

On the Convergence of Iterative Methods for Bearings-Only Tracking

J.-P. LE CADRE, Member, IEEE
IRISA/CNRS
France

C. JAUFFRET, Member, IEEE
GESSY/ISITV
France

This paper deals with the analysis of the convergence of iterative methods for bearings-only tracking (BOT). A geometric and unified framework is developed. Explicit sufficient conditions ensuring convergence of iterative methods for maximization of the likelihood functional are then derived.

Manuscript received May 8, 1996; revised April 27, 1997 and June 8, 1998.

IEEE Log No. T-AES/35/3/04049.

This work was supported by DCN/Ingénierie/Sud, (Dir. Const. Navales), France.

Authors' addresses: J.-P. Le Cadre, IRISA/CNRS, Campus de Beaulieu, 35 042, Rennes, France, E-mail: (lecadre@irisa.fr); C. Jauffret, GESSY/ISITV, Université de Toulon, Avenue G. Pompidou, B. P. 56, La Valette du Var, France, E-mail: (jauffret@isitiv.univ.tln.fr).

0018-9251/99/\$10.00 © 1999 IEEE

I. INTRODUCTION

Bearings-only tracking (BOT) is the determination of the trajectory of a target solely from bearing measurements. The BOT area has been widely investigated [1–4] and numerous solutions for this problem have been proposed [5].

However the measurement equation relating the target bearing to its state is highly nonlinear. Classical least square methods as the Kalman filter cannot be directly applied. This advocates for the use of extended Kalman filter for which convergence is improved by using modified polar coordinates [4]. Another approach is the pseudolinear estimator (PLE) formulation proposed in [1] which lumps the nonlinearities into the noise term, resulting in a linear measurement equation. However, the measurement matrix contains elements that are functions of noisy bearings and, overall, are correlated with the noise terms of the measurement equation. As a result, the PLE exhibits a bias which can be severe [1, 3].

To overcome this bias, maximum likelihood and instrumental variables have been developed [1, 3, 6] and give satisfactory results. These are gradient search based on a batch processing of all the available measurements. The latter feature may be quite useful if there are missing data (track interruption). So batch methods are generally considered as reliable. However, their proper convergence are reputed to be sensitive to initial conditions and step sizes.

So the objective of this work is to present sufficient conditions ensuring convergence of iterative (i.e., gradient-like, instrumental variable) search methods. A common point of all these algorithms is that they are of batch type. First, we give a general presentation of a basic result of Iltis and Anderson [7]. In fact, a general analysis based on linear and multilinear algebra reveals the fundamental nature of the problem and (almost) completely allows us to avoid boring calculations. It is then possible to consider a unique framework for target-observer scenarios of increasing complexity, common to all the methods. For instance, we consider successively the cases of a maneuvering observer, a maneuvering source (with known maneuvering instants) and, finally, unknown maneuvering instants. Thus, we can obtain sufficient convergence conditions for maneuvering source and observer. Besides their own theoretical interests, these conditions yield feasible methods for the general BOT problems. All these conditions come from a unique result which is strongly related to the quasiconcavity property (see Section III). Thus, even if the tribute to the seminal work of Iltis and Anderson is clear, different interpretations and extensions are developed here.

The paper is organized as follows. Section II deals with the general formulation of the BOT problem. The case of nonmaneuvering target and

observer is considered in Section III. The general formalism is then presented. Its generality allows us to straightforwardly handle more complex scenarios (Section IV), say maneuvering source and observer scenarios. A geometric interpretation of the convergence result is provided. Then the convergence of gradient methods is considered (Section V) and is followed by the analysis of the instrumental variable (Section VI). Extensions to constrained TMA (target motion analysis) problems are considered in Section V. In Section VII we deal with the estimation of the unknown source maneuver instants, while the multiple source case is considered in Section VIII. Finally, simulation results illustrate the interest of the global TMA approach.

The following standard notations are used throughout this work:

- 1) bold caps are used for vectors while a standard capital letter denotes a matrix,
- 2) the symbol (*) means transposition, \det represents the determinant,
- 3) r_x and r_y represent x and y relative coordinates, v_x and v_y denote relative x and y velocities,
- 4) integers $\{i, j, m, k\}$ are generally time variables, m or j generally represent a maneuver (source or observer) instant, l is an iteration index, p is the total number of measurements,
- 5) diag denotes a diagonal matrix, Id_n is the n -dimensional identity matrix, \otimes is the Kronecker product,
- 6) $\sin(\mathbf{A}, \mathbf{B})$ denotes the sine of the angle formed with the vectors \mathbf{A} and \mathbf{B} .

II. PROBLEM FORMULATION

Let us now briefly recall the BOT TMA model [1, 3] for, first, a constant velocity target. Also, for the sake of simplicity, the presentation is restricted to the planar problem.

Let \mathbf{X} be the relative state vector of the target (T), defined by

$$\mathbf{X} = \mathbf{X}_T - \mathbf{X}_{\text{obs}} \triangleq [r_x, r_y, v_x, v_y]^* \quad (1)$$

The discrete time equation of the relative motion takes the following form:

$$\mathbf{X}(k) = \Phi(k, k-1)\mathbf{X}(k-1) + \mathbf{U}(k)$$

where

$$\Phi(k, k-1) = \begin{pmatrix} Id_2 & \alpha Id_2 \\ 0 & Id_2 \end{pmatrix}, \quad (2)$$

$$Id_2 \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \alpha \triangleq t_k - t_{k-1}.$$

In the above formula the vector $\mathbf{U}(k)$ accounts for the effects of the observer accelerations. The matrix $\Phi(k, k-1)$ is the transition matrix, simply denoted

F in the sequel. Also, for the sake of simplicity, we assume that α is 1. The measurement equation is simply (without estimation noise):

$$\beta_k = \tan^{-1} \left(\frac{r_{x,k}}{r_{y,k}} \right). \quad (3)$$

From (2) we can see that the target trajectory is determined by a state vector \mathbf{X} defined at a given instant called the reference time. Note that this reference time is arbitrary. So, the TMA problem may be converted to the estimation of the state vector \mathbf{X} .

Let \mathbf{B} be the history of measured bearings, then the likelihood functional is [1]

$$p(\tilde{\mathbf{B}} | \hat{\mathbf{X}}) = \text{cst} \exp \left[-\frac{1}{2} \|\tilde{\mathbf{B}} - \mathbf{B}(\hat{\mathbf{X}})\|_{\Sigma^{-1}}^2 \right]$$

where¹

$$\tilde{\mathbf{B}} = (\tilde{\beta}_1, \dots, \tilde{\beta}_p)^* \quad (4)$$

$$\tilde{\beta}_k = \beta_k + w_k, \quad w : \mathcal{N}(0, \Sigma).$$

Obviously the likelihood functional depends nonlinearly on \mathbf{X} and there is no explicit solution for determining the maximum of the likelihood functional. This problem is usually solved by using a Gauss–Newton algorithm [1] which has the general form²

$$\hat{\mathbf{X}}_{\ell+1} = \hat{\mathbf{X}}_{\ell} - \rho_{\ell} \left[\left(\frac{\partial \tilde{\mathbf{B}}}{\partial \mathbf{X}} \right)^* \Sigma^{-1} \frac{\partial \tilde{\mathbf{B}}}{\partial \mathbf{X}} \right]^{-1} \times \left(\frac{\partial \tilde{\mathbf{B}}}{\partial \mathbf{X}} \right)^* \Sigma^{-1} (\tilde{\mathbf{B}} - \hat{\mathbf{B}}) \quad (5)$$

where ℓ is the iteration index, ρ_{ℓ} is the step size, and $\hat{\mathbf{B}} = \mathbf{B}(\hat{\mathbf{X}}_{\ell})$.

For the sake of simplicity we simply denote $\hat{\mathbf{B}}$ the vector $\mathbf{B}(\hat{\mathbf{X}})$. Analogously, we denote \mathbf{B} the (exact) measurement vector associated with the exact target state vector \mathbf{X} (no measurement noise).

First, we now consider the (exact) likelihood functional $p(\mathbf{B} | \hat{\mathbf{X}})$ and investigate a functional of its gradient vector. Then, the analysis will be straightforwardly extended to the (true) likelihood functional $p(\tilde{\mathbf{B}} | \hat{\mathbf{X}})$ (see Sect. V).

At this point, it is worth stressing that the likelihood functional is not concave. This is illustrated by Fig. 1. The exact Hessian matrix of the likelihood functional is calculated (see (40)) for various values of the components of the source state vector. For instance, we consider velocity components variations, v_x is varying between 0.9 m/s and 1.1 m/s while v_y values are comprised between 7.9 m/s and 8.1 m/s. The exact values of v_x and v_y are, respectively, 1 m/s and 8 m/s. The values of the couple v_x, v_y for which at

¹ w_k : white gaussian noise (WGN) sequence.

² $\Sigma = \text{Diag}(\sigma_i^2), \sigma_i^2$: noise variance of the i th measurement.

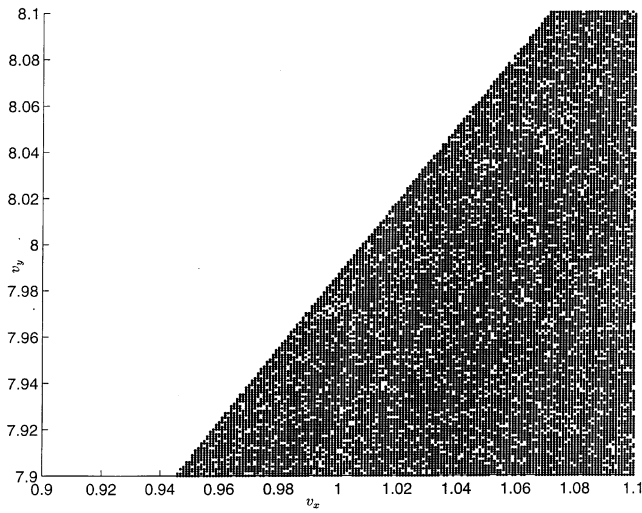


Fig. 1. Eigenvalues of exact Hessian matrix (exact values of v_x and v_y : 1 m/s and 8 m/s).

least one eigenvalue of the Hessian matrix is negative are represented by a point (in black). It is clear from Fig. 1 that the likelihood functional is not concave, even in a small neighborhood of the exact values of the state vector components. This constitutes the major motivation of this paper and we shall show that convergence of the iterative methods may be guaranteed despite the rather pathological shape of the Hessian eigenvalues.

III. CASE OF NONMANEUVERING TARGET AND OBSERVER

Similarly to Iltis and Anderson [7] we consider the following quadratic functional of $\hat{\mathbf{X}}$:

$$L(\hat{\mathbf{X}}) = \|\hat{\mathbf{X}} - \mathbf{X}\|^2. \quad (6)$$

Iltis, et al. show that this functional is a Lyapunov function [8] for the (continuous) differential equation:

$$\frac{d}{dt}\hat{\mathbf{X}} = \mathbf{G}(\hat{\mathbf{X}}). \quad (7)$$

In (7), $\mathbf{G}(\hat{\mathbf{X}})$ is the gradient vector of the (reduced) log-likelihood functional (4) at the point $\hat{\mathbf{X}}$.

Also, similarly to [7] we consider the following functional of $\hat{\mathbf{X}}$:

$$\begin{aligned} \dot{L}(\hat{\mathbf{X}}) &= 2(\hat{\mathbf{X}} - \mathbf{X})^* \frac{d}{dt}\hat{\mathbf{X}} \\ &= 2(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}}). \end{aligned} \quad (8)$$

Let us now detail the vector $\mathbf{G}(\hat{\mathbf{X}})$. Assuming a constant data rate (i.e., $\alpha = 1$) and a reference time equal to zero, elementary (classical) calculations yield [1]:

$$\mathbf{G}(\hat{\mathbf{X}}) = H^*(\hat{\mathbf{X}}) \begin{pmatrix} \beta_1 - \hat{\beta}_1 \\ \vdots \\ \beta_p - \hat{\beta}_p \end{pmatrix}$$

where

$$\beta_i \triangleq \beta_i(\mathbf{X}), \quad \hat{\beta}_i \triangleq \beta_i(\hat{\mathbf{X}})$$

and

$$H(\hat{\mathbf{X}}) = \begin{pmatrix} \frac{\cos \hat{\beta}_1}{\hat{r}_1} & -\frac{\sin \hat{\beta}_1}{\hat{r}_1} & \frac{\cos \hat{\beta}_1}{\hat{r}_1} & -\frac{\sin \hat{\beta}_1}{\hat{r}_1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\cos \hat{\beta}_p}{\hat{r}_p} & -\frac{\sin \hat{\beta}_p}{\hat{r}_p} & \frac{p \cos \hat{\beta}_p}{\hat{r}_p} & -\frac{p \sin \hat{\beta}_p}{\hat{r}_p} \end{pmatrix}$$

$$\hat{r}_i = (r_{x,i}^2(\hat{\mathbf{X}}) + r_{y,i}^2(\hat{\mathbf{X}}))^{1/2}, \quad H^*(\hat{\mathbf{X}}) : 4 \times p. \quad (9)$$

We now calculate the $(1 \times p)$ row matrix $(\hat{\mathbf{X}} - \mathbf{X})^* \cdot H^*(\hat{\mathbf{X}})$ of which the k th element (denoted I_k) takes the following form:

$$\begin{aligned} I_k &= \frac{1}{\hat{r}_k} (\hat{\mathbf{X}} - \mathbf{X})^* \begin{pmatrix} \cos \hat{\beta}_k \\ -\sin \hat{\beta}_k \\ k \cos \hat{\beta}_k \\ -k \sin \hat{\beta}_k \end{pmatrix} \\ &= \frac{1}{\hat{r}_k^2} (\hat{r}_{y,k}, -\hat{r}_{x,k}, k\hat{r}_{y,k}, -k\hat{r}_{x,k}) (\hat{\mathbf{X}} - \mathbf{X}). \end{aligned} \quad (10)$$

The second equality of (10) has been obtained by expliciting the values of $\cos \hat{\beta}_k$ and $\sin \hat{\beta}_k$ (ref. time = 0):

$$\cos \hat{\beta}_k = \hat{r}_{y,k} / \hat{r}_k, \quad \sin \hat{\beta}_k = \hat{r}_{x,k} / \hat{r}_k. \quad (11)$$

Let us now denote $\hat{\mathbf{W}}_k$ the four-dimensional vector defined by

$$\hat{\mathbf{W}}_k \triangleq (\hat{r}_{y,k}, -\hat{r}_{x,k}, k\hat{r}_{y,k}, -k\hat{r}_{x,k})^*$$

then, I_k (10) may be calculated by means of the two following elementary lemmas. First, let us rewrite I_k in a more compact form:

$$I_k = \frac{1}{\hat{r}_k^2} \hat{\mathbf{W}}_k^* (\hat{\mathbf{X}} - \mathbf{X})$$

then we have the following.

