What to Do with Biased Measurements?

Annie-Claude Pérez,
Université de Toulon, Aix Marseille
Univ, CNRS, IM2NP, Toulon, France
CS 60584, 83041 TOULON Cedex 9,
France
annie-claude.perez@univ-tln.fr

Claude Jauffret,
Université de Toulon, Aix Marseille
Univ, CNRS, IM2NP, Toulon, France
CS 60584, 83041 TOULON Cedex 9,
France
jauffret@univ-tln.fr

***Abstract*— Estimating a parameter with biased measurements is the topic of this paper. To do this, we propose three different models when the available data are corrupted by an additive constant bias. For each model, we detail the computation of the Cramér-Rao lower bound (CRLB). For each model, examples using the traditional bearing-only target motion analysis (BOTMA) are given, for which we propose an estimator whose empirical performance is compared to the CRLB.**

***Keywords*— Bias, Fisher information matrix, Cramér-Rao lower bound, estimation, bearings, TMA.**

I. INTRODUCTION

We presented a paper at the international FUSION conference in 2023 dealing with the observability of the bias in the problem of bearings-only target motion analysis (BOTMA) [1]. Our paper followed another one [2], presented at FUSION'20, in which the observability analysis was skipped. In a more recent article submitted to the *Journal of Advances in Information Fusion*, we addressed the estimation of the bias (together with the state vector) when observability is ensured [3]. But we did not take into account a prior about the bias, i.e., we considered the magnitude of the bias as an unknown deterministic extra-parameter. Doing this, we constructed what we will call, in this paper, Model I. This model was adopted in [4] when time difference of arrival (TDOA) measurements are biased in a problem of direction finding and in [5] to estimate the localization of stations' positions with biased range measurements in a station network.

Another approach consists in considering the bias as random. In this approach, two models can be proposed: In Model II, the bias is interpreted as a realization of a random variable, and it must be estimated – this is a Bayesian point of view. In Model III, the bias is assimilated as an "extra" noise: in this model, the measurements are correlated because of the presence of this common extra noise. The Bayesian approach was initially proposed by Gavish and Fogel in [6]. They addressed the problem of the "effect of bias on bearing-only target location." Model III was used in [7] to compute the Cramér-Rao lower bound (CRLB) when the measurements are the distances between a target and a set of beacons. Some beacons collect biased measurements, while the range measurements given by the other beacons are unbiased.

As far as we know, Model III has rarely been used in the TMA literature on this topic. This is why we propose to visit these three models of biased measurements and to implement these three models for (biased) bearings target motion analysis. We construct an estimate for each model and compare their performances with the CRLB.

We do insist on the fact that the topic of the paper is to estimate the bias of the measurements, not the bias of the estimators. This is why, the list of references is not very long.

This paper is organized as follows:

In the next section, the classic model of unbiased measurements is recalled, including the derivation of the Fisher information matrix (FIM). Section III is devoted to biased measurements, where we will propose three models. For each of them, we give the log-likelihood function, the FIM, and the CRLB. Examples coming from the bearings-only target motion analysis problems are given in Section IV. These examples confirm the previous computations, and intensive Monte Carlo simulations allow us to appreciate the performance of the proposed estimators.

II. REVIEW OF THE CASE OF UNBIASED MEASUREMENTS

Suppose that a sensor has made N unbiased measurements obeying the measurement equation

$$y_k = s_k(X) + \varepsilon_k,\ k = 1, \cdots, N \tag{1}$$

in which X is a d-dimensional vector, called the state vector or parameter, and $\varepsilon = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_N \end{bmatrix}^{\mathrm{T}}$ is the noise vector. We assume that it is a zero-mean Gaussian vector whose covariance matrix is $\mathbf{R}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_N$. And $s_k(X)$ is a mapping between $\mathbb{R}^d$ and $\mathbb{R}$. We define also the pure measurement vector

$$s(X) = \begin{bmatrix} s_1(X) & s_2(X) & \cdots & s_N(X) \end{bmatrix}^{\mathrm{T}}. \tag{2}$$

We assume that X is observable.

The probability density function (pdf) of $y = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}^{\mathrm{T}}$ is $K_U \exp\left[-C_U(X)\right]$, with

$$C_U(X) = \frac{1}{2}\left[s(X) - y\right]^{\mathrm{T}} \mathbf{R}_\varepsilon^{-1} \left[s(X) - y\right] \tag{3}$$

and

$$K_U = \frac{1}{\left(\sqrt{2\pi}\sigma_\varepsilon\right)^N}.$$

Under this assumption, the FIM of X, when y is available, is $\mathbf{F}_U(X) = \nabla_X s(X) \mathbf{R}_\varepsilon^{-1} \nabla_X^{\mathrm{T}} s(X)$. The CRLB of any unbiased estimator of X is given by $\mathbf{B}_U(X) = \mathbf{F}_U^{-1}(X)$. Note that the subscript U is for *unbiased*.

To make the following computation easier, we introduce the following notation:

$g_k = \nabla_X s_k(X)$, i.e., the gradient of $s_k(X)$

$\mathbf{h}_k = \nabla_X \nabla_X^{\mathrm{T}} s_k(X)$, i.e., the Hessian of $s_k(X)$ (see [8] p. 280)

$$\nabla_X s(X) = [g_1 \quad g_2 \quad \cdots \quad g_N]. \tag{4}$$

So the FIM is given by $\mathbf{F}_U(X) = \sigma_\varepsilon^{-2} \sum_{k=1}^{N} g_k g_k^{\mathrm{T}}$.

