid v_{x,k+1}) \sim \mathcal{N}(m_{r,k+1}, \sigma^2_{r,k+1})$$

with

$$\left|\begin{aligned} m_{r,k+1} &= r_{x,k} + \Delta T v_{x,k} + \frac{\Delta T^2}{2}\Delta T^{-1}(v_{x,k+1} - v_{x,k})\\ &= r_{x,k} + \frac{\Delta T}{2}(v_{x,k+1} + v_{x,k})\\ \sigma^2_{r,k+1} &= \sigma^2\left(\frac{\Delta T^3}{3} - \frac{\Delta T^2}{2}\Delta T^{-1}\frac{\Delta T^2}{2}\right) = \sigma^2\frac{\Delta T^3}{12}.\end{aligned}\right.$$

APPENDIX C. STOCHASTIC OBSERVABILITY

In order to study the system observability, it is worth reformulating the nonlinear system (eq. (1)–(4)) into the following one:

$$\left|\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{F}\mathbf{x}_k + \mathbf{u}_k + \omega_k\\ z_k &= \mathbf{H}_k\mathbf{x}_k\end{aligned}\right.$$

with

$$\mathbf{H}_k = (\cos\theta_k, -\sin\theta_k, 0, 0). \tag{54}$$

Thus the nonlinear measurements θ_k (eq. (4)) can be manipulated to provide a pseudomeasurement z_k which is linearly related to the target state. This constitutes the basis of the pseudolinear estimation methods [24, 31].

Let us now consider the problem of stochastic observability for a Markovian state (eq. (1)) and the BOT measurements. The benefit of using measurements in state observation is normally manifested by the reduction of a certain cost function with respect to their values when no such signals are used.

According to the definition of Boguslavskij [7], we say that a stoschastic linear system is stochastically observable if, in estimating its state from its output, the posterior error variances of the state components are strictly smaller than the priors.

Let $\hat{\mathbf{x}}_k$ be the linear least-mean-square estimate of $\mathbf{x}_k$ given the measurements $\{y_k, y_{k-1}, \ldots, y_0\}$ and define the matrices Π_k and $\mathbf{P}_k$:

$$\mathbf{\Pi}_k = \mathbb{E}\{\mathbf{x}_k\mathbf{x}_k^*\}$$
$$\mathbf{P}_k = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^*\}.$$

Then we consider the following definition of observability.

DEFINITION 1 The system (eqs. (1)–(4)) is stochastically observable if and only if (iff):

$$\mathbf{e}_i^*\mathbf{P}_k\mathbf{e}_i < \mathbf{e}_i^*\mathbf{\Pi}_k\mathbf{e}_i, \qquad 1 \le i \le n$$

$\{\mathbf{e}_i\}_{i=1}^n$, usual orthogonal basis of $\mathbb{R}^n$, $n = \dim \mathbf{x}$.

It can be easily shown [7, ch. 4] that the general form of $\mathbf{P}_k$ is

$$\mathbf{P}_k = \mathbf{\Pi}_k - \mathbf{L}_k\mathbf{\Sigma}_k^{-1}\mathbf{L}_k^*. \tag{55}$$

The rectangular matrix $(n \times kn)$ $\mathbf{L}_k$ is defined later.

Since the matrix $\mathbf{\Sigma}_k$ is the covariance matrix of the noise measurements, the matrix $\mathbf{\Pi}_k - \mathbf{P}_k$ is positive semidefinite. So, the general inequality $\mathbf{e}_i^*\mathbf{P}_k\mathbf{e}_i \le \mathbf{e}_i^*\mathbf{\Pi}_k\mathbf{e}_i$ always holds. The values of the index i ensuring a strict inequality are the observable state composents. Actually, the above definition is rather arbitrary and does not take into account the possible coupling between the estimate of the state components. Then, a convenient definition may be the estimability definition of Baram and Kailath [3].

DEFINITION 2 The system (eq. (1)–(4)) is estimable if:

$$\mathbf{\Pi}_k - \mathbf{P}_k \quad \text{is positive definite.}$$

Denote $\theta(\mathbf{L}_k)$ the number of rows of the matrix $\mathbf{L}_k$ with non-zero elements, then a direct consequence of (55) is that the system ((1)–(4)) is stochastically observable (Def. 1) iff $\theta(\mathbf{L}_k)$ is equal to n.

Since $\mathbf{\Sigma}_k$ is positive definite, it follows from (55) that the system is estimable (Def. 2) iff the rank of $\mathbf{L}_k$ is equal to n.