LEMMA 1 *The following equality holds:*

$$\hat{\mathbf{W}}_k^* \hat{\mathbf{X}} = 0, \quad k = 1, \dots, p.$$

PROOF The following equality is straightforwardly deduced from the definition (see (10)) of $\hat{\mathbf{W}}_k$, i.e.,

$$\hat{\mathbf{W}}_k = M_k \hat{\mathbf{X}}$$

where

$$\begin{aligned} M_k &= \begin{pmatrix} 0 & 1 & 0 & k \\ -1 & 0 & -k & 0 \\ 0 & k & 0 & k^2 \\ -k & 0 & -k^2 & 0 \end{pmatrix} = \begin{pmatrix} J & kJ \\ kJ & k^2J \end{pmatrix} \\ &= \begin{pmatrix} 1 & k \\ k & k^2 \end{pmatrix} \otimes J \end{aligned} \quad (12)$$

with

$$J \triangleq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \otimes : \text{Kronecker product [9].}$$

From (12), we have

$$\hat{\mathbf{W}}_k^* \hat{\mathbf{X}} = \hat{\mathbf{X}}^* M_k \hat{\mathbf{X}}$$

but the matrix M_k is antisymmetric [9] so that $\hat{\mathbf{X}}^* M_k \hat{\mathbf{X}}$ is null.

LEMMA 2 *The following equality holds:*

$$-\frac{1}{\hat{r}_k^2} \hat{\mathbf{W}}_k^* \mathbf{X} = \frac{r_k}{\hat{r}_k} \sin(\hat{\beta}_k - \beta_k).$$

PROOF First, the vectors \mathbf{X} and $\hat{\mathbf{X}}$ are partitioned in position (\mathbf{R}) and velocity (\mathbf{V}) subvectors as below:

$$\mathbf{X} \triangleq \begin{pmatrix} \mathbf{R} \\ \mathbf{V} \end{pmatrix}, \quad \hat{\mathbf{X}} \triangleq \begin{pmatrix} \hat{\mathbf{R}} \\ \hat{\mathbf{V}} \end{pmatrix}$$

we then obtain

$$\begin{aligned} \hat{\mathbf{W}}_k^* \mathbf{X} &= (\hat{\mathbf{R}}^*, \hat{\mathbf{V}}^*) \begin{pmatrix} J & kJ \\ kJ & k^2 J \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{V} \end{pmatrix} \\ &= (\hat{\mathbf{R}} + k\hat{\mathbf{V}})^* J (\mathbf{R} + k\mathbf{V}). \end{aligned} \quad (13)$$

Now the following elementary remark is instrumental. If \mathbf{U} and \mathbf{U}' are two vectors of \mathbb{R}^2 , then:

$$\begin{aligned} \mathbf{U}^* J \mathbf{U}' &= u_x u'_y - u'_x u_y \\ &= \det(\mathbf{U}, \mathbf{U}'). \end{aligned} \quad (14)$$

Collecting (13) and (14), we obtain

$$\hat{\mathbf{W}}_k^* \mathbf{X} = \det(\hat{\mathbf{R}} + k\hat{\mathbf{V}}, \mathbf{R} + k\mathbf{V})$$

and therefore

$$\begin{aligned} \frac{1}{\hat{r}_k^2} \hat{\mathbf{W}}_k^* \mathbf{X} &= \frac{r_k}{\hat{r}_k} \frac{\det(\hat{\mathbf{R}}_k, \mathbf{R}_k)}{\|\mathbf{R}_k\| \|\hat{\mathbf{R}}_k\|} \\ &= \frac{r_k}{\hat{r}_k} \sin(\hat{\mathbf{R}}_k, \mathbf{R}_k) \end{aligned} \quad (15)$$

where

$$\mathbf{R}_k \triangleq \mathbf{R} + k\mathbf{V}, \quad \hat{\mathbf{R}}_k \triangleq \hat{\mathbf{R}} + k\hat{\mathbf{V}}.$$

Now we have from the azimuth definition:

$$\sin(\hat{\mathbf{R}}_k, \mathbf{R}_k) = \sin(\beta_k - \hat{\beta}_k)$$

which proves Lemma 2. The following result has thus been proved.

PROPOSITION 1 *From Lemmas 1 and 2 the time derivative $\dot{L}(\hat{\mathbf{X}})$ of the Liapunov function $L(\hat{\mathbf{X}})$ (eq. (6))*

is

$$\begin{aligned} \dot{L}(\hat{\mathbf{X}}) &= 2(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}}) \\ &= -2 \sum_{k=1}^p (\hat{\beta}_k - \beta_k) I_k \\ &= -2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k). \end{aligned} \quad (16)$$

The above result (16) is quite identical to the Iltis one [7] and may also be proved by direct trigonometric manipulations. However the main advantage of the matrix formulation lies in the fact that it can be straightforwardly extended to more complex source-observer scenarios.

When *all* the $\{\beta_k - \hat{\beta}_k\}_{k=1}^p$ lie in the interval $]-\pi, \pi[$, all the scalars $(\beta_k - \hat{\beta}_k) \sin(\hat{\beta}_k - \beta_k)$ are positive which implies that $\dot{L}(\hat{\mathbf{X}})$ cannot be null and in turn the gradient vector $\mathbf{G}(\hat{\mathbf{X}})$ cannot be the zero vector. Therefore, a local maximum of the likelihood functional cannot exist on all the $\hat{\mathbf{X}}$ domain verifying the above condition.

According to (16), the above condition can imply the convergence of the gradient algorithm only if the stepsize is correctly chosen and if the gradient is a *descent direction*. This is the object of Section V, where sufficient convergence conditions are derived. Further, note that the observability conditions do not appear in the previous analysis. In fact, the problems are not of the same nature and there is no contradiction since convergence means convergence towards *observable parameters*. However, practically, it is worthy that the TMA problem be observable. For instance, the classical TMA problem will be (generally) observable if multiple measurements (e.g. multiple arrays) are available at each scan. We refer to [10] for an elementary proof.

Up to now, our analysis was largely inspired by the work of Iltis, et al. The Lyapunov function (6) played a main role. Another point of view is that of numerical analysis. The basic scalar product $(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}})$ will then receive another (perhaps more natural) interpretation. At this point, it is worth to recall the general definition of a quasiconvex functional [13].

DEFINITION 1 Let $f : S \rightarrow E$, where S is a nonempty convex set. The functional f is said to be quasiconvex if, for each \mathbf{X}_1 and $\mathbf{X}_2 \in S$, the following inequality holds true for each $\lambda \in [0, 1]$:

$$f[\lambda \mathbf{X}_1 + (1 - \lambda) \mathbf{X}_2] \leq \max\{f(\mathbf{X}_1), f(\mathbf{X}_2)\}.$$

Analogously, the functional f is said quasiconcave if $-f$ is quasiconvex. Practically, the following (classical) characterization of a differentiable quasiconvex functional [13] is very useful.

THEOREM 1 [13] *Let S be a nonempty open convex set and let f be differentiable on S . Then f is quasiconvex if the following statement holds:*

$$\begin{aligned} \text{If } \mathbf{X}_1, \mathbf{X}_2 \in S \quad \text{and} \quad f(\mathbf{X}_1) \leq f(\mathbf{X}_2), \\ \text{then} \quad \nabla^* f(\mathbf{X}_2)(\mathbf{X}_2 - \mathbf{X}_1) \geq 0. \end{aligned} \quad (17)$$

The definition of $\dot{L}(\hat{\mathbf{X}})$ (see (8)) is thus quite similar to the differential characterization of the quasiconvexity. Actually, the aim of the ‘‘main results’’ (see Sections III–VI) is to prove that the likelihood functional is quasiconcave on a ‘‘reasonable’’ domain of the state space. In fact, these results (see Properties 1, 2, 3) imply the validity of the following proposition:

$$\begin{aligned} \text{If: } -L(\hat{\mathbf{X}}) \geq -L(\mathbf{X}), \quad \text{then:} \\ -\nabla^* L(\hat{\mathbf{X}})(\hat{\mathbf{X}} - \mathbf{X}) = \mathbf{G}^*(\hat{\mathbf{X}})(\hat{\mathbf{X}} - \mathbf{X}) \leq 0. \end{aligned} \quad (18)$$

From which, we deduce that the log-likelihood functional L is quasiconcave. A general property of quasiconcave functionals is that a local maximum is actually a global maximum [13], on the quasiconcavity domain. However, note [7] that the $\tan^{-1}(\cdot)$ function is only defined without ambiguity in the range $[-\pi/2, \pi/2]$.

IV. EXTENSION TO MANEUVERING SOURCE AND OBSERVER SCENARIOS

We now show how the previous formalism may be extended to more complex scenarios. First, we restrict our attention to a *nonmaneuvering target*.

We consider that the *observer trajectory* is a leg-by-leg one and we denote m its *maneuver instant*. According to (2), (3), we have ($j \geq 0$):

$$\begin{aligned} \hat{r}_x(m+j) &= \hat{r}_x(0) + (m+j)\hat{v}_x + ju_x \\ \hat{r}_y(m+j) &= \hat{r}_y(0) + (m+j)\hat{v}_y + ju_y \end{aligned} \quad (19)$$

u_x, u_y : observer change of velocity (maneuver) and $k = m + j, j \geq 0$.

Using (19), the previous formalism yields

$$\hat{r}_k^2 \{ \hat{\mathbf{X}}^* H^*(\hat{\mathbf{X}}) \}_{j_k} = (\hat{\mathbf{R}}^*, \hat{\mathbf{V}}^*) \left[M_k \begin{pmatrix} \hat{\mathbf{R}} \\ \hat{\mathbf{V}} \end{pmatrix} + j \begin{pmatrix} J\mathbf{U} \\ kJ\mathbf{U} \end{pmatrix} \right]$$

where

$$\begin{aligned} M_k &= \begin{pmatrix} 1 & k \\ k & k^2 \end{pmatrix} \otimes J \\ \mathbf{U} &= \begin{pmatrix} u_x \\ u_y \end{pmatrix}, \quad k = m + j \end{aligned}$$

so that³

$$\hat{r}_k^2 \{ \hat{\mathbf{X}}^* H^*(\hat{\mathbf{X}}) \}_{j_k} = j(\hat{\mathbf{R}} + k\hat{\mathbf{V}})^* J\mathbf{U}. \quad (20)$$

³ $\{ \hat{\mathbf{X}}^* H^*(\hat{\mathbf{X}}) \}_{j_k}$ denotes the k th element of $\hat{\mathbf{X}}^* H^*(\hat{\mathbf{X}})$.