We define $\mathbf{f} = \sum_{k=1}^{N} g_k g_k^{\mathrm{T}}$. So $\mathbf{F}_U(X) = \sigma_\varepsilon^{-2}\mathbf{f}$, and $\mathbf{B}_U(X) = \sigma_\varepsilon^{2}\mathbf{f}^{-1}$

$\mathbf{1}_L = [1 \quad \cdots \quad 1]^{\mathrm{T}}$ (L-dimensional vector of ones)

$\mathbf{0}_{L,1} = [0 \quad \cdots \quad 0]^{\mathrm{T}}$ (L-dimensional vector of zeros)

$\mathbf{0}_{p,d}$ is the $p \times d$ matrix of zeros.

III. WHEN THE MEASUREMENTS ARE BIASED

In this section, we assume that the data are polluted by an additive bias, say b. The measurement equations are now

$$z_k = y_k + b = s_k(X) + b + \varepsilon_k, \quad k = 1, \cdots, N. \tag{5}$$

We define also the measurement vector $z = [z_1 \quad z_2 \quad \cdots \quad z_N]^{\mathrm{T}}$.

A. First model: The bias is viewed as a unknown deterministic parameter (with no prior)

In this model, the bias is assumed to be constant. The state vector must hence be augmented by the unknown magnitude of the bias: $\begin{bmatrix} X^{\mathrm{T}} & b \end{bmatrix}^{\mathrm{T}}$. We assume that b is observable.

We introduce the vector

$r(X,b) = \begin{bmatrix} s_1(X)+b & s_2(X)+b & \cdots & s_N(X)+b \end{bmatrix}^{\mathrm{T}}$. The pdf of z is then

$$p_{b,1}(z|X,b) = K_{b,1} \exp\left[-C_{b,1}(X,b)\right],$$

with

$$C_{b,1}(X,b) = \frac{1}{2}\left[r(X,b) - z\right]^{\mathrm{T}} \mathbf{R}_\varepsilon^{-1} \left[r(X,b) - z\right]$$

$$= \frac{1}{2\sigma_\varepsilon^2} \sum_{k=1}^{N} \left[s_k(X) + b - z_k\right]^2, \tag{6}$$

and

$K_{b,1} = K_U$.

The gradient of $r(X,b)$ is

$$\nabla_{X,b} r(X,b) = \begin{bmatrix} \nabla_X s_1(X) & \nabla_X s_2(X) & \cdots & \nabla_X s_N(X) \\ \dfrac{\partial s_1(X)+b}{\partial b} & \dfrac{\partial s_2(X)+b}{\partial b} & \cdots & \dfrac{\partial s_N(X)+b}{\partial b} \end{bmatrix} \tag{7}$$

So, $$\nabla_{X,b} r(X,b) = \begin{bmatrix} g_1 & g_2 & \cdots & g_N \\ 1 & 1 & \cdots & 1 \end{bmatrix}. \tag{8}$$

Then, the FIM is

$$\mathbf{F}_{b,1}(X,b) = \nabla_{X,b} r(X,b) \mathbf{R}_\varepsilon^{-1} \nabla_{X,b}^{\mathrm{T}} r(X,b)$$

$$= \sigma_\varepsilon^{-2} \begin{bmatrix} \sum_{k=1}^{N} g_k g_k^{\mathrm{T}} & \sum_{k=1}^{N} g_k \\ \sum_{k=1}^{N} g_k^{\mathrm{T}} & N \end{bmatrix} = \sigma_\varepsilon^{-2} \begin{bmatrix} \mathbf{f} & \mathbf{g} \\ \mathbf{g}^{\mathrm{T}} & N \end{bmatrix} \tag{9}$$

with $\mathbf{g} = \sum_{k=1}^{N} g_k$.

We recognize in $\sigma_\varepsilon^{-2}\mathbf{f} = \mathbf{F}_U(X)$ the FIM when the measurement are unbiased.

The classic blockwise inversion formula [9] yields the CRLB

$$B_{b,1}(X,b) = \sigma_\varepsilon^2 \begin{bmatrix} \mathbf{f} & \mathbf{g} \\ \mathbf{g}^{\mathrm{T}} & N \end{bmatrix}^{-1} = \sigma_\varepsilon^2 \begin{bmatrix} \mathbf{f}^{-1} + \eta \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} & -\eta \mathbf{f}^{-1}\mathbf{g} \\ -\eta \mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} & \eta \end{bmatrix}, \tag{10}$$

with $$\eta = \frac{1}{N - \mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1}\mathbf{g}}. \tag{11}$$

We deduce that the CRLB of X is

$$B_{b,1}(X) = \sigma_\varepsilon^2 \left(\mathbf{f}^{-1} + \eta \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1}\right) = B_U(X) + \sigma_\varepsilon^2 \eta \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} \tag{12}$$

B. Second model: The bias is viewed as a unknown parameter with a prior (Bayesian model)

The measurement equations and the augmented state vector are the same as in the first approach. Now, we assume that b is the realization of a 0-mean Gaussian random variable, independent conditionally on X of the measurements, whose variance is σ_b^2: the law $G(0,\sigma_b^2)$ describes the prior knowledge of b. The pdf of the couple (z,b) is then

