On the Effect of Data Contamination for Multitarget Tracking, Part 1

Adrien Ickowicz
IRISA/INRIA, 35042, Rennes, France

J.-Pierre Le Cadre,
IRISA/CNRS,35042, Rennes,France

Pierre Minvielle
CEA CESTA, BP 2 33114 Le Barp, France

***Abstract*—This paper is concerned with performance prediction of multiple target tracking system. Effects of misassociation are considered in a simple (linear) framework so as to provide closed-form expressions of the probability of correct association. In this paper, we focus on the development of explicit approximations of this probability for a unique false measurements. Rigorous calculations allow us to determine the dimensioning parameters.**

I. Introduction

A fundamental problem in multi-target tracking is to evaluate the performance of the association algorithms. However, it is quite obvious that tracking and association are completely entangled in multi-target tracking. In this context, a key performance measure is the probability of correct association. Generally, track accuracy has been considered without consideration of the association problem. A remarkable exception is the work of K.C. Chang, C.Y. Chong and S. Mori [4], [6].

However, this work is essentially oriented toward a modelling of misassociations via the effect of permutations. Here, we focus on the effect of the ”contamination” of a target track due to extraneous measurements. In fact, a ”contamination” results in a change of estimates of the track parameters, which would render misassociations more likely. It is certain that only measurements situated in the immediate vicinity of the target track would have a severe effect. This the case for dense target environment or (e.g.) decoys.

Here, our analysis is devoted to multiscan association analysis. For this part, the target motion is assumed to be deterministic, while we are concerned with batch performance. In this setup, a linear estimation framework is a simple but efficient way to perform caculations. This paper is organized as follows. In Section 2 the association scenario is presented. We have then to calculate the association costs under the two hypotheses (correct and false association). This is the object of Section 3. The major result of this section is the calculation of a closed-form of these association costs.

The true problem is now to derive from this result an accurate closed-form approximation of the probability of correct association. This is precisely the aim of Section 4, which plays the central role in this paper. The way we derive this approximation is detailed. It is based upon an approximation of the normal density by sums of indicator functions and statistical considerations. The final result is a very simple closed-form approximation, whose accuracy is testimonied by Section 5 (simulation results). Note, however, that this result is limited to a single false association within the whole batch period. It will be shown that the results of Section 3 allow us to consider the general case study. This will be the aim of the companion paper (i.e. Part 2).

II. Problem formulation

A target is moving with a rectilinear and uniform motion. Noisy measurements consisting of Cartesian positions are represented by the points:

$$\tilde{P}_1 = (\tilde{x}_1, \tilde{y}_1)\ ,\ \tilde{P}_2 = (\tilde{x}_2, \tilde{y}_2), \cdots,\ \tilde{P}_N = (\tilde{x}_N, \tilde{y}_N)\ , \quad (1)$$

at time periods t_1, t_2,$\cdots$,t_N, which are called ”*scans*”. Under the correct association hypothesis, the position measurements are the exact Cartesian positions $P_i = (x_i, y_i)$, corrupted by a sequence of independent and identically normally distributed noises (denoted $\varepsilon_{x_i},\ \varepsilon_{y_i}$), i.e.:

$$\tilde{P}_i = (\tilde{x}_i,\, \tilde{y}_i) = (x_i + \varepsilon_{x_i},\, y_i + \varepsilon_{y_i})\ . \quad (2)$$

When a target is (*sufficiently*) isolated from others, there is no ambiguity about the measurement origin. It is not true any more if it happens that a second target comes to stand in the vicinity of the first target. In this case, it becomes possible to make a mistake about the origin of an observation by associating it to the wrong target, thus corrupting target trajectory estimation. But the question is to give a more precise meaning to the term ”sufficiently isolated”.