The matrix $\mathbf{P}_k$ is actually the covariance matrix of the conditional expectation of the state $\mathbf{x}_k$ given the measurements $\{y_0, y_1, \ldots, y_k\}$ and direct calculations [3, 7] give the following expression of $\mathbf{L}_k$:

$$\mathbf{L}_k = \{\mathbf{\Phi}_{k,0}N_0, \mathbf{\Phi}_{k,1}N_1, \ldots, \mathbf{\Phi}_{k,k}N_k\}$$

where

$$\mathbf{\Phi}_{k,j} = \mathbf{F}^{k-j}, \qquad j < k$$
$$\vdots$$
$$\mathbf{\Phi}_{k,k} = \mathbf{Id}$$
$$N_j = \mathbf{\Pi}_j\mathbf{H}_j^*.$$

and $\mathbf{\Pi}_j$ satisfies the Lyapunov equation:

$$\mathbf{\Pi}_{j+1} = \mathbf{F}\mathbf{\Pi}_j\mathbf{F}^* + \mathbf{Q}. \tag{56}$$

From (56) and quite direct but lengthly calculations we see that the BOT system (eq. (1)–(4)) is stochastically observable except very pathological cases, i.e., the sequence $\{\cos\theta_0, \ldots, \cos\theta_k\}$ or $\{\sin\theta_0, \ldots, \sin\theta_k\}$ are identically null.

Stochastic observability is thus less demanding than deterministic observability [24] which requires a maneuver of the observer. Furthermore, direct calculations prove that the rank of $\mathbf{L}_k$ is generally equal to 4 except for a zero bearing-rate scenario proving thus that the BOT system is generally estimable.

REFERENCES

[1] Averbuch, A., Itzikovitz, S., and Kapon, T. (1991) Radar target tracking—Viterbi versus IMM. *IEEE Transactions on Aerospace and Electronics Systems*, **27**, 3 (May 1991), 550.

[2] Bagchi, A. (1993) *Optimal Control of Stochastic Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[3] Baram, Y., and Kailath, T. (1988) Estimability and regulability of linear systems. *IEEE Transactions on Automatic Control*, **33**, 12 (Dec. 1988), 1116–1121.

[4] Barniv, Y. (1985) Dynamic programming solution for detecting dim moving target. *IEEE Transactions on Aerospace and Electronic Systems*, **21**, 1 (Jan. 1985), 144–156.

[5] Basseville, M., and Nikiforov, I. (1992) *Detection of Abrupt Changes.* Englewood Cliffs, NJ: Prentice-Hall, 1992.

[6] Blom, H. A. P., and Bar-Shalom, Y. (1988) The interacting multiple model algorithm with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, **33**, 8 (Aug. 1988), 780–783.

[7] Boguslavskij, I. A. (1988) *Filtering and Control*, (n.y. ed.). Optimization Software Inc., Publication Division, 1988.

[8] Burdic, W. S. (1991) *Underwater Acoustic System Analysis* (2nd ed.), Signal Processing Series. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[9] Chang, C.-B., and Tabaczinski, J. A. (1984) Application of state estimation to target tracking. *IEEE Transactions on Automatic Control*, **29**, 2 (Feb. 1984), 98.

[10] Forney, G. D. (1973) The viterbi algorithm. *Proceedings of the IEEE*, **61**, 3 (Mar. 1973), 268–278.

[11] Gholson, N. H., and Moose, R. L. (1977) Maneuvering target tracking using adaptive state estimation. *IEEE Transactions on Aerospace and Electronic Systems*, **13**, 3 (May 1977), 310–317.

[12] Hurd, H. L. (1977) A Bayesian approach to simultaneous tracking of multiple targets. In *Proceedings of the ONR Passive Tracking Conference*, Monterey, CA, May 1977.

[13] Kumar, P. R., and Varaiya, P. (1986) *Stochastic Systems (Estimation, Identification and Adaptive Control).* Englewood Cliffs, NJ: Prentice-Hall, 1986.

[14] Lagunov, V. N. (1985)
Introduction to Differential Games and Control Theory (Sigma Series in Applied Sciences).
Berlin: Helderman Verlag, 1985.

[15] Lane, D. E. (1989)
A partially observable model of decision making by fisherman.
Operations Research, **37**, 2 (Mar.–Apr. 1989), 240–254.

[16] Larson, R. E., and Peschon, J. (1966)
A dynamic programming approach to trajectory estimation.
IEEE Transactions on Automatic Control (July 1966), 537–540.

[17] Le Cadre, J.-P., and Trémois, O. (1997)
The matrix dynamic programming property and its implications.
To be published in *SIAM Journal on Matrix Analysis*, 1997.