$p_{b,2}(z,b|X) = K_{b,2} \exp\left[-C_{b,2}(X,b)\right]$, with

$$C_{b,2}(X,b) = C_{b,1}(X,b) + \frac{b^2}{2\sigma_b^2}$$

$$= \frac{1}{2\sigma_\varepsilon^2} \sum_{k=1}^{N} \left[s_k(X) + b - z_k\right]^2 + \frac{b^2}{2\sigma_b^2}, \tag{13}$$

and

$$K_{b,2} = \frac{1}{\left(\sqrt{2\pi}\sigma_\varepsilon\right)^N \sqrt{2\pi}\sigma_b}$$

Hence,

$$p_{b,2}(z,b|X) = K_{b,2} \exp\left[-\frac{1}{2} C_{b,1}(X,b)\right] \exp\left(-\frac{b^2}{2\sigma_b^2}\right)$$

Following [10] p. 84, we compute the *a priori* information about b as $\begin{bmatrix} \mathbf{0}_{d,d} & \mathbf{0}_{d,1} \\ \mathbf{0}_{1,d} & \sigma_b^{-2} \end{bmatrix}$, which must be added to $\mathbf{F}_{b,1}(X,b)$:

$$\mathbf{F}_{b,2}(X,b) = \sigma_\varepsilon^{-2} \begin{bmatrix} \mathbf{f} & \mathbf{g} \\ \mathbf{g}^{\mathrm{T}} & N + \left(\sigma_\varepsilon / \sigma_b\right)^2 \end{bmatrix}. \tag{14}$$

Its inverse is

$$\mathbf{B}_{b,2}(X,b) = \sigma_\varepsilon^2 \begin{bmatrix} \mathbf{f}^{-1} + \kappa \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} & -\kappa \mathbf{f}^{-1}\mathbf{g} \\ -\kappa \mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} & \kappa \end{bmatrix}, \tag{15}$$

with $$\kappa = \frac{1}{N + \left(\sigma_\varepsilon / \sigma_b\right)^2 - \mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1}\mathbf{g}}. \tag{16}$$

We conclude from this that

$$\mathbf{B}_{b,2}(X) = \sigma_\varepsilon^2 \mathbf{f}^{-1} + \sigma_\varepsilon^2 \kappa \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1}$$

$$= \mathbf{B}_U(X) + \sigma_\varepsilon^2 \kappa \mathbf{f}^{-1}\mathbf{g}\mathbf{g}^{\mathrm{T}}\mathbf{f}^{-1} \tag{17}$$

<u>Remark:</u>

- Since $\kappa < \eta$, $\mathbf{B}_{b,1}(X) - \mathbf{B}_{b,2}(X)$ is a non-negative symmetrical matrix, i.e., $\mathbf{B}_{b,2}(X) \leq \mathbf{B}_{b,1}(X)$.

- When $\sigma_b \to 0$, then $\kappa \to 0$, and $\mathbf{B}_{b,2}(X) \to \mathbf{B}_U(X)$.
- When $\sigma_b \to \infty$, then $\kappa \to \eta$, and $\mathbf{B}_{b,2}(X) \to \mathbf{B}_{b,1}(X)$.

C. Third model: The bias is viewed as an extra noise

In this model, b is viewed as an additive noise. Hence, it must not be estimated, and, unlike the Bayesian model, the state vector is not augmented. Only the covariance matrix of the measurements changes:

Since $\mathrm{Var}(z_k) = \sigma_\varepsilon^2 + \sigma_b^2$, and $\mathrm{Cov}(z_i, z_j) = \sigma_b^2$, we get

$$\mathbf{R}_{\varepsilon,b} = \begin{bmatrix} \sigma_\varepsilon^2+\sigma_b^2 & \sigma_b^2 & \cdots & \sigma_b^2 \\ \sigma_b^2 & \sigma_\varepsilon^2+\sigma_b^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma_b^2 \\ \sigma_b^2 & \cdots & \sigma_b^2 & \sigma_\varepsilon^2+\sigma_b^2 \end{bmatrix}$$

$$= \sigma_\varepsilon^2 \mathbf{I}_N + \sigma_b^2 \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} = \sigma_\varepsilon^2 \left[\mathbf{I}_N + \left(\sigma_b / \sigma_\varepsilon \right)^2 \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right] \quad (18)$$

The pdf of z is, this time,

$p_{b,3}(z|X) = K_{b,3} \exp\left[-C_{b,3}(X) \right]$,

with

$$C_{b,3}(X) = \frac{1}{2} \left[s(X) - z \right]^{\mathrm{T}} \mathbf{R}_{\varepsilon,b}^{-1} \left[s(X) - z \right], \quad (19)$$

and

$$K_{b,3} = \frac{1}{\sqrt{\det(2\pi R_{\varepsilon,b})}}$$

Note that in this model, the measurements are in some way unbiased!

Putting this together with the Sherman-Morrison formula [11], we obtain a closed form for the inverse of $\mathbf{R}_{\varepsilon,b}$:

$$\mathbf{R}_{\varepsilon,b}^{-1} = \sigma_\varepsilon^{-2} \left(\mathbf{I}_N + \sigma_b^2 \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right)^{-1},$$

$$= \sigma_\varepsilon^{-2} \left[\mathbf{I}_N - \frac{\left(\sigma_b / \sigma_\varepsilon \right)^2}{1 + \left\| \sigma_b / \sigma_\varepsilon \mathbf{1}_N \right\|^2} \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right] = \sigma_\varepsilon^{-2} \left[\mathbf{I}_N - \alpha \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right] \quad (20)$$

with $\alpha = \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2}$ (21)

Note that α increases with σ_b, and goes from 0 to $\dfrac{1}{N}$.