Similarly, the other term of $\dot{L}(\hat{\mathbf{X}})$ is obtained

$$\begin{aligned} \hat{r}_k^2 \{ \mathbf{X}^* H^*(\hat{\mathbf{X}}) \}_{j_k} &= (\mathbf{R}, \mathbf{V})^* \left[M_k \begin{pmatrix} \hat{\mathbf{R}} \\ \hat{\mathbf{V}} \end{pmatrix} + j \begin{pmatrix} J\mathbf{U} \\ kJ\mathbf{U} \end{pmatrix} \right] \\ &= (\mathbf{R} + k\mathbf{V})^* J(\hat{\mathbf{R}} + k\hat{\mathbf{V}}) + j(\mathbf{R} + k\mathbf{V})^* J\mathbf{U}. \end{aligned} \quad (21)$$

Let us now calculate the following bilinear form:

$$\begin{aligned} (\mathbf{R} + k\mathbf{V} + j\mathbf{U})^* J(\hat{\mathbf{R}} + k\hat{\mathbf{V}} + j\mathbf{U}) \\ = (\mathbf{R} + k\mathbf{V})^* J(\hat{\mathbf{R}} + k\hat{\mathbf{V}}) + j\mathbf{U}^* J(\hat{\mathbf{R}} + k\hat{\mathbf{V}}) \\ + j(\mathbf{R} + k\mathbf{V})^* J\mathbf{U} + j\mathbf{U}^* J\mathbf{U}. \end{aligned} \quad (22)$$

Now, the antisymmetry property of J yields

$$\begin{aligned} \mathbf{U}^* J(\hat{\mathbf{R}} + k\hat{\mathbf{V}}) &= -(\hat{\mathbf{R}} + k\hat{\mathbf{V}})^* J\mathbf{U} \\ \mathbf{U}^* J\mathbf{U} &= 0. \end{aligned} \quad (23)$$

From (20), (21), (22), and (23) the following equality is deduced

$$\begin{aligned} \hat{r}_k^2 \{ (\hat{\mathbf{X}} - \mathbf{X})^* H^*(\hat{\mathbf{X}}) \}_{j_k} \\ = -(\hat{\mathbf{R}} + k\hat{\mathbf{V}} + j\mathbf{U})^* J(\mathbf{R} + k\mathbf{V} + j\mathbf{U}) \\ = -\det[(\hat{\mathbf{R}} + k\hat{\mathbf{V}} + j\mathbf{U}), (\mathbf{R} + k\mathbf{V} + j\mathbf{U})] \end{aligned}$$

or, more simply

$$\begin{aligned} \hat{r}_k^2 \{ (\hat{\mathbf{X}} - \mathbf{X})^* H^*(\hat{\mathbf{X}}) \}_{j_k} \\ = -\frac{r_k}{\hat{r}_k} \sin(\hat{\mathbf{R}} + k\hat{\mathbf{V}} + j\mathbf{U}, \mathbf{R} + k\mathbf{V} + j\mathbf{U}) \\ = \frac{r_k}{\hat{r}_k} \sin(\mathbf{R} + k\mathbf{V} + j\mathbf{U}, \hat{\mathbf{R}} + k\hat{\mathbf{V}} + j\mathbf{U}) \\ = \frac{r_k}{\hat{r}_k} \sin(\hat{\beta}_k - \beta_k) \end{aligned}$$

so that finally

$$\dot{L}(\hat{\mathbf{X}}) = -2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k). \quad (24)$$

Note the similarity of (24) with (16). Analogously to the case of a nonmaneuvering observer, $\dot{L}(\hat{\mathbf{X}})$ will be (strictly) negative if *all* the $\{\hat{\beta}_k - \beta_k\}_{k=1}^p$ lie in the interval $]-\pi, \pi[$. Therefore a local minimum of the likelihood functional cannot exist on all the corresponding $\hat{\mathbf{X}}$ domain. Note that observability is generally ensured for a leg-by-leg observer trajectory [14, 15]. Further, from (24) we note that the speed of convergence of an iterative algorithm is proportional to $\sum_{k=1}^p r_k / \hat{r}_k (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k)$.

Assuming now that the target maneuver instant is *known*, the previous analysis can be extended to the case of a *maneuvering source*. Indeed, the main change is the dimension of the state space (eqs. (2), (3)).

Consider, first, a two-leg path of the source. Opposite to the case of a nonmaneuvering source, two transition

matrices (F_1 and F_2) are now required [15]. The state equations then take the following general form:

$$\begin{cases} \mathbf{X}_1 = F_1 \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_m = F_1^m \mathbf{X}_0 \leftarrow \text{source maneuver instant} \\ \mathbf{X}_{m+1} = F_1^m F_2 \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_{m+m_{\text{obs}}+1} = F_1^m F_2^{m_{\text{obs}}+1} \mathbf{X}_0 + \mathbf{U} \\ \leftarrow \text{obs. maneuver instant} \\ \vdots \end{cases}$$

with⁴

$$\begin{aligned} F_1 &= \begin{pmatrix} Id & Id & 0 \\ 0 & Id & 0 \\ 0 & 0 & Id \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \otimes Id_2 \\ F_2 &= \begin{pmatrix} Id & 0 & Id \\ 0 & Id & 0 \\ 0 & 0 & Id \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \otimes Id_2 \\ \mathbf{X}_0 &= (r_{x,0}, r_{y,0}, v_{x,1}, v_{y,1}, v_{x,2}, v_{y,2})^* \end{aligned} \quad (25)$$

This time, the dimension of the state vector \mathbf{X} is equal to 6. The parameters $(r_{x,0}, r_{y,0})$ represent the initial source position and $(v_{x,i}, v_{y,i})_{i=1,2}$ are the components of the *relative* velocity vector on source legs 1 and 2. The matrices F_1 and F_2 are the transition matrices for respectively the source leg 1 and leg 2. Note their different structures. This formalism makes it possible to define parameterizations of the state vectors \mathbf{X}_k involving only the state vector \mathbf{X}_0 .

We state without proof the following elementary lemma.

LEMMA 3

$$\begin{aligned} \text{For } \{m, j\} \in \mathbb{N}^2, \quad F_2^j F_1^m &= \begin{pmatrix} Id & mId & jId \\ 0 & Id & 0 \\ 0 & 0 & Id \end{pmatrix} \\ \text{and: } F_2^j F_1^m &= F_1^m F_2^j. \end{aligned} \quad (26)$$

According to the new state equations, we must replace (12) by

$$\hat{\mathbf{W}}_k = M_k \hat{\mathbf{X}} \quad (27)$$

where, this time $k = m + j$, (m instant of source maneuver):

$$\begin{aligned} M_k &= \begin{pmatrix} 1 & m & j \\ m & m^2 & mj \\ j & mj & j^2 \end{pmatrix} \otimes J \\ \hat{\mathbf{X}} &= (\hat{r}_x, \hat{r}_y, \hat{v}_x^1, \hat{v}_y^1, \hat{v}_x^2, \hat{v}_y^2)^* \end{aligned}$$

⁴ m_{obs} is the maneuver instant of the observer.

In (25) the index m denotes the target maneuver instant while (\hat{r}_x, \hat{r}_y) is its initial position and $(\hat{v}_x^i, \hat{v}_y^i)$ are its velocity components on the i th leg.

As previously, the matrix M_k is antisymmetric so that the previous reasoning holds and once again gives an identical result (cf. (24)). Let us briefly detail the corresponding calculations.

First, assuming the observer *not* to be maneuvering, then

$$\hat{\mathbf{W}}_k^*(\hat{\mathbf{X}} - \mathbf{X}) = -\hat{\mathbf{X}}^* M_k \mathbf{X}. \quad (28)$$

Now

$$M_k = \begin{pmatrix} 1 \\ m \\ j \end{pmatrix} \otimes (1, m, j) \otimes J$$

so that

$$\begin{aligned} \hat{\mathbf{W}}_k^*(\hat{\mathbf{X}} - \mathbf{X}) &= -(\hat{\mathbf{R}} + m\hat{\mathbf{V}}_1 + j\hat{\mathbf{V}}_2)^* J (\mathbf{R} + m\mathbf{V}_1 + j\mathbf{V}_2) \\ &= r_k \hat{r}_k \sin(\hat{\beta}_k - \beta_k). \end{aligned} \quad (29)$$

The case of a *maneuvering observer* is treated in the same way. Let us denote m_{obs} , the instant of observer maneuver, then we have ($k' \triangleq (k - m_{\text{obs}})$):

$$\begin{aligned} \hat{r}_k^2 \{\hat{\mathbf{X}}^* H^*(\hat{\mathbf{X}})\}_k &= (\hat{\mathbf{R}}^*, \hat{\mathbf{V}}_1^*, \hat{\mathbf{V}}_2^*) \left[M_k \begin{pmatrix} \hat{\mathbf{R}} \\ \hat{\mathbf{V}}_1 \\ \hat{\mathbf{V}}_2 \end{pmatrix} + k' \begin{pmatrix} JU \\ mJU \\ jJU \end{pmatrix} \right] \\ &= k' (\hat{\mathbf{R}} + m\hat{\mathbf{V}}_1 + j\hat{\mathbf{V}}_2)^* JU. \end{aligned} \quad (30)$$

Similarly, we have

$$\begin{aligned} \hat{r}_k^2 \{\mathbf{X}^* H^*(\hat{\mathbf{X}})\}_k &= (\mathbf{R} + m\mathbf{V}_1 + j\mathbf{V}_2)^* J (\hat{\mathbf{R}} + m\hat{\mathbf{V}}_1 + j\hat{\mathbf{V}}_2) \\ &+ k' (\mathbf{R} + m\mathbf{V}_1 + j\mathbf{V}_2)^* JU \end{aligned} \quad (31)$$

so that, finally, the following equality is still valid:

$$\hat{r}_k^2 \{(\hat{\mathbf{X}} - \mathbf{X})^* H^*(\hat{\mathbf{X}})\}_k = \frac{r_k}{\hat{r}_k} \sin(\hat{\beta}_k - \beta_k).$$

Obviously, this reasoning can be also extended to any number of legs. For the sake of brevity, such straightforward extensions are not detailed. The following general result has thus been obtained.

PROPOSITION 2 *Suppose that the instants of source maneuvers are known, then the time derivative $L(\hat{\mathbf{X}})$ of the Liapunov function $L(\hat{\mathbf{X}})$ (eq. 6) is*

$$\begin{aligned} \dot{L}(\hat{\mathbf{X}}) &= 2(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}}) \\ &= -2 \sum_{k=1}^p (\hat{\beta}_k - \beta_k) I_k \\ &= -2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k). \end{aligned} \quad (32)$$

We stress that Propositions 1–3 are valid under the assumption that the source maneuver instants (or

change-points) are (exactly) known. Extensions to the case of unknown change-points are considered in Section VII. More precisely, a global optimization procedure for estimating the trajectory (including the change-points) of a maneuvering source are presented. If the change-points are *not* the exact ones, the scalar $L_k \stackrel{\Delta}{=} \hat{r}_k^2 \{(\hat{\mathbf{X}} - \mathbf{X})^* H^*(\hat{\mathbf{X}})\}_k$ (see (24)) becomes⁵

$$L_k = \frac{r_k}{\hat{r}_k} \sin(\hat{\mathbf{R}} + m' \hat{\mathbf{V}}_1 + (k - m') \hat{\mathbf{V}}_2 + j_0 \mathbf{U}, \mathbf{R} + m \mathbf{V}_1 + (k - m) \mathbf{V}_2 + j_0 \mathbf{U}). \quad (33)$$

Thus, the previous result is still valid when the change-points are not the exact ones, but the corresponding sign condition becomes more constraining (i.e., it is not the sign of $\sum_i x_i \sin(x_i)$).

REMARKS

1) We have noted the fundamental role (in convergence analysis) of the sum: $\sum_{k=1}^p (r_k / \hat{r}_k) \cdot (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k)$. A geometric interpretation of this sum is that it represents (approximately) a weighted sum of the areas of the parallelograms spanned by the vector couples $\{\hat{\mathbf{R}} + k \hat{\mathbf{V}} + j \mathbf{U}, \mathbf{R} + k \mathbf{V} + j \mathbf{U}\}$.

2) The validity of the proof of Propositions 1–3 is due to the special structure of the matrix M_k (see (25), (26)). More precisely, we have seen that

$$M_k = \begin{pmatrix} 1 & m & j \\ m & m^2 & mj \\ j & mj & j^2 \end{pmatrix} \otimes J = T_{m,j} \otimes J. \quad (34)$$

The matrix $T_{m,j}$ is the “trajectory” matrix of the target. It admits the following factorization:

$$T_{m,j} = \begin{pmatrix} 1 \\ m \\ j \end{pmatrix} \otimes (1, m, j).$$

Thus, $T_{m,j}$ is a *rank one*, semidefinite matrix, while M_k is *antisymmetric*. These two (elementary) properties are the *fundamental ingredients* of the previous calculations.

3) The uncertainty about the maneuver instant τ can be modeled by a randomization (of τ). This leads to replace the expression (4) of $p(\mathbf{B} | \mathbf{X})$ by the following:

$$p(\tilde{\mathbf{B}} | \mathbf{X}, \bar{\tau}) = \sum_i p(\tilde{\mathbf{B}} | \mathbf{X}, \tau_i(\bar{\tau})) p(\tau_i(\bar{\tau})).$$

For this modeling, the distribution of $\{\tau_i\}$ is centered around $\bar{\tau}$. The problem then consists in determining

⁵ m' assumed change-point, m actual change-point

the parameters $\{\mathbf{X}, \bar{\tau}\}$ which maximizes $p(\tilde{\mathbf{B}} | \mathbf{X}, \bar{\tau})$. This may be achieved by means of a classic iterative algorithm. The analysis of convergence is similar to the deterministic case.