[18] Logothetis, A. (1994)
Bearings-only tracking.
Master's thesis, School of Information Technology and Electrical Engineering, University of Melbourne, Australia, 1994.

[19] Logothetis, A., Evans, R. J., and Sciacca, L. J. (1994)
Bearing-only tracking using hidden Markov models.
In *Proceedings of 33rd Conference on Decision and Control*, Lake Buena-Vista, FL, Dec. 1994, 3301–3302.

[20] Lovejoy, W. S. (1991)
Computationally feasible bounds for partially observed Markov decision processes.
Operations Research, **39**, 1 (Jan.–Feb. 1991), 162–175.

[21] Martinerie, F., and Forster, P. (1992)
Data association and tracking from distributed sensors using hidden Markov models and dynamic programming.
ICASSP, Mar. 1992.

[22] Martinerie, F., and Forster, P. (1992)
Data association and tracking from distributed sensors using hidden Markov models and evidential reasoning.
CDC, Dec. 1992.

[23] McAulay, R. J., and Delinger, E. J. (1973)
A decision-directed adaptive tracker.
IEEE Transactions on Aerospace and Electronic Systems, **AES-9** (Mar. 1973), 229–236.

[24] Nardone, S. C., Lindgren, A. G., and Gong, K. F. (1984)
Fundamental properties and performance of conventional bearings-only target motion analysis.
IEEE Transactions on Automatic Control, **AC-29**, 9 (Sept. 1984), 775–787.

[25] Papoulis, A. (1965)
Probability, Random Variables and Stochastic Processes.
New York: McGraw-Hill, 1965.

[26] Rabiner, L. R. (1989)
A tutorial on hidden Markov models and selected applications in speech recognition.
Proceedings of the IEEE, **77**, 2 (Feb. 1989), 257.

[27] Ross, S. (1992)
Introduction to Stochastic Dynamic Programming.
New York: Academic Press, 1992.

[28] Singer, R. A. (1970)
Estimation optimal tracking filter performance for manned maneuvering targets.
IEEE Transactions on Aerospace and Electronic Systems, **AES-6** (July 1970), 473–483.

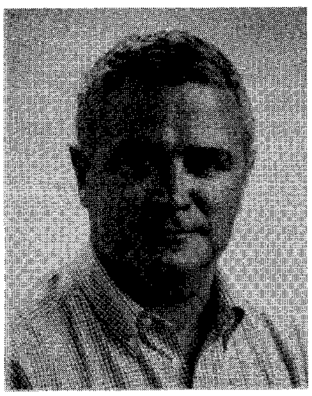
[29] Smallwood, R. D., and Sondik, E. J. (1973)
The optimal control of partially observable Markov processes over a finite horizon.
Operations Research, **21** (1973), 1071–1088.

[30] Thorp, J. S. (1973)
Optimal tracking of maneuvering targets.
IEEE Transactions on Aerospace and Electronic Systems, **AES-9** (July 1973), 512–519.

[31] Trémois, O. (1995)
Étude de méthodes de trajectographie pour des sources maneuvrantes.
Ph.D. thesis, Université de Rennes I, France, June, 1995.

[32] Trémois, O., and Le Cadre, J. P. (1994)
Maneuvering target motion analysis using hidden Markov model.
ICASSP, 1994.

[33] Xie, X., and Evans, R. J. (1991)
Multiple target tracking and multiple frequency line tracking using hidden Markov models.
IEEE Transactions on Signal Processing, **39**, 12 (Dec. 1991), 2659–2676.



J. P. Le Cadre (M'93) received the M.S. degree in mathematics in 1977, the Doctorat de 3^{-eme} cycle in 1982, and the Doctorat d'Etat in 1987, both from INPG, Grenoble, France.

From 1980 to 1989, he worked at the GERDSM (Groupe d'Etudes et de Recherche en Detection Sous-Marines), a laboratory of the DCN (Direction des Constructions Navales), mainly on array processing. Since 1989, he has been with IRISA/CNRS, where he is a CNRS "Directeur de Recherche". His interests have now moved toward topics like system analysis, detection, data association and operations research.

He received (with O. Zugmeyer) the Eurasip *Signal Processing* best paper award in 1993. He is a member of various societies of IEEE and of SIAM.

Olivier Trémois graduated from the *École Nationale Supérieure de Physique de Marseille* in 1990 and received his *Diplôme d'Étude Approfondie* in signal processing from the University of Rennes 1 in 1992.

He joined the IRISA in 1992 and received the Ph.D. degree in signal processing in June 1995. His research interests are target tracking and target motion analysis.

Since November 1995, he has been with Thomson-Marconi Sonars, Brest, working on sonor projects.