We deduce from (20) that

$$C_{b,3}(X) = \frac{1}{2\sigma_\varepsilon^2} \left[s(X) - z \right]^{\mathrm{T}} \left(\mathbf{I}_N - \alpha \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right) \left[s(X) - z \right]$$

$$C_{b,3}(X) = \frac{1}{2\sigma_\varepsilon^2} \sum_{k=1}^{N} \left[s_k(X) - z_k \right]^2 - \frac{\alpha}{2\sigma_\varepsilon^2} \left\{ \sum_{k=1}^{N} \left[s_k(X) - z_k \right] \right\}^2 \quad (22)$$

We deduce that the FIM is

$$\mathbf{F}_{b,3}(X) = \sigma_\varepsilon^{-2} \nabla_X s(X) \nabla_X^{\mathrm{T}} s(X) - \frac{\sigma_\varepsilon^{-2} \sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2} \nabla_X s(X) \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \nabla_X^{\mathrm{T}} s(X)$$

$$= \sigma_\varepsilon^{-2} \left(\mathbf{f} - \frac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \right). \quad (23)$$

Hence, $\mathbf{B}_{b,3}(X) = \sigma_\varepsilon^2 \left(\mathbf{f} - \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \right)^{-1}$

We use again the Sherman-Morrison formula to get the inverse of $\mathbf{F}_{b,3}(X)$:

$$\mathbf{B}_{b,3}(X) = \sigma_\varepsilon^2 \left[\mathbf{f}^{-1} + \frac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2} \frac{1}{1 - \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2} \mathbf{g}^T \mathbf{f}^{-1} \mathbf{g}} \mathbf{f}^{-1} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \mathbf{f}^{-1} \right] \quad (24)$$

$$\mathbf{B}_{b,3}(X) = \sigma_\varepsilon^2 \mathbf{f}^{-1} + \frac{\sigma_\varepsilon^2 \sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2 - \sigma_b^2 \mathbf{g}^T \mathbf{f}^{-1} \mathbf{g}} \mathbf{f}^{-1} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \mathbf{f}^{-1} \quad (25)$$

$$\mathbf{B}_{b,3}(X) = \sigma_\varepsilon^2 \mathbf{f}^{-1} + \frac{\sigma_\varepsilon^2}{\left(\sigma_\varepsilon / \sigma_b \right)^2 + N - \mathbf{g}^T \mathbf{f}^{-1} \mathbf{g}} \mathbf{f}^{-1} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \mathbf{f}^{-1}, \text{ in}$$

which we recognize κ (see (16)):

$$\mathbf{B}_{b,3}(X) = \mathbf{B}_U(X) + \sigma_\varepsilon^2 \kappa \mathbf{f}^{-1} \mathbf{g}\,\mathbf{g}^{\mathrm{T}} \mathbf{f}^{-1} \quad (26)$$

We conclude that, surprisingly, the CRLB of X is the same in the second and third model (see (17)). Still, we have to estimate the bias in the second model, whereas in the third model, the bias has not to be estimated. However, the bias can be estimated, once X has been estimated by $\hat{X}$ via the formula (for example): $\hat{b} = \dfrac{1}{N} \sum_{k=1}^{N} z_k - s_k(\hat{X})$.

Remark:

- In the three models, the magnitude of b does not intervene in the computation of the FIMs. However in the two last models, the variance σ_b^2 matters.
- The links between the three models are the following:

  $p_{b,3}(z|X) = \int p_{b,2}(z,b|X)\,db = \int p_{b,1}(z|X,b)\,p(b)\,db$

  since $p(b|X) = p(b)$.
- The estimation (for each model) can be made by minimizing either $C_{b,1}(X,b)$, $C_{b,2}(X,b)$ or $C_{b,3}(X)$. Note that in Model III, the bias has not to be estimated.

IV. EXAMPLE: THE BEARINGS-ONLY TMA.

We propose some examples coming from the BOTMA literature. The BOTMA is completely described in [12].

The assumption of the classic BOTMA is the following:

Two vehicles are moving in the same plane. One is the observer, and the second is the source (or the target).

During the scenario, the target has a constant velocity, while the observer regularly measures the azimuths in which it detects the source. The observer maneuvers to get observability [1] [3]. To define the positions of each vehicle, the plane has a referential $(O, East, North)$. The angles are clockwise and quantified from North. The notations to be used will now be presented.

The state vector characterizing the trajectory of the target is $X = \left[x_T \quad y_T \quad \dot{x}_T \quad \dot{y}_T \right]^{\mathrm{T}}$. The first two components are those of its position at a chosen time t^*, and the last two are those of its velocity. The positions of the target and of the observer at time t are $P_{T,t}$ and $P_{O,t}$, respectively.

The theoretical bearing at time t is

$$\theta_t(X) = \angle(P_{O,t}, P_{T,t})$$
$$= \arctan 2\left(x_{O,t} - x_T - (t-t^*)\dot{x}_T, y_{O,t} - y_T - (t-t^*)\dot{y}_T\right) \qquad (27)$$

The bearings are collected at time $t_k = (k-1)\Delta t$, and are biased:

$$\beta_k = \theta_k(X) + b + \varepsilon_k, \quad k = 1, \cdots, N. \qquad (28)$$

In this equation, β_k and $\theta_k(X)$ play the role of z_k and $s_k(X)$, respectively (see (5)).

In our simulation, we chose t^* as the final time of the scenario $\Delta t = 1$ s, and $\sigma_\varepsilon = 1°$.