4) For the 3-D TMA problem, the receiver measurements are the bearing (β_k) and the elevation angles (γ_k). The previous calculations then yield

$$\dot{L}(\hat{\mathbf{X}}) = -2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k) - 2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\gamma}_k - \gamma_k) \cos(\hat{\gamma}_k - \gamma_k).$$

5) The main result is readily extended to the case of a target with a constant acceleration vector. More generally, this result holds for any measurement (β_t) equation of the type:

$$\tan(\beta_t) = \frac{\mathbf{A}_t^* \mathbf{X}}{\mathbf{B}_t^* \mathbf{X}}$$

where \mathbf{A}_t and \mathbf{B}_t are (time-varying) vectors associated with the definition of source motion, and while \mathbf{X} is the state vector. Thus for the rectilinear, uniform motion model, $\mathbf{A}_t = (1, 0, t, 0)^*$ and $\mathbf{B}_t = (0, 1, 0, t)^*$. Extension to the more general case $f(\beta_t) = \mathbf{A}_t^* \mathbf{X} / \mathbf{B}_t^* \mathbf{X}$ requires (at least) the monotonicity of f . It is clear that the special form of $f(\beta_t)$ is instrumental.

V. DISCUSSION AND CONVERGENCE ANALYSIS

A. Unconstrained Estimation

Let us now consider the convergence of a gradient algorithm for the exact case, i.e., the $\{\beta_i\}_{i=1}^p$ are the exact ones. The convergence of a variable stepsize gradient algorithm on the compact domain Δ can be proved if the two following conditions are verified [16].

1) The functional $f(\hat{\mathbf{X}}) = -\log p(\mathbf{B} | \hat{\mathbf{X}})$ (cf. (4)) satisfies a Lipschitz condition on Δ , i.e.,

$$\|\mathbf{G}(\hat{\mathbf{X}}) - \mathbf{G}(\hat{\mathbf{Y}})\| \leq L \|\hat{\mathbf{X}} - \hat{\mathbf{Y}}\|, \quad L > 0, \quad \hat{\mathbf{X}}, \hat{\mathbf{Y}} \in \Delta. \quad (35)$$

2) For any $\varepsilon > 0$, there exists a δ ($\delta(\varepsilon) > 0$) such that

$$\|\mathbf{G}(\hat{\mathbf{X}})\| \geq \delta \quad \text{for all } \hat{\mathbf{X}} \text{ s.t. } \|\mathbf{X} - \hat{\mathbf{X}}\| > \varepsilon. \quad (36)$$

Actually, the following inequality is easily deduced (see Appendix A) from the above conditions:

$$f(\hat{\mathbf{X}}_l) - f(\hat{\mathbf{X}}_{l+1}) \geq \frac{1}{2L} \alpha_l^2 \|\mathbf{G}(\hat{\mathbf{X}}_l)\|^2$$

where⁶

$$\left\{ \begin{array}{l} L \text{ is defined by (35)} \\ \alpha_l \triangleq \frac{(\mathbf{G}^*(\hat{\mathbf{X}}_l)\mathbf{S}_l)}{\|\mathbf{G}^*(\hat{\mathbf{X}}_l)\|\|\mathbf{S}_l\|} \\ \hat{\mathbf{X}}_{l+1} = \hat{\mathbf{X}}_l - \rho_l \mathbf{S}_l. \end{array} \right. \quad (37)$$

In (37) the scalar α_l represents the cosine of the angle between the descent direction \mathbf{S}_l and the gradient \mathbf{G}_l . This allows us to consider a general form of an iterative algorithm (Gauss–Newton, pseudo-Newton, etc.). Practically, thanks to this result and the intermediate value theorem [16], a *sufficient* condition of convergence is readily deduced. Indeed, the inequality (37) asserts that the gradient is a descent direction when the conditions (35) and (36) are satisfied. Note that the difficulty of the problem stems from the nonconvexity of the functional f . Then, it is not surprising that the convergence analysis relies on the conditions (35) and (36), which may be demanding, even for a (general) convex functional.

We now restrict our attention to the above conditions. First, let us examine the condition (35). From the intermediate value theorem, we have [17]

$$\|\mathbf{G}(\hat{\mathbf{X}}) - \mathbf{G}(\hat{\mathbf{Y}})\| \leq \|\hat{\mathbf{X}} - \hat{\mathbf{Y}}\| \sup_{0 \leq t \leq 1} \|\mathcal{H}((1-t)\hat{\mathbf{X}} + t\hat{\mathbf{Y}})\| \quad (38)$$

where \mathcal{H} is the differential matrix of \mathbf{G} (i.e., the Hessian matrix).

Elementary calculations yield the matrix (denoted $D_2\beta$) of second-order partial derivatives of the bearing $\beta_k(\mathbf{X})$:

$$D_2\beta_k = \begin{pmatrix} 1 & k \\ k & k^2 \end{pmatrix} \otimes \begin{pmatrix} -a & b \\ b & a \end{pmatrix}$$

where

$$a = \frac{\sin 2\beta_k}{r_k^2}, \quad b = -\frac{\cos 2\beta_k}{r_k^2} \quad (39)$$

$$(D_2\beta)(1,3) \triangleq \frac{\partial^2 \beta}{\partial r_x \partial v_x}, \quad (D_2\beta)(2,4) \triangleq \frac{\partial^2 \beta}{\partial r_y \partial v_y}, \quad \text{etc.}$$

From (9) and (39), the matrix \mathcal{H} is easily deduced, yielding⁷

$$\mathcal{H}(\hat{\mathbf{X}}) = \left(\frac{\partial \hat{\mathbf{B}}}{\partial \hat{\mathbf{X}}} \right)^* \left(\frac{\partial \hat{\mathbf{B}}}{\partial \hat{\mathbf{X}}} \right) + (\mathbf{C}_1 \mid \mathbf{C}_2 \mid \mathbf{C}_3 \mid \mathbf{C}_4)$$

⁶ \mathbf{S}_l is the descent direction.

⁷From (40), we note that $\mathcal{H}(\hat{\mathbf{X}})$ is the sum of a nonnegative definite matrix $((\partial \hat{\mathbf{B}}/\partial \hat{\mathbf{X}})^*(\partial \hat{\mathbf{B}}/\partial \hat{\mathbf{X}}))$ and another matrix which is not generally nonnegative definite. Hence, it is impossible to obtain a general conclusion about the positive definiteness of the Hessian matrix $\mathcal{H}(\hat{\mathbf{X}})$.

where

$$\frac{\partial \hat{\mathbf{B}}}{\partial \hat{\mathbf{X}}} = H(\hat{\mathbf{X}}) \quad (\text{cf. (9)})$$

$$\mathbf{C}_1 = \begin{pmatrix} \frac{\sin 2\hat{\beta}_1}{\hat{r}_1^2} & \cdots & \cdots & \frac{\sin 2\hat{\beta}_p}{\hat{r}_p^2} \\ \frac{\cos 2\hat{\beta}_1}{\hat{r}_1^2} & \cdots & \cdots & \frac{\cos 2\hat{\beta}_p}{\hat{r}_p^2} \\ \frac{\sin 2\hat{\beta}_1}{\hat{r}_1^2} & \cdots & \cdots & \frac{p \sin 2\hat{\beta}_p}{\hat{r}_p^2} \\ \frac{\cos 2\hat{\beta}_1}{\hat{r}_1^2} & \cdots & \cdots & \frac{p \cos 2\hat{\beta}_p}{\hat{r}_p^2} \end{pmatrix} (\tilde{\mathbf{B}} - \hat{\mathbf{B}}). \quad (40)$$

Expressions of vectors \mathbf{C}_i stem from (39). We deduce from (9) and (39), (40) that $\|\mathcal{H}((1-t)\hat{\mathbf{X}} + t\hat{\mathbf{Y}})\|$ is bounded on any compact subset of \mathbb{R}^4 excluding the origin, hence the Lipschitz condition (35) holds on such a subset.

Let us now examine the second condition (36). From the main results (i.e., Property 1 and 2), we have

$$(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}}) = \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k). \quad (41)$$

The following inequality is then deduced from (41) and the Cauchy–Schwartz inequality [9]:

$$\begin{aligned} \|\mathbf{G}(\hat{\mathbf{X}})\| &\geq \frac{1}{\|\hat{\mathbf{X}} - \mathbf{X}\|} \left| \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k) \right| \\ &\geq \frac{1}{M_\Delta} \left| \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k) \right|. \end{aligned} \quad (42)$$

Now since the compact domain Δ does not contain the origin and since the sum $\sum_{k=1}^p r_k/\hat{r}_k (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k)$ admits a strictly positive lower bound when the condition $\|\hat{\mathbf{X}} - \mathbf{X}\| \geq \varepsilon$ holds, then (36) holds true.

Up to now our attention has been restricted to the case of noise-free measurements. Let us now consider *noisy measurements*, (3) becomes

$$\tilde{\beta}_k = \tan^{-1} \left(\frac{r_{x,k}}{r_{y,k}} \right) + w_k \quad (43)$$

w_k : wgn.

Denote $\tilde{\mathbf{X}}$ the state vector which maximizes the likelihood functional (4), then $L(\hat{\mathbf{X}})$ (eq. (6)) is naturally replaced by

$$L'(\hat{\mathbf{X}}) = \|\hat{\mathbf{X}} - \tilde{\mathbf{X}}\|^2. \quad (44)$$

Using the previous calculations, we then obtain

$$(\hat{\mathbf{X}} - \tilde{\mathbf{X}})^* \mathbf{G}(\hat{\mathbf{X}}) = -2 \sum_{k=1}^p \frac{\tilde{r}_k}{r_k} (\hat{\beta}_k - \tilde{\beta}_k) \sin(\hat{\beta}_k - \tilde{\beta}_k), \quad (45)$$

$$\text{where: } \tilde{r}_k \triangleq r_k(\tilde{\mathbf{X}}) \quad \text{and} \quad \tilde{\beta}_k \triangleq \beta_k(\tilde{\mathbf{X}}).$$

The analysis of a gradient method is thus quite similar to the previous one. In particular (45) is a sufficient condition which guarantees the *numerical* convergence towards $\tilde{\mathbf{X}}$ (for convenient stepsizes). It remains to examine the consistency of the estimator $\tilde{\mathbf{X}}$. Usually, such analysis is conducted under the assumption that observations are identically distributed.

Here, this assumption *is not valid*. This problem has been considered in [20, pp. 186–191]. With our notations, it has been shown that consistency is ensured if the following conditions are satisfied⁸ ($\forall \mathbf{X}, \mathbf{X}'$):

$$\begin{aligned} \lim_{p \rightarrow \infty} \frac{1}{d_p^2} \left(\sum_{k=1}^p \mathbf{H}_k(\mathbf{X}') \mathbf{H}_k^*(\mathbf{X}') \right) &= \mathbf{O} \\ \lim_{p \rightarrow \infty} \frac{1}{d_p} \left(\sum_{k=1}^p (\beta_k(\mathbf{X}) - \beta_k(\mathbf{X}')) \mathbf{H}_k(\mathbf{X}) \right) & \end{aligned} \quad (46)$$

exists and is non zero.

In (46), $\{d_p\}$ is an arbitrary sequence of positive numbers, while \mathbf{H}_k is the k th column of the matrix H (i.e., $\mathbf{H}_k^* = 1/r_k (\cos(\beta_k), -\sin(\beta_k), k \cos(\beta_k), -k \sin(\beta_k))$). For the sequence $\{d_p\} = \{p^2\}$, the first condition of (46) is thus satisfied. The second stems from the form of $\mathbf{H}_k(\mathbf{X})$ (see (9)). When p tends towards infinity, the bearing-rate approaches zero thus implying that the second condition holds. However, note that the true statistical problem is to determine how far $\tilde{\mathbf{X}}$ is from the exact state vector \mathbf{X} . But, this is beyond the scope of this work and we refer to [1, 10].

B. Constrained Estimation

It is easily seen [1, 10] that the possibility of large errors grows with the range-to-baseline ratio. When unreasonable estimates are obtained, additional information may be available which, if properly used, could improve the TMA solution. Perhaps the two most readily available constraints on the state vector are relative to target range and speed.⁹ More specifically, it is quite reasonable to assume that the state vector components are bounded. This leads to consider the following type (*primal problem P*) of constrained optimization problems:

$$\begin{cases} \text{Minimize} & f(\mathbf{X}) = \|\tilde{\mathbf{B}} - \mathbf{B}(\hat{\mathbf{X}})\|_{\Sigma^{-1}}^2 \\ \text{subject to} & g(\mathbf{X}) \leq 0. \end{cases} \quad (47)$$

⁸ \mathbf{O} is the zero matrix.

⁹A pioneering work about this problem can be found in [1, 18].