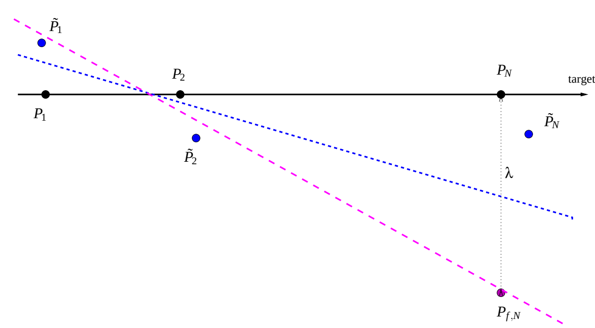

Figure 1. *The association scenario*

Thus, the aim of this article is to give a closed-form expression for the probability of correct association of measurements to a target track, as a function of the number of scans and the distance of the outlier observations. In order to simplify the scenario, we consider that the outlier measurement P_f is located close to the true target position $P_l = (x_l, y_l)$ at time period t_l, with a distance λ^1. The general problem setting and definitions are depicted in fig. 1. Let us denote $\delta_i = t_{i+1} - t_i$, the inter-measurement time, and:

$$\mathbf{v} = (v_x, v_y)^T \ ,$$

the two components of the constant target velocity on the Cartesian axis. In the deterministic case, the target trajectory is then defined by the state vector (x_1, y_1, v_x, v_y).

III. Problem analysis

Under the correct association (ca) hypothesis and denoting $\tau_i \stackrel{\Delta}{=} \delta_1 + \delta_2 + \cdots + \delta_i$, the position measurements $\tilde{P}_i$ are represented by the following equation[2]:

$$\underbrace{\begin{pmatrix} \tilde{x}_1 \\ \tilde{y}_1 \\ \tilde{x}_2 \\ \tilde{y}_2 \\ \vdots \\ \tilde{x}_N \\ \tilde{y}_N \end{pmatrix}}_{\tilde{Z}_{\mathsf{ca}}} = \underbrace{\begin{pmatrix} I_2 & 0_2 \\ I_2 & \tau_1 I_2 \\ \vdots & \vdots \\ I_2 & \tau_{N-1} I_2 \end{pmatrix}}_{\mathcal{X}} \underbrace{\begin{pmatrix} x_1 \\ y_1 \\ v_x \\ v_y \end{pmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{pmatrix} \varepsilon_{x_1} \\ \varepsilon_{y_1} \\ \varepsilon_{x_2} \\ \varepsilon_{y_2} \\ \vdots \\ \varepsilon_{x_N} \\ \varepsilon_{y_N} \end{pmatrix}}_{\tilde{\varepsilon}_{\mathsf{ca}}} \quad (3)$$

With these definitions and under the correct association hypothesis, the measurement model simply stands as follows:

$$\tilde{Z}_{\mathsf{ca}} = \mathcal{X}\ \boldsymbol{\beta} + \tilde{\varepsilon}_{\mathsf{ca}} \ . \quad (4)$$

A. The regression model [2]

Consider the following linear regression model:

$$\tilde{Z} = \mathcal{X}\ \boldsymbol{\beta} + \tilde{\varepsilon} \ , \quad (5)$$

where $\tilde{Z}$ are the data, $\mathcal{X}$ are the regressors and $\boldsymbol{\beta}$ is the vector of parameters, to be estimated. Generally, the estimation of $\boldsymbol{\beta}$ is made via the quadratic loss function:

$$\mathcal{L}_2(\boldsymbol{\beta}) = \left(\tilde{Z} - \mathcal{X}\ \boldsymbol{\beta}\right)^T \ (Z - \mathcal{X}\ \boldsymbol{\beta}) = \|\tilde{Z} - \mathcal{X}\ \boldsymbol{\beta}\|^2 \ . \quad (6)$$

If the matrix $\mathcal{X}^T\mathcal{X}$ is non-singular, then $\mathcal{L}_2(\boldsymbol{\beta})$ is minimum for the unique value $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ such that:

$$\hat{\boldsymbol{\beta}} = (\mathcal{X}^T\ \mathcal{X})^{-1}\mathcal{X}^T\ \tilde{Z} \ . \quad (7)$$