The trajectory of the observer is composed of three legs (to insure observability [3]): the first leg begins at the initial time and finishes at 250 s; the second one ends at 350 s; and the last one finishes at 600 s. The three successive headings are 90°, -60°, and 20°. Its speed is 8 m/s.

The speed of the target is 7 m/s, and its course is 45°. Its initial position is $[4000 \quad 4000]^T$ (m).

For the first model, the bias is $b = 3°$. For the last two models, the standard deviation of the bias is $\sigma_b = 5°$.

We chose the maximum likelihood estimator (MLE) for Model I and Model III, and the maximum a posteriori (MAP) estimator for Model II. They minimize $C_{b,1}(X,b)$, $C_{b,2}(X,b)$ and $C_{b,3}(X)$. Their computations were made via the Gauss-Newton routine (any other numerical routine should work, as well).

In the coming subsections, the results of our Monte Carlo simulations (10,000 runs) are given in tables. Each table contains four columns. The first gives the "true" values of the components of the state vector X and b (when the magnitude of b is fixed). The second contains the empirical means of each component of the estimator (given by the Monte Carlo simulations). We computed the square roots of the diagonal of the CRLB and present then in the third column. The fourth column presents the empirical standard deviation of the estimator. The content of the second column must be compared to the first column (line by line), and the fourth column has to be compared to column 3. We computed also the mean values of the normalized estimation error squared (NEES) [13], restricted to *X*, and compared them to its 2σ-confidence interval (4 ± 0.06).

A. First model: The bias is deterministic

In this subsection, the criterion to minimize is $C_{b,1}(X,b) = \frac{1}{2\sigma_\varepsilon^2}\sum_{k=1}^{N}\left[\theta_k(X) + b - \beta_k\right]^2$ (see eq. (6)).

The results are presented in Fig.1 and Table I.

In Fig. 1, the 500 final location estimates are plotted together with the 90% confidence ellipsoid.

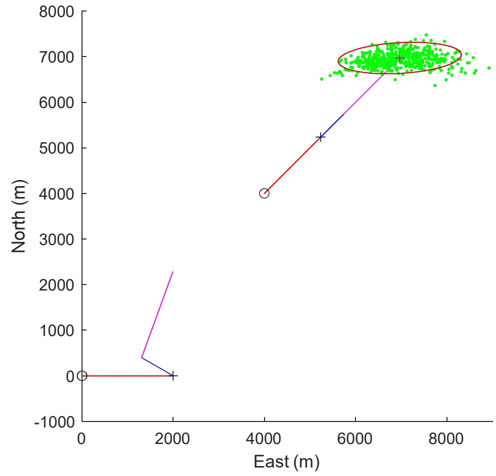

Fig. 1. The trajectories of the observer and of the target, the cloud of position MLEs at final time, together with the 90% confidence ellipse, for Model I.

Table I gives details about the performance.

TABLE I. PERFORMANCE OF THE MLE IN MODEL I

(X,b)	$<\hat{X}>, <\hat{b}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	6967.3	629.3	614.2
6969.8 (m)	6931.8	161.0	168.8
4.95 (m/s)	4.98	0.98	0.96
4.95 (m/s)	4.89	0.84	0.83
3 (degrees)	3.00	3.55	3.49

The NEES being equal 30.03, we cannot conclude that the MLE is Gaussian, even though its empirical mean and standard deviation are close to *X* and the corresponding asymptotical standard deviation given by the CRLB.

However, it is worth noting that the greatest difference between the elements of the CRLB and the empirical covariance matrix of the MLE is less than 9.86%.

B. Second model: The bias is a realization of a random variable

This model was initially proposed in [7] for a problem of localization of a target from a moving passive observer collecting bearings. The criterion is now (see eq. (13))

$$C_{b,2}(X,b) = \frac{1}{2\sigma_\varepsilon^2}\sum_{k=1}^{N}\left[\theta_k(X) + b - \beta_k\right]^2 + \frac{b^2}{2\sigma_b^2}.$$

1) To appreciate the contribution of the prior to the bias, we fixed the value of the bias at 3° (as in the case of Model I) in our Monte Carlo simulations. In this way, we can compare the respective behavior of the MLE and of the MAP, with Table I and Table II.

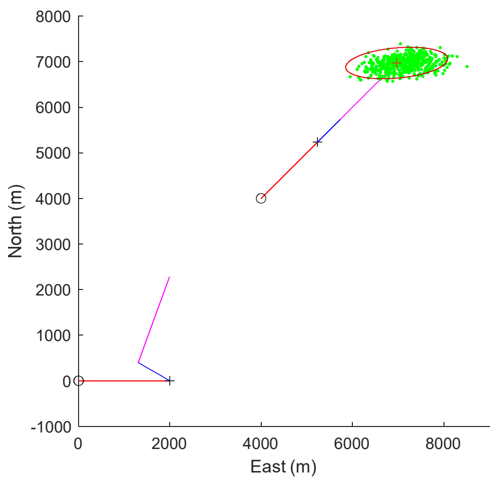

Fig. 2. The trajectories of the observer and of the target, the cloud of position MAPs at final time, together with the 90% confidence ellipse, for Model II, when $b = 3°$.

TABLE II. PERFORMANCE OF THE MAP IN MODEL II, WHEN THE MAGNITUDE OF THE BIAS IS FIXED

(X,b)	$<\hat{X}>,<\hat{b}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	7122	520.8	426.5
6969.8 (m)	6951.1	161.0	152.3
4.95 (m/s)	5.19	0.84	0.72
4.95 (m/s)	5.10	0.73	0.62
3 (degrees)	2.12	2.89	2.3

The presence of a constant bias in the measurements (the same in each simulation run) yields a bias of estimation (see column $<\hat{X}>,<\hat{b}>$ of Table II.