In (47), the m -dimensional function g is generally linear and represents the velocity and position constraints (e.g., $-v_{\max} \leq v_x, v_y \leq v_{\max}$). This problem can be efficiently solved by means of Lagrangian duality [13], leading to consider the following Lagrangian *dual problem*:

$$\begin{cases} \text{Maximize} & \theta(\mathbf{U}) \\ \text{subject to} & \mathbf{U} \geq \mathbf{0} \\ \text{where:} & \theta(\mathbf{U}) = \inf_{\mathbf{X}} \{f(\mathbf{X}) + \mathbf{U}^* g(\mathbf{X})\}. \end{cases} \quad (48)$$

Numerous methods are available for solving the dual problem. Among them, one of the more feasible is the cutting plane method, detailed in [13, pp. 224–230]. However, for our application, Uzawa's method has attractive features. In its spirit, it consists in solving the (difficult) constrained problem by means of a sequence of unconstrained problems. Let us briefly recall the principle of this method. A couple $(\mathbf{X}_k, \mathbf{U}_k)$ is recursively defined by¹⁰

$$\begin{cases} \mathbf{X}_k = \arg \inf_{\mathbf{X}} [f(\mathbf{X}) + \mathbf{U}_k^* g(\mathbf{X})] \\ u_{i,k+1} = \max\{u_{i,k} + \rho g_i(\mathbf{X}_k), 0\}, \quad 1 \leq i \leq m \end{cases} \quad (49)$$

where ρ is the algorithm stepsize, u_i and g_i are the i th components of \mathbf{U} and $g(\mathbf{X})$, the vector \mathbf{U}_k is the vector of Lagrangian multipliers at the iteration k . Thus, $u_{i,k+1}$ may be interpreted as the projection of the element $u_{i,k} + \rho g_i(\mathbf{X}_k)$ on \mathbb{R}^{+m} (i.e., $u_{k+1} = P_+(u_k + \rho g(\mathbf{X}_k))$).

The sum of a quasi-convex (f) and a convex function being quasi-convex, the convergence of an iterative algorithm is a consequence of the previous results. Convergence towards the exact state vector can occur only if this state vector belongs to the constraint subset.

Now, invoking the strong duality theorem [13], there exists a saddle point (let us denote \mathbf{X} and \mathbf{U} the corresponding values of the state vectors and multipliers) for which the values of the primal and dual problems are equal (no duality gap). Assuming, further, the constraints linear and since the projection operator does not increase the distances ($g(\mathbf{X}) = C\mathbf{X}$), we then obtain [19]

$$\begin{cases} \nabla f(\mathbf{X}_k) - \nabla f(\mathbf{X}) + C^*(\mathbf{U}_k - \mathbf{U}) = \mathbf{0} \\ \|\mathbf{U}_{k+1} - \mathbf{U}\| \leq \|\mathbf{U}_k - \mathbf{U} + \rho C(\mathbf{X}_k - \mathbf{X})\|. \end{cases} \quad (50)$$

From (50), the following inequality is straightforwardly deduced [19]

$$\begin{aligned} \|\mathbf{U}_{k+1} - \mathbf{U}\|^2 &\leq \|\mathbf{U}_k - \mathbf{U}\|^2 + \rho^2 \|C(\mathbf{X}_k - \mathbf{X})\|^2 \\ &\quad - 2\rho (\nabla f(\mathbf{X}_k) - \nabla f(\mathbf{X}), \mathbf{X}_k - \mathbf{X}). \end{aligned} \quad (51)$$

Usually, an upper bound of the scalar product $(\nabla f(\mathbf{X}_k) - \nabla f(\mathbf{X}), \mathbf{X}_k - \mathbf{X})$ is a direct consequence

¹⁰In this section k denotes the index of the iteration.

of the convexity (and more precisely coercitivity) of f . We cannot use this argument here. Instead we can use the previous results for calculating an explicit expression of this scalar product, i.e.,

$$\begin{aligned} & (\nabla f(\mathbf{X}_k) - \nabla f(\mathbf{X}), \mathbf{X}_k - \mathbf{X}) \\ &= 2 \sum_i \left\{ (\beta_i - \beta_{i,k}) \sin(\beta_i - \beta_{i,k}) \left(\frac{r_i}{r_{i,k}} + \frac{r_{i,k}}{r_i} \right) \right\}. \end{aligned} \quad (52)$$

Invoking (52), it is always possible to determine an upper bound of the stepsize ρ such that the scalar product $(\nabla f(\mathbf{X}_k) - \nabla f(\mathbf{X}), \mathbf{X}_k - \mathbf{X})$ is bounded by a given scalar α . We then have

$$\begin{aligned} \|\mathbf{U}_{k+1} - \mathbf{U}\|^2 &\leq \|\mathbf{U}_k - \mathbf{U}\|^2 \\ &\quad - \rho \{2\alpha - \rho \|C\|^2\} \|\mathbf{U}_k - \mathbf{U}\|^2. \end{aligned} \quad (53)$$

Now, if the stepsize ρ is such that $0 \leq \rho \leq 2\alpha/\|C\|^2$, then the sequence $\{\|\mathbf{U}_k - \mathbf{U}\|\}_k$ is decreasing. Therefore, this sequence will converge for a convenient choice of the stepsize. The convergence of the sequence $\|\mathbf{X} - \mathbf{X}_k\|$ is then a direct consequence of (53) and the strong duality theorem [13].

In practice, however, the adequacy of the stepsize is a salient feature of this method. So, it is worth using (52) and (53) in order to obtain an acceptable upper bound of ρ .

VI. ANALYSIS OF INSTRUMENTAL VARIABLE METHOD

In the BOT context a classical variant of the gradient algorithm is the modified instrumental variable (MIV) method. It has been carefully studied by various researchers [1, 3, 6].

This method relies upon a quite subtle reformulation of the measurement equation which results in a pseudomeasurement derived from the known observer state and the available measured bearings, and is ‘‘linearly’’ related to the target state. In this spirit the more direct approach is the pseudolinear estimate (PLE). However, the PLE method is known to produce a (very) biased state estimate [1, 3].

An MIV method has been examined that uses successive state estimates to generate the instrumental variables in an iterative algorithm [21, 22]. By using the so-called instruments [22, 23] which in our scheme are the predictions of the values of the present regressors based on past measurements, consistent estimates can be obtained.

Let us now recall the general form of the MIV in the BOT context. Given the linear dynamic system and measurement equation ((1) and (2)), the pseudomeasurements z_k are defined by

$$z_k = r_{\text{obs},x}(k) \cos \tilde{\beta}_k - r_{\text{obs},y}(k) \sin \tilde{\beta}_k \quad (54)$$

where

- 1) $\tilde{\beta}_k$ is estimated bearings,
- 2) $(r_{\text{obs},x}(k), r_{\text{obs},y}(k))$ is Cartesian coordinates of the observer at time k .

Useful relations are then [1, 3]

$$z_k = \tilde{\mathbf{A}}^*(k) \mathbf{X}_{\text{obs}}(k) = \tilde{\mathbf{A}}^*(k) \mathbf{X}_T(k) + \eta_k$$

where:

$$\begin{cases} \tilde{\mathbf{A}}^*(k) = (\cos \tilde{\beta}_k, -\sin \tilde{\beta}_k, 0, 0), \\ \eta_k = r_k \sin w_k \simeq r_k w_k, \\ w_k : \text{measurement noise}, \\ \mathbf{X}_{\text{obs}} : \text{observer state}, \mathbf{X}_T : \text{target state}. \end{cases} \quad (55)$$

Given (55) the problem of determining the target state from an history \mathbf{Z}_p of pseudomeasurements ($\mathbf{Z}_p = (z_1, z_2, \dots, z_p)^*$) takes the form of a linear state estimation technique. The vector \mathbf{Z}_p can be written as

$$\mathbf{Z}_p = \tilde{\mathcal{A}}_p \mathbf{X}_T(0) + \mathbf{N}$$

where

$$\begin{cases} k\text{th row of } \tilde{\mathcal{A}}_p = \tilde{\mathbf{A}}^*(k) F^k \\ = (\cos \tilde{\beta}_k, -\sin \tilde{\beta}_k, k \cos \tilde{\beta}_k, -k \sin \tilde{\beta}_k) \\ \mathbf{N} = (\eta_1, \dots, \eta_p)^*. \end{cases} \quad (56)$$

Minimizing the norm of the error $e(e = \mathbf{Z}_p - \tilde{\mathcal{A}}_p \mathbf{X}_T)$ yields the PLE but a major drawback of this estimator is its bias due to the correlation between the terms of $\tilde{\mathcal{A}}_p$ and \mathbf{N} . The MIV approach consists in replacing the classical optimality equation in the minimization of e^2 by the following one:

$$\mathcal{A}_p^*(\hat{\mathbf{X}}) [\mathbf{Z}_p - \tilde{\mathcal{A}}_p \hat{\mathbf{X}}_T] = \mathbf{0}. \quad (57)$$

In (57) $\mathcal{A}_p^*(\hat{\mathbf{X}})$ is the \mathcal{A} matrix associated with $\hat{\mathbf{X}}$. An iterative estimator of the $\{target\}$ state vector is deduced from (57) and stands as follows:

$$\hat{\mathbf{X}}_{T,\ell+1} = (\mathcal{A}_p^*(\hat{\mathbf{X}}_\ell) \tilde{\mathcal{A}}_p)^{-1} \mathcal{A}_p^*(\hat{\mathbf{X}}_\ell) \mathbf{Z}_p$$

or taking into account the variations of the source-observer range ($\hat{R} = \text{diag}(\{r_i\}_{i=1}^p)$):

$$\hat{\mathbf{X}}_{T,\ell+1} = (\mathcal{A}_p^*(\hat{\mathbf{X}}_\ell) \hat{R}^{-2} \tilde{\mathcal{A}}_p)^{-1} (\mathcal{A}_p^*(\hat{\mathbf{X}}_\ell) \hat{R}^{-2} \mathbf{Z}_p). \quad (58)$$

First let us examine the unicity of the MIV estimate. Let \mathbf{X} be the exact source state vector then in the exact case (no measurement noise) (57) becomes

$$[\mathcal{A}_p^*(\hat{\mathbf{X}}) \mathcal{A}_p(\mathbf{X})] (\mathbf{X}_T - \hat{\mathbf{X}}_T) = \mathbf{0}. \quad (59)$$

Invoking the following classical lemma [9].

LEMMA 4 Let A and B be respectively a $(m \times k)$ matrix and a $(k \times m)$ matrix, then the following inequality holds:

$$\begin{aligned} &(\text{rank } A + \text{rank } B) - k \\ &\leq \text{rank } (AB) \leq \min(\text{rank } A, \text{rank } B). \end{aligned}$$

Applied to $A = \mathcal{A}_p^*(\hat{\mathbf{X}})$ and $B = \mathcal{A}_p(\mathbf{X})$, this lemma implies the unicity of the MIV solution when the BOT problem is observable.

Consider now the noisy case, using the previous calculations (see Sections III, IV), we obtain the following result.

PROPOSITION 3 The following equation holds:

$$\begin{aligned} &(\hat{\mathbf{X}} - \mathbf{X})^* \mathcal{A}_p^*(\hat{\mathbf{X}}) [\mathbf{Z}_p - \mathcal{A}_p(\hat{\mathbf{X}}) \mathbf{X}_T] \\ &= \sum_{k=1}^p r_k \hat{r}_k [\sin(\beta_k - \hat{\beta}_k)]^2. \end{aligned} \quad (60)$$

The conclusion is thus quite identical to the convergence analysis of the gradient method.

It remains to consider the behavior of the MIV algorithm itself. For that aim, we consider the function f defined below¹¹

$$f(\hat{\mathbf{X}}_T) = [\mathcal{A}_p^*(\hat{\mathbf{X}}) \tilde{\mathcal{A}}_p]^{-1} \mathcal{A}_p^*(\hat{\mathbf{X}}) \mathbf{Z}_p$$

now [1, 3, 6]

$$\mathbf{Z}_p = \tilde{\mathcal{A}}_p \mathbf{X}_T + \mathbf{N}$$

so that

$$f(\hat{\mathbf{X}}_T) = \mathbf{X}_T + (\mathcal{A}_p^*(\hat{\mathbf{X}}) \tilde{\mathcal{A}}_p)^{-1} \mathcal{A}_p^*(\hat{\mathbf{X}}) \mathbf{N}. \quad (61)$$

Thanks to the fixed point theorem [16, 17] the convergence of the MIV algorithm will be guaranteed if we can prove that f is k -Lipschitz ($k < 1$) (see (35)). Indeed, the fixed point theorem will imply the convergence of the sequence $\mathbf{X}_{k+1} = f(\mathbf{X}_k)$.

The two following hypotheses are usual¹² [21–23]:

$$\begin{aligned} &P \left(\lim_{p \rightarrow \infty} \frac{1}{p} \mathcal{A}_p^*(\hat{\mathbf{X}}) \mathbf{N} \right) = \mathbf{0} \\ &P \left(\lim_{p \rightarrow \infty} \frac{1}{p} \mathcal{A}_p^*(\hat{\mathbf{X}}) \tilde{\mathcal{A}}_p \right) \text{ is non singular.} \end{aligned} \quad (62)$$

The second condition is guaranteed if the observability conditions [1, 6, 15] are satisfied. For instance, if an observer maneuver occurs. The first one results from [22, Theorem 2, p. 709] which states that (62) holds if the elements of $\mathcal{A}_p^*(\hat{\mathbf{X}})$ are uniformly bounded and statistically independent of the elements of N .