2) To fulfill the mathematical assumption, we now draw a value of b at each simulation (according to the Gaussian law $G(0,\sigma_b^2)$). In this case, we have to compute the so-called mean-square error of $\hat{b}$, which is $\mathrm{MSE}(\hat{b}) = \mathrm{E}\left[\left(\hat{b}-b\right)^2\right]$ (see [10], p. 72), and its square root $\mathrm{RMSE}(\hat{b}) = \sqrt{\mathrm{MSE}(\hat{b})}$ The theoretical computation of this quantity being impossible, we give an approximate value of this using its empirical version. We got $\mathrm{RMSE}_{\mathrm{emp}}(\hat{b}) = 2.86°$. This value must be compared to the square-root of the (5,5)th element of $B_{b,1}(X,b)$, which is $\sqrt{\eta} = 2.89°$ (see Eq. (11)). We denote it by σ_{CRLB} in Table III.

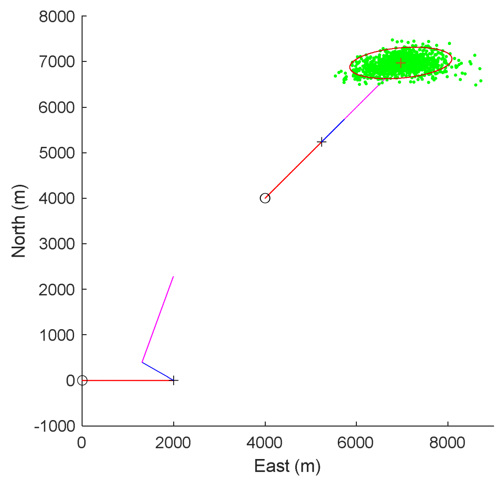

Fig. 3. The trajectories of the observer and of the target, the cloud of position MLEs at final time, together with the 90% confidence ellipse, for Model II, when b is randomized according to $G(0,\sigma_b^2)$.

TABLE III. PERFORMANCE OF THE MAP IN MODEL II, WHEN THE MAGNITUDE OF THE BIAS IS RANDOMIZED

X	$<\hat{X}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	6968.5	520.8	517.6
6969.8 (m)	6945.4	161,0	164.9
4.95 (m)	4.98	0.84	0.83
4.95 (m)	4.91	0.73	0.74
b	$\hat{b}$	σ_{CRLB}	$\mathrm{RMSE}_{\mathrm{emp}}$
Random	-	2.89°	2.86°

We observe now that the estimator are no longer biased. But here again, the magnitude of the NEES (15.83) is too high, and the relative difference between the CRLB and the empirical covariance matrix of the MAP is less than 7%.

C. *Third model: The bias is an extra noise*

As in Section IV.B, the magnitude of the bias is randomly selected at each simulation. The criterion is

$$C_{b,3}(X) = \frac{1}{2\sigma_\varepsilon^2}\sum_{k=1}^{N}\left[\theta_k(X)-\beta_k\right]^2 - \frac{\alpha}{2\sigma_\varepsilon^2}\left\{\sum_{k=1}^{N}\left[\theta_k(X)-\beta_k\right]\right\}^2$$

with $\alpha = \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + N\sigma_b^2}$. This is the version of eq. (22) for TMA.

The results presented in Fig. IV and Table IV have to be compared with those in Fig. 3 and Table III.

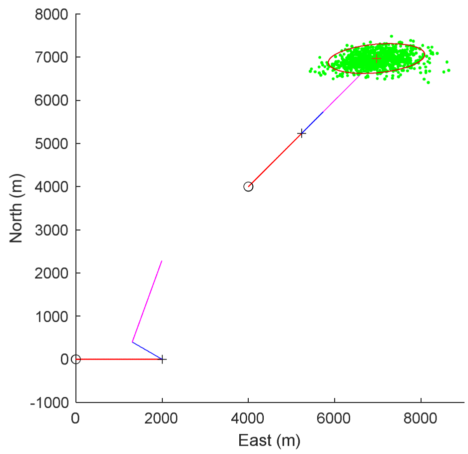

Fig. 4. The trajectories of the observer and of the target, the cloud of position MLEs at final time, together with the 90% confidence ellipse, for Model III, when b is an extra noise.

TABLE IV. PERFORMANCE OF THE MLE IN MODEL III, WHEN THE BIAS IS AN EXTRA NOISE

X	$<\hat{X}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	6968.7	520.8	518.6
6969.8 (m)	6947.7	161.0	164.1
4.95 (m/s)	4.98	0.84	0.83
4.95 (m/s)	4.92	0.73	0.73

The NEES is equal 15.63, out of its 2σ-confidence interval (4 ± 0.06).
Nevertheless, the empirical standard deviation and the empirical mean of the MLE are very close to the corresponding asymptotical CRLB and the true value of the state vector (see Table IV). The relative difference between the CRLB and the covariance matrix of the MLE is about 3.96 %.

Note, that the bias can be estimated once the state vector X has been estimated by $\hat{b} = \frac{1}{N}\sum_{k=1}^{N}\beta_k - \theta_k(\hat{X})$. Then, via Monte Carlo simulation, we compute its empirical MSE: $\mathrm{MSE}_{\mathrm{Emp}}(\hat{b}) = 2.89°$, the same value as σ_{CRLB} of $\hat{b}$ given in Table III.

Remark
Ignoring the presence of the bias when estimating the state vector X can lead to performance degradation. We computed the least squares estimate (LSE), i.e. we sake the argmin of $\frac{1}{2\sigma_\varepsilon^2}\sum_{k=1}^{N}\left[\theta_k(X) - \beta_k\right]^2$. The results are illustrated by Fig. 5.

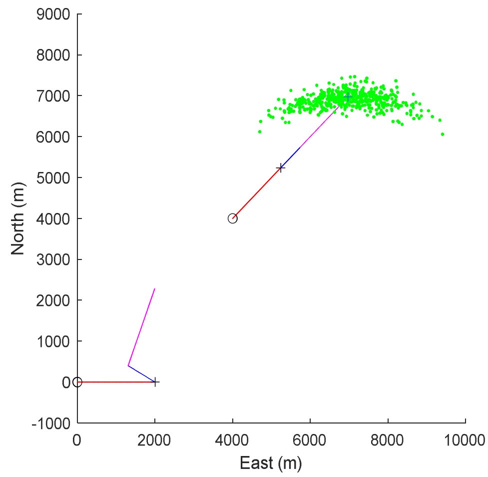

Fig. 5. The trajectories of the observer and of the target, the cloud of estimated positions at final time, when b is ignored in the estimation process.

The gain of the model III can be visually assessed by comparing Fig. 4 and Fig. 5.

The detail can be read in Table V:

TABLE V. PERFORMANCE OF THE LSE, WHEN THE DATA ARE GENERATED ACCORDINGLY TO MODEL III, AND WHEN THE BIAS IS IGNORED.

X	$<\hat{X}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	6982.5	520.8	859.9
6969.8 (m)	6885.8	161.0	201.5
4.95 (m/s)	5.03	0.84	1.29
4.95 (m/s)	4.83	0.73	1.08

D. Discussion

First of all, the three estimates (the MLE in Model I, the MAP in Model II, and the MLE in Model III) are unbiased and their covariance matrices are very close to the CRLB. However, we cannot conclude that they have a Gaussian behavior since the mean values of the normalized estimation error squared (NEES) are out of their respective 2σ-confidence intervals (in our Monte Carlo simulations). So, we sake another coordinates system in which the convergence toward the Gaussian distribution is insured: for example, with the semi-polar coordinates system $[\rho \quad \theta \quad \dot{x}_T \quad \dot{y}_T]^{\mathrm{T}}$ in which ρ and θ are respectively the range between the observer and the target, and the azimuth at the reference time. Things are better, but the NEES is still greater that the upper bound of the confidence interval.

Concerning the comparison between the three models, we can say that the advantage of Model I is that it does not require prior information about b, but its disadvantage resides in the fact that the CRLB is the same whatever the magnitude of the bias. It turns out that in practice the bias is less than 10°. Therefore, using a prior such as the one proposed here is reasonable. Now, a choice must be made between Model II and Model III. In Model III, we do not estimate b, for the same performance, and still obtain results much better than Model I: with $\sigma_b = 20°$ (which is a pessimistic choice), we got promising results (much more accurate than Model I). This is why we recommend adopting Model III in the case of biased measurements.

E. Extention to several legs

The third model can be readily extended to the case of one bias by leg, each leg being independent of the others.
For example, we consider again the same scenario. This time, during leg #i, a bias b_i is added to the unbiased measurements (for *i*=1, 2, 3):

$\beta_k = \theta_k(X) + b_1 + \varepsilon_k,\ \ k = 1, \cdots, M_1$, for leg #1,

$\beta_k = \theta_k(X) + b_2 + \varepsilon_k,\ \ k = M_1 + 1, \cdots, M_2$, for leg #2,

$\beta_k = \theta_k(X) + b_3 + \varepsilon_k,\ \ k = M_2 + 1, \cdots, N$, for leg #3.

The random vector $[b_1 \ \ b_2 \ \ b_3]^{\mathrm{T}}$ is assumed to be distributed as $G(\mathbf{0}_{3,1}, \sigma_b^2 \mathbf{I}_3)$.

The criterion to minimize is then

$$C_{b,3,ext}(X) = \frac{1}{2\sigma_\varepsilon^2}\sum_{k=1}^{N}\left[\theta_k(X) - \beta_k\right]^2 - \frac{\alpha_1}{2\sigma_\varepsilon^2}\left\{\sum_{k=1}^{M_1}\left[\theta_k(X) - \beta_k\right]\right\}^2 - \frac{\alpha_2}{2\sigma_\varepsilon^2}\left\{\sum_{k=M_1+1}^{M_2}\left[\theta_k(X) - \beta_k\right]\right\}^2 - \frac{\alpha_3}{2\sigma_\varepsilon^2}\left\{\sum_{k=M_2+1}^{N}\left[\theta_k(X) - \beta_k\right]\right\}^2$$

with $\alpha_1 = \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + M_1\sigma_b^2}$, $\alpha_2 = \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + (M_2 - M_1)\sigma_b^2}$, and $\alpha_3 = \dfrac{\sigma_b^2}{\sigma_\varepsilon^2 + (N - M_2)\sigma_b^2}$, from which we deduce the expression of the FIM:

$$\mathbf{F}_{b,3,ext} = \sigma_\varepsilon^{-2}\left(\sum_{k=1}^{N} g_k g_k^{\mathrm{T}} - \alpha_1 \sum_{k=1}^{M_1} g_k \sum_{k=1}^{M_1} g_k^{\mathrm{T}} - \alpha_2 \sum_{k=M_1+1}^{M_2} g_k \sum_{k=M_1+1}^{M_2} g_k^{\mathrm{T}} - \alpha_3 \sum_{k=M_2+1}^{N} g_k \sum_{k=M_2+1}^{N} g_k^{\mathrm{T}}\right)$$

Note, however, that we cannot give a closed form of the CRLB as we did in subsection IV.C.
The result of 500 runs of Monte Carlo simulation are displayed in Fig. 6, and presented in Table VI.
Again, the behavior of the MLE is very satisfactory (no bias, and its covariance matrix is close to the CRLB).