¹¹ $\tilde{\mathcal{A}}$ is the matrix \mathcal{A} where β_k is replaced by $\hat{\beta}_k$.

¹²The symbol P here denotes the probability.

Then from Slutski's theorem [22, 25] we have

$$\begin{aligned} &P \left[\lim_{p \rightarrow \infty} (\mathcal{A}_p^*(\hat{\mathbf{X}}) \tilde{\mathcal{A}}_p)^{-1} \mathcal{A}_p^*(\hat{\mathbf{X}}) \mathbf{N} \right] \\ &= P \left[\left(\lim_{p \rightarrow \infty} \frac{1}{p} \mathcal{A}_p^*(\hat{\mathbf{X}}) \tilde{\mathcal{A}}_p \right)^{-1} \right] P \left[\lim_{p \rightarrow \infty} \frac{1}{p} \mathcal{A}_p^*(\hat{\mathbf{X}}) \mathbf{N} \right] = \mathbf{0}. \end{aligned} \quad (63)$$

so that $\hat{\mathbf{X}}$ defined by (61) is a consistent estimator.

VII. UNKNOWN CHANGE-POINTS OF SOURCE TRAJECTORY

Although it has been more than 25 years since Mc Aulay and Denlinger first published their paper on tracking a maneuvering target [26], this problem still motivates a great deal of fruitful research. For instance, see [27] and the related references. However, a large part of these efforts is devoted to radar applications.

The sonar context is rather different. Sudden maneuvers and leg-by-leg source trajectory is a realistic hypothesis in the sonar context. Also, the observations are nonlinearly depending of the source state. Moreover, the detection of source maneuvers requires a suitable estimation of the state, itself needing a sufficient signal-to-noise ratio and a sufficient number of measurements. All these considerations plead for a global modeling of the source trajectory, including its maneuvers and their detection process. As noticed by Bogler [27] and accordingly to our analysis, the complete solution of this problem must be divided into consecutive steps. First, the maneuver must be detected. Second, the source trajectory must be estimated. Third, after detection and correction, the source trajectory parameters are adjusted in anticipation of future target maneuvers.

We examine now the case of a maneuvering target with unknown instant of velocity change. Assuming this instant (τ) real,¹³ a direct calculation yields (2 legs, $\mathbf{X} = (r_x, r_y, v_{x,1}, \dots, v_{y,2}, \tau)^*$):

$$\begin{aligned} \dot{L}(\hat{\mathbf{X}}) &= -2 \sum_{k=1}^p \frac{r_k}{\hat{r}_k} (\hat{\beta}_k - \beta_k) \sin(\hat{\beta}_k - \beta_k) \\ &\quad - 2(\hat{\tau} - \tau) \sum_{k=1}^p (\hat{\beta}_k - \beta_k) \frac{\partial \hat{\beta}_k}{\partial \tau} \end{aligned} \quad (64)$$

$$\text{where } \frac{\partial \hat{\beta}_k}{\partial \tau} = \frac{\|\hat{\mathbf{V}}_1 - \hat{\mathbf{V}}_2\|}{\hat{r}_k} \sin(\hat{\mathbf{V}}_1 - \hat{\mathbf{V}}_2, \hat{\mathbf{R}}).$$

From (64) we see that, unlike the case of known maneuver instants, it is not possible to give simple conditions ensuring that $L(\hat{\mathbf{X}})$ is non-zero. The main

¹³Differentiability w.r.t. τ is thus assumed.

result must be replaced by less affirmative discussions. In particular, we see from (64) that when $(\hat{\tau} - \tau)$ is small in regard to 1 ($\alpha = 1$), then the second term of (64) may be neglected in regard to the first one. This result means that the main results are still valid as long as the error $(\hat{\tau} - \tau)$ is sufficiently small.

Obviously, the maneuver instant may be included in the set of parameters and an iterative (e.g., gradient) method can be easily derived. However, this approach suffers from a major drawback: the presence of local minima of the likelihood functional. Instead of it, a generalized likelihood ratio (GLR) approach is quite preferable. In fact, this approach consists in combining an exhaustive search for the change-points (τ) with an iterative method for estimating the source kinematic parameters (r_x, r_y, v_x, v_y) . Sequential detection of source maneuvers will render the approach feasible. The convergence of the method stems from the GLR definition and the previous results.

Consider, for instance, a two-leg source trajectory, the GLR takes the following form:

$$\Lambda(\tilde{\mathbf{B}}) = \frac{\sup_{\mathbf{R}, \mathbf{V}} p(\tilde{\mathbf{B}} | \mathbf{R}, \mathbf{V})}{\sup_{\mathbf{R}, \mathbf{V}_1, \mathbf{V}_2, \tau} p(\tilde{\mathbf{B}} | \mathbf{R}, \mathbf{V}_1, \mathbf{V}_2, \tau)}. \quad (65)$$

Maximization relatively to $\{\mathbf{R}, \mathbf{V}_1, \mathbf{V}_2\}$ is achieved by means of iterative methods (gradient or MIV) with τ fixed. The change point τ itself is estimated by exhaustive search in an “acceptable” interval. Notice that this approach provides simultaneously estimates of $\mathbf{R}, \mathbf{V}_i, \tau$ and detection of the source maneuvers and avoids numerical problems (local minima).

Practically, it performs very well, allowing us to obtain a very reliable detection of the source maneuver as well as an accurate estimation of the maneuver instant.¹⁴ An important drawback may be its computation cost since it needs multiple maximizations.

To overcome this problem,¹⁵ an approximate localization of the source maneuvers can be obtained by sequential detection [20] based, for instance, on the following type of test [28] which may be viewed as an approximation¹⁶ of a likelihood ratio around the parameter $\theta_0 \in \Theta_0$:

$$\Lambda_n(\tilde{\mathbf{B}}_2) = (1/n)\mathbf{e}_2^* U_2^{-1} \mathbf{e}_2 \quad (66)$$

¹⁴Approximating the asymptotic performance of this test is a difficult task, since the sequence of observations is *not* identically distributed.

¹⁵Computation cost of an exhaustive search for τ .

¹⁶More precisely, denoting $\hat{\theta}$ the MLE over Θ and θ^* the MLE over Θ_0 we consider the following expansion of the likelihood: $l(\theta^*) = l(\hat{\theta}) + l_{\hat{\theta}}(\theta^* - \hat{\theta}) - (\theta^* - \hat{\theta})^* I(\hat{\theta})(\theta^* - \hat{\theta})$, $I(\hat{\theta})$ Fisher matrix.

where

$$\begin{aligned} \mathbf{e}_2 &\triangleq \tilde{\mathbf{B}}_2 - \mathbf{B}_2(\hat{\mathbf{X}}_1) \\ U_2 &= \Sigma_2 + H_2(\hat{\mathbf{X}}_1)F_1^{-1}H_2^*(\hat{\mathbf{X}}_1) \\ n &= \dim(\mathbf{e}_2). \end{aligned}$$

In (66), $\hat{\mathbf{X}}_1$ represents the source state vector estimated using the first batch of measurements ($\tilde{\mathbf{B}}_1$) of the first source leg which assumes that no source change-point occurs during this first set. The vector $\mathbf{B}_2(\hat{\mathbf{X}}_1)$ is the vector of extended (from \mathbf{X}_1) bearings on the second set of measurements while $\tilde{\mathbf{B}}_2$ is the second set of measurements. The matrix U_2 is the covariance matrix of \mathbf{e}_2 where the matrix $H_2(\hat{\mathbf{X}}_1)$ is defined in (9).

Actually, the expression of Λ_n (66) stems from the following considerations

$$\mathbf{e}_2 = \underbrace{\tilde{\mathbf{B}}_2 - \mathbf{B}_2(\mathbf{X}_2)}_{\text{meas. noise}} + \underbrace{\mathbf{B}_2(\mathbf{X}_2) - \mathbf{B}_2(\hat{\mathbf{X}}_1)}_{\text{mod. error noise}}$$

so that

$$\begin{aligned} \text{cov}(\mathbf{e}_2) &= \text{cov}[\tilde{\mathbf{B}}_2 - \mathbf{B}_2(\mathbf{X}_2)] + \text{cov}[\mathbf{B}_2(\mathbf{X}_2) - \mathbf{B}_2(\hat{\mathbf{X}}_1)] \\ &= \Sigma_2 + H_2(\hat{\mathbf{X}}_1)F_1^{-1}H_2^*(\hat{\mathbf{X}}_1). \end{aligned} \quad (67)$$

The equality $\text{cov}[\mathbf{B}_2(\mathbf{X}_2) - \mathbf{B}_2(\hat{\mathbf{X}}_1)] = H_2(\hat{\mathbf{X}}_1) \cdot F_1^{-1}H_2^*(\hat{\mathbf{X}}_1)$ stems from a theorem of Cramér [25, (p. 47)] and the implicit function theorem [17]. Let us denote H_0 and H_1 the hypotheses associated, respectively, with no source maneuver and a maneuver. Under H_0 , $n\Lambda$ is (asymptotically) distributed as a central chi-squared random variable χ_n^2 , while under H_1 it is distributed as a noncentral chi-squared $\chi_n^2(\delta)$ where the noncentrality parameter δ is

$$\delta = (\mathbf{B}_2 - \mathbf{B}_2(\mathbf{X}_1))^* U_2^{-1} (\mathbf{B}_2 - \mathbf{B}_2(\mathbf{X}_1)). \quad (68)$$

The power of the test may then be easily derived [20, 29] from (66) and (68). Classical calculations yield (η : value of the threshold):

$$\begin{cases} P_{\text{fa}} = \exp(-\eta/2) \sum_{i=0}^{n/2-1} \frac{(\eta/2)^i}{i!} \\ P_{\text{d}} = 1 - \exp\left(-\frac{\eta + \delta}{2}\right) \sum_{j=0}^{\infty} \frac{\delta^j}{j!} \left[\sum_{i=0}^{(n/2)+j-1} \frac{(\eta/2)^i}{i!} \right]. \end{cases} \quad (69)$$

Using the Neyman–Pearson lemma, the above formula allows us to determine the threshold η for a given level of false alarm. The receiver operating characteristic (ROC) curves [20] of the test may be calculated from the above equations. However, the convergence of the infinite sum may be problematic when δ is great. From (69), we note that the probability of detection is tightly related to the value of δ . As δ is (roughly speaking) proportional to n^3 ,

the test performance is fastly increasing with n . For instance, even a small change in the source trajectory may be detected, with a quite acceptable confidence, for values of n as low as 4. However, practically, this test is also quite sensitive to the presence of a (measurement) bias. This may be a serious drawback, which reinforces the importance of the registration step [11, 28].

VIII. SOME EXTENSIONS

A. Multiple Receivers

Assuming that the (multiple) measurements are independent at each scan (m receivers), the scalar product $(\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}(\hat{\mathbf{X}})$ is replaced by

$$\sum_{j=1}^m (\hat{\mathbf{X}} - \mathbf{X})^* \mathbf{G}_j(\hat{\mathbf{X}})$$

yielding [7, 10]:

$$\dot{L}(\hat{\mathbf{X}}) = -2 \sum_{j=1}^m \sum_{k=1}^p \frac{r_k^j}{\hat{r}_k^j} (\hat{\beta}_k^j - \beta_k^j) \sin(\hat{\beta}_k^j - \beta_k^j) \quad (70)$$

$1 \leq j \leq m$: index of the j th subarray. The previous convergence analysis may thus be straightforwardly extended to multiple receivers, as well as to registration problems [11, 12].

B. Multiple Source Case

Let us now consider the case of multiple sources. First, for the sake of brevity, our analysis is restricted to the case of two sources and two independent measurements. The state vector is made of the source state vectors (\mathbf{X}_1 and \mathbf{X}_2) and the mixture parameters (π_1 and $\pi_2 = 1 - \pi_1$). At each scan, two measurement models are observed; with an assignment probability π_1 , resp. π_2 . The likelihood functional $L(p)$ (p scans) then takes the following form [30, 31]:

$$\begin{aligned} L(p) &= \prod_{t=1}^p \left\{ \prod_{j=1}^2 (\pi_1 p(\tilde{\beta}_{j,t} | \mathbf{X}_1) + \pi_2 p(\tilde{\beta}_{j,t} | \mathbf{X}_2)) \right\} \\ &= \prod_{t=1}^p L_t \end{aligned} \quad (71)$$

where $\tilde{\beta}_{1,t}$ and $\tilde{\beta}_{2,t}$ are the measurements for scan t .

Let us now briefly detail the term L_t . Omitting the time index (in L_t), we have

$$\begin{aligned} L_t &= (\pi_1 p(\tilde{\beta}_1 | \mathbf{X}_1) + \pi_2 p(\tilde{\beta}_1 | \mathbf{X}_2)) \\ &\quad \times (\pi_1 p(\tilde{\beta}_2 | \mathbf{X}_1) + \pi_2 p(\tilde{\beta}_2 | \mathbf{X}_2)) \\ &\triangleq p(\tilde{\beta}_1, \mathbf{X}_1, \mathbf{X}_2) p(\tilde{\beta}_2, \mathbf{X}_1, \mathbf{X}_2). \end{aligned} \quad (72)$$

Quite analogously to the unique source case, we obtain $(p'(\beta_j | \mathbf{X}_1) = (1/\sigma)p(\beta_j | \mathbf{X}_1))$

$$\begin{aligned} r_1^2 \nabla_{\mathbf{X}_1} L_t &= p(\tilde{\beta}_2, \mathbf{X}_1, \mathbf{X}_2) p'(\tilde{\beta}_1 | \mathbf{X}_1) \pi_1 [\tilde{\beta}_1 - \beta(\mathbf{X}_1)] \mathbf{W}_1 \\ &\quad + p(\tilde{\beta}_1, \mathbf{X}_1, \mathbf{X}_2) p'(\tilde{\beta}_2 | \mathbf{X}_1) \pi_1 [\tilde{\beta}_2 - \beta(\mathbf{X}_1)] \mathbf{W}_1. \end{aligned} \quad (73)$$

From (73) and Properties 1 and 2, the following equalities are straightforwardly deduced

$$\begin{aligned} &\left(\sigma^2 \pi_1^{-1} \frac{\hat{r}_1}{r_1} \right) (\hat{\mathbf{X}}_1 - \mathbf{X}_1)^* \nabla_{\hat{\mathbf{X}}_1} L_t \\ &= p(\tilde{\beta}_2, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2) p(\tilde{\beta}_1 | \hat{\mathbf{X}}_1) (\tilde{\beta}_1 - \hat{\beta}_1) \sin(\tilde{\beta}_1 - \hat{\beta}_1) \\ &\quad + p(\tilde{\beta}_1, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2) p(\tilde{\beta}_2 | \hat{\mathbf{X}}_1) (\tilde{\beta}_2 - \hat{\beta}_1) \sin(\tilde{\beta}_2 - \hat{\beta}_1) \end{aligned}$$

and, similarly

$$\begin{aligned} &\left(\sigma^2 \pi_2^{-1} \frac{\hat{r}_2}{r_2} \right) (\hat{\mathbf{X}}_2 - \mathbf{X}_2)^* \nabla_{\hat{\mathbf{X}}_2} L_t \\ &= p(\tilde{\beta}_2, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2) p(\tilde{\beta}_1 | \hat{\mathbf{X}}_2) (\tilde{\beta}_1 - \hat{\beta}_2) \sin(\tilde{\beta}_2 - \hat{\beta}_2) \\ &\quad + p(\tilde{\beta}_1, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2) p(\tilde{\beta}_2 | \hat{\mathbf{X}}_2) (\tilde{\beta}_2 - \hat{\beta}_2) \sin(\tilde{\beta}_2 - \hat{\beta}_2). \end{aligned} \quad (74)$$

We thus have finally:

$$\begin{aligned} &(\hat{\mathbf{X}}_1 - \mathbf{X}_1)^* \nabla_{\hat{\mathbf{X}}_1} L(T) \\ &= \sum_t L_1 \cdots L_{t-1} (\hat{\mathbf{X}}_1 - \mathbf{X}_1)^* \nabla_{\hat{\mathbf{X}}_1} L(t) \cdot L_{t+1} \cdots L_T \end{aligned} \quad (75)$$

where the term $(\hat{\mathbf{X}}_1 - \mathbf{X}_1)^* \nabla_{\hat{\mathbf{X}}_1} L_t$ is given by (74).

Equation (74) extends the main results to the case of a multiple source mixture. However, the conclusions are much less affirmative. In fact, let us consider the sign of the elementary term $(\hat{\mathbf{X}}_1 - \mathbf{X}_1)^* \nabla_{\hat{\mathbf{X}}_1} L_t$.

Analogously to the unique source case, the term $(\tilde{\beta}_1 - \hat{\beta}_1) \sin(\tilde{\beta}_1 - \hat{\beta}_1)$ is positive but the trouble comes from the term $(\tilde{\beta}_2 - \hat{\beta}_1) \sin(\tilde{\beta}_1 - \hat{\beta}_1)$ which represents an *interaction* term between the two components of the mixture. The analysis of the term $(\hat{\mathbf{X}}_2 - \mathbf{X}_2)^* \nabla_{\hat{\mathbf{X}}_2} L_t$ leads to an identical conclusion. In fact, this is due to the combinatorial nature of the data association problem in the multiple source case. To overcome this problem, various approaches have been investigated, either based on mixture modeling and identification [30, 31] or, gating and pruning [32]. But, the combinatorial problem is always subjacent. This problem has been thoroughly considered by Pattipati, et al. [33] using Lagrangian relaxation.

IX. SIMULATION RESULTS

The previous results are now illustrated by simulation results. The following scenario is considered. The source trajectory is a leg by leg one

TABLE I
Parameters of Observer Trajectory (see Fig. 2)

Simulation Parameters														
Observer		rx0	ry0	vx0	vy0	t1	vx1	vy1	t2	vx2	vy2	t3	vx3	vy3
		0	0	2	2	50	2	-2	134	2	0	150	2	-0,5
Bearings	snapshots	SNR	number of sensors				number of Trials							
200	10	1	64				100							

TABLE II
Comparisons Between Local and Global Estimation Methods, SNR = 0 dB

Parameters	Exact Value	Estimated Mean	Estimated Standard Deviation	Cramer-Rao Bound	method
rx0	10 000,0	10 022,4	1,4E+02	1,4E+02	M
ry0	10 000,0	10 022,4	1,3E+02	1,3E+02	I
vx0	2,0	2,0	8,6E-02	8,5E-02	V
vy0	3,0	3,0	7,4E-02	7,3E-02	
t1	120,0	120,0	2,5E-01	2,9E-01	G
vx1	-2,0	-2,0	6,4E-01	8,6E-01	L
vy1	2,0	1,9	1,1E+00	1,4E+00	O
t2	145,0	145,0	1,7E-01	5,1E-01	B
vx2	2,0	2,0	2,7E-01	3,1E-01	A
vy2	3,0	3,0	5,1E-01	5,9E-01	L
rx0	10 000,0	10 023,4	1,4E+02	1,4E+02	
ry0	10 000,0	10 023,4	1,4E+02	1,4E+02	
vx0	2,0	2,0	8,8E-02	8,8E-02	L
vy0	3,0	3,0	7,5E-02	7,5E-02	
O					
rx1	12 400,0	124 659,6	1,6E+04	3,1E+03	
ry1	12 400,0	13 691,6	2,2E+04	4,4E+03	C
vx1	-2,0	-2,2	6,3E+00	1,3E+00	
vy1	2,0	1,7	2,6E+01	5,0E+00	A
rx2	11 900,0	-9 214,7	1,2E+05	6,7E+03	L
ry2	14 100,0	-20 525,4	2,0E+05	1,1E+04	
vx2	2,0	2,9	8,4E+00	3,9E+00	
vy2	3,0	-3,4	3,4E+01	2,3E+01	

as well as the observer trajectory. This hypothesis is quite reasonable in the Sonar context.

The observer trajectory is a 4 legs one whose parameters are given in Table I. The source trajectory parameters are given in Tables II and III. Notice that the scenario is identical for Table II and III which differ only by the source signal-to-noise ratio. The source change points are also included in the Tables II and III. The scenario is depicted in Fig. 2, a circle indicating the source and observer starts. We assume that each source bearings is estimated by using 10 array snapshots, each snapshot corresponding to 1 s. Thus, each bearing is estimated on a total duration of 10 s. Since the total scenario involve 200 estimated bearings, its total duration is 2000 s. Furthermore successive estimated bearings are assumed uncorrelated.

The source bearings themselves are simulated by adding a white noise to the exact bearings whose variance is given by the Woodward's formula. In

order to obtain empirical distribution of the estimators of the source trajectory we consider 100 runs of the source trajectory estimation for each scenario. The estimation itself is achieved by means of the MIV algorithm (see Section VI). This choice avoids the problems induced by the computation of a convenient stepsize in the classical gradient algorithm (MLE) which become rather boring when the source state dimension increases. Note however that the advantage is only computational since the results of MLE and MIV are identical.

We compare two estimation strategies. The first one is global. The source trajectory is parameterized by a 8-dimensional state vector whose components are $(r_{x,0}, r_{y,0}, v_{x,0}, \dots, v_{y,2})$. The parameters $(r_{x,0}, r_{y,0})$ are the source coordinates at the time origin, while $(v_{x,i}, v_{y,i})$ are the source velocity parameters on the source leg i ($0 \leq i \leq 2$). The source change points are *estimated* by using the GLR around change points estimated by a sequential detection scheme (see

TABLE III
Comparisons Between Local and Global Estimation Methods, SNR = -10 dB

Parameters	Exacte Value	Estimated Mean	Estimated Standard Deviation	Cramer-Rao Bound	method
rx0	10 000,0	10 040,0	4,9E+02	4,9E+02	M
ry0	10 000,0	10 040,8	4,9E+02	4,5E+02	I
vx0	2,0	2,0	3,1E-01	2,9E-01	V
vy0	3,0	3,0	2,7E-01	2,5E-01	
t1	120,0	119,8	1,1E+00	9,8E-01	G
vx1	-2,0	-2,3	3,2E+00	2,9E+00	L
vy1	2,0	1,5	5,4E+00	4,8E+00	O
t2	145,0	145,0	1,0E+00	1,7E+00	B
vx2	2,0	2,1	9,8E-01	1,1E+00	A
vy2	3,0	3,1	2,0E+00	2,0E+00	L
O					
rx0	10 000,0	10 028,1	4,9E+02	4,7E+02	
ry0	10 000,0	10 028,9	4,8E+02	4,6E+02	
vx0	2,0	2,0	3,1E-01	3,0E-01	L
vy0	3,0	3,0	2,6E-01	2,5E-01	
O					
rx1	12 400,0	12 157,8	2,6E+04	1,1E+04	
ry1	12 400,0	-2 049,7	3,6E+04	1,5E+04	C
vx1	-2,0	-6,6	9,3E+00	4,5E+00	
vy1	2,0	-16,6	3,6E+01	1,7E+01	A
L					
rx2	11 900,0	-2 212,4	3,8E+04	2,3E+04	
ry2	14 100,0	-9 013,9	6,2E+04	3,7E+04	
vx2	2,0	2,2	1,4E+00	1,3E+00	
vy2	3,0	-2,1	1,7E+01	7,6E+00	

Simulation Parameters														
Observer		rx0	ry0	vx0	vy0	t1	vx1	vy1	t2	vx2	vy2	t3	vx3	vy3
		0	0	2	2	50	2	-2	134	2	0	150	2	-0,5
Bearings	snapshots	SNR	number of sensors				number of Trials							
200	10	0,1	64				100							

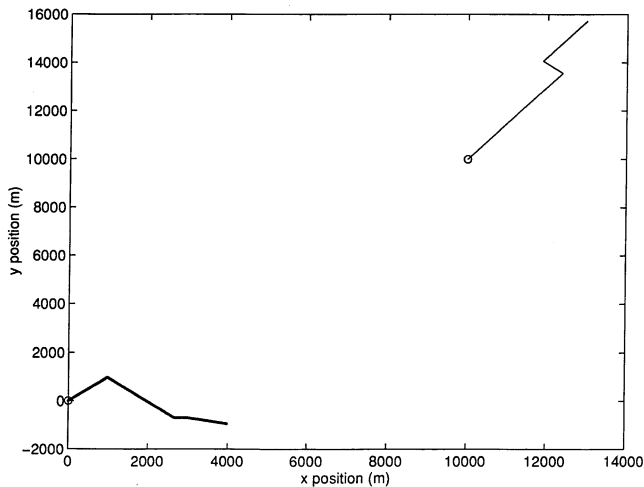


Fig. 2. Target and observer trajectories. Observer trajectory (in bold) originates at (0,0). Target trajectory originates at (10000,10000). Exact parameter values of scenarios in Tables II and III.

Section VII). The global optimization is initialized by local estimates, i.e., obtained on a part of one source leg. Even if these local estimates may be very poor,

no convergence problem occurs for global MIV as far as the assumed change points are sufficiently close to the exact ones (generally 10 bearing samples). The second strategy is local, i.e., source parameters are estimated on each source leg. The source change points are assumed to be known. Furthermore, the TMA problem is observable for each source leg which means, practically, that an observer maneuver occurs for each source leg.

Results of these two estimation strategies are compared, between them, on the first hand, and with the Cramér-Rao bound on the other (100 trials). Notice that the calculation of the Fisher information matrix (FIM) components relative to the source change-points corresponds to a formal liberty. In view of Tables II and III, the results of the global MIV estimation are very close to the bounds, even if a slight bias appears for the estimation of $r_{x,1}, r_{y,1}$ (especially at -10 dB). The results of the local estimation are also quite acceptable on the first leg and very similar to the global optimization results. However, these good results are due to the strong observer maneuver which occurs during the first

source leg. On the contrary, the performance of local estimation is rather poor on source legs 2 and 3. In particular, bias and variance of the local estimator are very important on leg 3. This time, the poor results of the local approach are due to the fact that the amplitude of the observer maneuver is weak on the second leg and very weak on the third one.