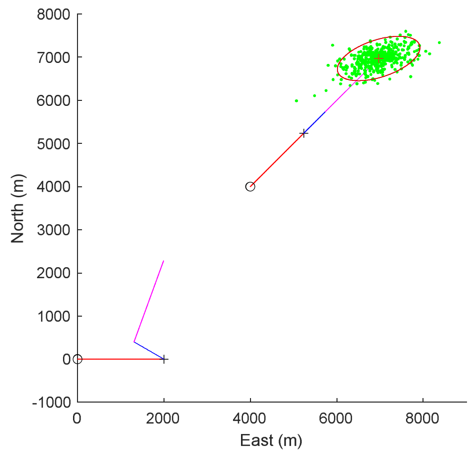

Fig. 6. The trajectories of the observer and of the target, the cloud of position MLEs at final time, together with the 90% confidence ellipse, for Model III, with a random bias during each leg.

We observe that the orientation and the size of the ellipsoid plotted in Fig. 6 differ from those of the ellipsoid given in Fig. 4. This is due to the fact that now, we have to "face" three random biases (the criterion to minimize is not the same). This can be evaluated by comparing the columns "σ_{CRLB}" of Table IV and Table VI.

TABLE VI. PERFORMANCE OF THE MLE IN MODEL III (THE THREE BIASES ARE AN EXTRA NOISE)

X	$<\hat{X}>$	σ_{CRLB}	σ_{emp}
6969.8 (m)	6965.8	445.8	440.1
6969.8 (m)	6959.8	233.8	220.7
4.95 (m/s)	4.96	0.92	0.91
4.95 (m/s)	4.93	0.82	0.79

We computed the NEES: this time, its magnitude is close to the upper bound of its 2σ-confidence interval (4.06): NEES = 4.37.
The first two empirical moments of the MLE are close to the true value of the state vector and the corresponding asymptotical standard deviation given by the CRLB.

V. CONCLUSION

In this paper, we proposed three models to take into account an additive and constant bias of the available data. With the first model, the bias is an unknown and deterministic value. If we admit that the bias benefits from some prior, two models can be constructed: either the magnitude is viewed as a realization of a random variable (of zero mean), or the bias is viewed as an extra noise (of zero mean, too). With this assumption, the measurements are no longer biased, but correlated. The MAP and the MLE for the last two models behave similarly: their performances are equal. However, in the last model, the magnitude of the bias is not directly estimated. These three models are very general: they can be used any time the data are biased. We illustrated our proposal by an example of a bearing-only target motion analysis. Because of the simplicity of the computation and of the good behavior of the MLE in Model III, we definitively recommend using it.

Acknowledgements

The authors are very grateful to the anonymous reviewers (with a special mention to reviewer 3): their comments allowed them to improve the form and substance of the paper.

REFERENCES

[1] A. C. Pérez and C. Jauffret, "Observability of bias of measurements in bearings-only target motion analysis," in: 26th International Conference on Information Fusion, Charleston, July 2023.

[2] S. R. Martin, M. R. Albernathy, and N. Moshtagh, "A Cramér-Rao bound for the estimation of bias with single bearings-only sensor," in: 23rd International Conference on Information Fusion, Pretoria, South Africa, July 2020.

[3] A. C. Pérez and C. Jauffret, "Target motion analysis with biased bearings," unpublished.

[4] X. Cui, K. Yu, and S. Lu, "Direction finding for transient acoustic source based on biased TDOA measurement," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 11, pp. 2442-2453, November 2016.

[5] A. J. Weiss and J. S. Picard, "Network localization with biased range measurements," IEEE Transactions on Wireless Communications, vol. 7, no. 1, pp. 298-304, January 2008.

[6] M. Gavish and E. Fogel, "Effect of bias on bearing-only target location", IEEE Transactions on Aerospace and Electronic Systems, vol. 26, no. 1, pp. 22-26, January 1990.

[7] T. Wang, "Cramer-Rao bound for localization with a priori knowledge on biased range measurements," IEEE Transactions on Aerospace and Electronic Systems, vol. 48, no. 1, pp. 468-476, January 2012.

[8] Y. Bar-Shalom and T.E. Fortmann, Tracking and Data Association, Academic Press, 1988.

[9] Invertible matrix, equation (1)
Homepage: https://en.wikipedia.org/wiki/Invertible_matrix

[10] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Part I. Wiley, 1968

[11] K.B. Petersen and M.S. Pedersen
The matrix cookbook
Homepage:
https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf

[12] S. Nardone, A. Lindgren, and K. Gong, "Fundamental properties and performance of conventional bearings-only target motion analysis," IEEE Transactions on Automatic Control, vol. 29, no. 9, pp. 775-787, September 1984.

[13] Y. Bar-Shalom and X.-R Li , Estimation and Tracking: Principles, Techniques and Software,
Artech House, 1993.