On the contrary, the global estimation approach takes advantage of the observer maneuver occurring on the first source leg for estimating the kinematic parameters of the second and third ones. The performance of the TMA algorithm is thus greatly improved. This is the common situation for this type of scenario. Improvement gained by a global estimation may be quite important.

X. CONCLUSION

A general formalism for the analysis of the convergence of iterative methods for BOT has been developed. General results have been obtained. Extensions of previous results to the case of maneuvering source and observer constitutes the major contribution of this work and demonstrate the effectiveness of our approach. Further theoretical interest, simple and feasible (sufficient) convergence criteria have thus been derived which should allow to handle complex TMA scenarios.

APPENDIX A

This Appendix is devoted to the proof of the inequality (37) which is instrumental for obtaining sufficient convergence conditions. This part is inspired by [16]. We begin with the following basic lemma.

LEMMA 5 *Assume that f satisfies the Lipschitz condition (35) then the following inequality holds:*

$$f(\mathbf{X}) - f(\mathbf{Y}) \geq \mathbf{G}^*(\mathbf{X})(\mathbf{X} - \mathbf{Y}) - \frac{L}{2} \|\mathbf{X} - \mathbf{Y}\|^2. \quad (76)$$

PROOF Using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned} f(\mathbf{X}) - f(\mathbf{Y}) &= \int_0^1 \mathbf{G}^*[\mathbf{Y} + \tau(\mathbf{X} - \mathbf{Y})](\mathbf{X} - \mathbf{Y}) d\tau \\ &\geq \mathbf{G}^*(\mathbf{X})(\mathbf{X} - \mathbf{Y}) - \int_0^1 \|\mathbf{G}[\mathbf{Y} + \tau(\mathbf{X} - \mathbf{Y})] \\ &\quad - \mathbf{G}(\mathbf{X})\| \|\mathbf{X} - \mathbf{Y}\| d\tau \\ &\geq \mathbf{G}^*(\mathbf{X})(\mathbf{X} - \mathbf{Y}) - L \|\mathbf{X} - \mathbf{Y}\|^2 \int_0^1 (1 - \tau) d\tau \\ &= \mathbf{G}^*(\mathbf{X})(\mathbf{X} - \mathbf{Y}) - \frac{L}{2} \|\mathbf{X} - \mathbf{Y}\|^2. \quad (77) \end{aligned}$$

Let us now recall the general form of a descent algorithm:

$$\left\{ \begin{array}{l} \mathbf{X}_{l+1} = \mathbf{X}_l - \rho_l \mathbf{S}_l \\ \text{where the stepsize } \rho \text{ satisfies:} \\ f(\mathbf{X}_l - \rho_l \mathbf{S}_l) \leq (1 - \lambda_l) f(\mathbf{X}_l) + \lambda_l \omega_l \\ \text{with: } 0 < \lambda_l \leq 1 \quad \text{and} \quad \omega_l \triangleq \min_{\rho} f(\mathbf{X}_l - \rho \mathbf{S}_l). \end{array} \right. \quad (78)$$

The following equality then holds for any value of the stepsize ρ :

$$\begin{aligned} f(\mathbf{X}_l) - f(\mathbf{X}_{l+1}) &\geq \lambda_l [f(\mathbf{X}_l) - \omega_l] \\ &\geq \lambda_l [f(\mathbf{X}_l) - f(\mathbf{X}_l - \rho \mathbf{S}_l)]. \quad (79) \end{aligned}$$

Using the above lemma and the definition of α_l (cosine of the angle between $\mathbf{G}(\mathbf{X}_l)$ and \mathbf{S}_l), we deduce

$$\begin{aligned} f(\mathbf{X}_l) - f(\mathbf{X}_{l+1}) &\geq \lambda_l [\rho \mathbf{G}^*(\mathbf{X}_l) \mathbf{S}_l - \frac{1}{2} L \rho^2 \|\mathbf{S}_l\|^2] \\ &= \lambda_l [\rho \alpha_l \|\mathbf{G}(\mathbf{X}_l)\| \|\mathbf{S}_l\| - \frac{1}{2} L \rho^2 \|\mathbf{S}_l\|^2] \end{aligned} \quad (80)$$

valid for every value of ρ . Choosing $\rho = (\alpha_l / L) (\|\mathbf{G}(\mathbf{X}_l)\| / \|\mathbf{S}_l\|)$, the following inequality is obtained

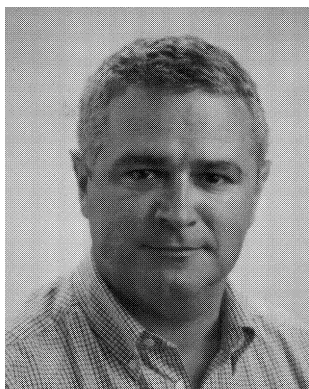
$$f(\mathbf{X}_l) - f(\mathbf{X}_{l+1}) \geq \frac{1}{2L} \lambda_l \alpha_l^2 \|\mathbf{G}(\mathbf{X}_l)\|^2 \quad (81)$$

and (37) is thus proved.

REFERENCES

- [1] Nardone, S., Lindgren, A., and Gong, K. (1984) Fundamental properties and performance of conventional bearings-only target motion analysis. *IEEE Transactions on Automatic Control*, **AC-29**, 9 (Sept. 1984), 775–787.
- [2] Kolb, R. C., and Hollister, F. H. (1967) Bearings-only target motion estimation. In *Proceedings of the 1st Asilomar Conference Circuits and Systems*, Monterey, CA, 1967.
- [3] Chan, Y. T., and Rudnicki, S. W. (1992) Bearings-only and Doppler-bearing tracking using instrumental variables. *IEEE Transactions on Aerospace and Electronic Systems*, **28**, 4 (Oct. 1992), 1076–1083.
- [4] Aidala, V. J., and Hammel, S. E. (1983) Utilization of modified polar coordinates for bearings-only tracking. *IEEE Transactions on Automatic Control*, **AC-28** (Mar. 1983), 283–294.
- [5] Shensa, M. J. (1981) The MLE tracking algorithm: A summary and error analysis. Technical report 689, Naval Ocean Systems Center, May 1981, 26.

- [6] Passerieux, J. M., Pillon, D., Blanc-Benon, P., and Jauffret, C. (1989)
Target motion analysis with bearings and frequencies measurements via instrumental variable estimator.
In *Proceedings of the IEEE International Conference on Acoustics and Signal Processing*, Glasgow, Scotland, May 1989, 2645–2648.
- [7] Iltis, R. A., and Anderson, K. L. (1996)
A consistent estimation criterion for multisensor bearings-only tracking.
IEEE Transactions on Aerospace and Electronic Systems, **32**, 1 (Jan. 1996), 108–121.
- [8] Barnett, S., and Cameron, R. G. (1985)
Introduction to Mathematical Control Theory (2nd ed.), (Applied Mathematics Series).
New York: Oxford University Press, 1985, ch. 5.
- [9] Horn, R. A., and Johnson, C. R. (1987)
Matrix Analysis.
New York: Cambridge University Press, 1987.
- [10] Trémois, O., and Le Cadre, J. P. (1996)
Target motion analysis with multiple arrays: Performance analysis.
IEEE Transactions on Aerospace and Electronic Systems (July 1996), 1030–1046.
- [11] Dana, M. P. (1990)
Registration: A prerequisite for multiple sensor tracking.
In Y. Bar-Shalom (Ed.), *Multitarget-Multisensor Tracking: Advanced Applications*.
Dedham, MA: Artech House, 1990.
- [12] Mc Michael, D. W., and Okello, N. N. (1996)
Maximum likelihood registration of dissimilar sensors.
In *Proceedings of the 1st Australian Data Fusion Symposium*, ADFS-96, Adelaide, Nov. 1996, 31–34.
- [13] Bazaraa, M. S., Sherali, H. D., and Shetty, C. M. (1993)
Non Linear Programming Theory and Algorithms (2nd ed.).
New York: Wiley, 1993.
- [14] Nardone, S. C., and Aidala, V. J. (1981)
Observability criteria for bearings-only target motion analysis.
IEEE Transactions on Aerospace and Electronic Systems, **17**, 2 (Mar. 1981), 162–166.
- [15] Le Cadre, J. P., and Jauffret, C. (1997)
Discrete-time observability and estimability for bearings-only target motion analysis.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 21 (Jan. 1997), 178–201.
- [16] Karmanov, V. (1977)
Programmation Mathématique.
Editions de Moscou, 1977 (in French).
- [17] Cartan, H. (1977)
Calcul Différentiel.
Paris: Hermann, 1977.
- [18] Gong, K. F., Lindgren, A. G., and Nardone, S. C. (1981)
Passive localization and motion analysis with a state parameter constraint.
In *Proceedings of the 15th Asilomar Conference on Circuits, Systems and Computers*, Nov. 9–11, 1981, 366–374.
- [19] Ciarlet, P. G. (1982)
Introduction à L'analyse Numérique Matricielle et à L'optimisation.
Paris: Masson, 1982, 226–230.
- [20] Poor, H. V. (1994)
An Introduction to Signal Detection and Estimation (2nd ed.). (Springer Texts in Electrical Engineering.)
New York: Springer-Verlag, 1994.
- [21] Lindgren, A. G., and Gong, K. F. (1978)
Properties of bearings-only motion analysis estimator: An interesting case study in system observability.
In *Proceedings of the 12th Asilomar Conference on Circuits Systems*, Monterey, CA, Nov. 1978.
- [22] Wong, K. Y., and Polak, E. (1967)
Identification of linear discrete time systems using the instrumental variable.
IEEE Transactions on Automatic Control, **AC-12** (Dec. 1967), 707–718.
- [23] Browden, R. J., and Turkington, D. A. (1984)
Instrumental Variables. (Econometric Series in Quantitative Economics).
New York: Cambridge University Press, 1984.
- [24] Karr, A. F. (1993)
Probability. (Springer Texts in Statistics).
New York: Springer-Verlag, 1993, 146–147.
- [25] Ferguson, T. S. (1996)
A Course in Large Sample Theory. (Texts in Statistical Science).
London: Chapman and Hall, 1996.
- [26] Mc Aulay, R. J., and Denlinger, E. J. (1973)
A decision-directed adaptive tracker.
IEEE Transactions on Aerospace and Electronic Systems, **9**, 2 (Mar. 1973), 229–236.
- [27] Bogler, P. (1987)
Tracking a maneuvering target using input estimation.
IEEE Transactions on Aerospace and Electronic Systems, **23**, 3 (May 1987), 298–310.
- [28] Passerieux, J. M. (1987)
Comparaison des performances des méthodes non récursives en trajectographie passive sous-marine.
In *Actes du 11-ème colloque GRETSI*, Nice, France, June 1987, 349–352.
- [29] Mood, A. M., Graybill, F. A., and Boes, D. C. (1974)
Introduction to the Theory of Statistics (3rd ed.). (Series in Probability and Statistics.)
New York: McGraw-Hill, 1974.
- [30] Streit, R. L., and Luginbuhl, T. E. (1994)
Maximum likelihood method for probabilistic multi-hypothesis tracking.
In *Proceedings of SPIE International Symposium on Signal and Data Processing of Small Targets 1994*, SPIE Proceedings, **2335-24**, Orlando, FL, Apr. 5–7, 1994.
- [31] Gauvrit, H., Jauffret, C., and Le Cadre, J. P. (1997)
A formulation of multitarget tracking as an incomplete data problem.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 4 (Oct. 1997), 1242–1257.
- [32] Bar Shalom, Y., and Fortmann, T. E. (1988)
Tracking and Data Association.
New York: Academic Press, 1988.
- [33] Pattipati, K. R., Deb, S., Bar-Shalom, Y., and Washburn, R. B. (1992)
A new relaxation algorithm and passive sensor data association.
IEEE Transactions on Automatic Control, **27**, 2 (Feb. 1992), 198–213.



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 multitarget tracking 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.



Claude Jauffret (M'95) was born on March 29, 1957, in Ollioules, France. He received the Diplôme d'Études Approfondies in applied mathematics from the Saint Charles University, Marseille, France in 1981, the diplôme d'Ingénieur from the École Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble, Grenoble, France in 1983, the title of Doctor of the University of Toulon, Toulon, France, in 1993 and the Habilitation à Diriger des Recherches.

From Nov. 1983 to Sept. 1984, he worked on image processing used in passive sonar systems at the Groupe d'Étude et de Recherche en Détection Sous-Marine (GERDSM), Six-Fours-les-Plages, France. From 1984 to 1988, he worked on target motion analysis (TMA) problems at the GERDSM. A sabbatical year at the University of Connecticut, Storrs, allowed him to work on normalization, estimation and data fusion encountered in sonar systems. In September 1996, he joined the University of Toulon-Var (France), where he is professor.

Dr. Jauffret has published papers about observability and estimation in nonlinear systems and TMA.