From the estimation $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$, let $\widehat{Z}$ be the estimator of the mean $\mathcal{X}\ \boldsymbol{\beta}$ of the random vector $\tilde{Z}$ defined by:

$$\begin{array}{lcl} \widehat{Z} & = & \mathcal{H}\ Z \ , \\ \text{with:} & & \\ \mathcal{H} & \stackrel{\Delta}{=} & \mathcal{X}(\mathcal{X}^T\mathcal{X})^{-1}\ \mathcal{X}^T \ . \end{array}$$

[1]For the sake of brevity, we assume that measurements are resolved (see [5])

[2]I: identity matrix

The vector of the residuals $\hat{\varepsilon} \stackrel{\Delta}{=} \tilde{Z} - \widehat{Z}$ is given by:

$$\hat{\varepsilon} = \mathcal{M}\ \tilde{Z} \ , \quad (8)$$

with $\mathcal{M} = I - \mathcal{H}$, and I the identity matrix. It is easy to check that $\mathcal{M}$ is a projection matrix (i.e $\mathcal{M}^T = \mathcal{M}$ and $\mathcal{M}^2 = \mathcal{M}$). We also recall the following classical identities, which will be used subsequently [1]:

$$\mathcal{M}\ \mathcal{X} = 0 \ , \quad \text{and: } \hat{\varepsilon} = \mathcal{M}\ \tilde{\varepsilon} \ . \quad (9)$$

B. Evaluation of the correct association probability

Assume that the outlier measurement $P_{f,l} = (x_f, y_f)$ (the lowercase f stands for false association) is located at the time-period l ($1 \leq l \leq N$, see fig. 1):

$$\left\{ \begin{array}{lcl} x_f & = & x_l \ , \\ y_f & = & y_l - \lambda \ . \end{array} \right.$$

The correct association (ca) is then defined by the association of points $\left\{\tilde{P}_1, \cdots, \tilde{\mathbf{P}}_l, \cdots, \tilde{P}_N\right\}$, whereas the wrong association (fa) is defined by $\left\{\tilde{P}_1, \cdots, \mathbf{P}_{f,l}, \cdots, P_N\right\}$.

The vectors of residuals are $\hat{\varepsilon}_{\mathsf{ca}} = \tilde{Z}_{\mathsf{ca}} - \hat{Z}_{\mathsf{ca}}$ under the correct association hypothesis (ca) and $\hat{\varepsilon}_{\mathsf{fa}} = \tilde{Z}_{\mathsf{fa}} - \hat{Z}_{\mathsf{fa}}$ under the false asssociation hypothesis (fa). They are deduced from a linear regression, leading to the following definition of the costs of correct association (denoted $\mathcal{C}_{\mathsf{ca}}$) and false association (denoted $\mathcal{C}_{\mathsf{fa}}$) :

$$\begin{array}{lcl} \mathcal{C}_{\mathsf{ca}} & = & (\tilde{Z}_{\mathsf{ca}} - \widehat{Z}_{\mathsf{ca}})^T\ (\tilde{Z}_{\mathsf{ca}} - \widehat{Z}_{\mathsf{ca}}) \ , \\ & = & \tilde{\varepsilon}_{\mathsf{ca}}^T\ \mathcal{M}\ \tilde{\varepsilon}_{\mathsf{ca}} \ . \end{array} \quad (10)$$

In the same way, we have also:

$$\mathcal{C}_{\mathsf{fa}} = \tilde{\varepsilon}_{\mathsf{fa}}^T\ \mathcal{M}\ \tilde{\varepsilon}_{\mathsf{fa}} \ . \quad (11)$$

Let us define now $\Delta_{f,c}$ the difference between the correct and wrong costs, i.e.:

$$\Delta_{f,c} \stackrel{\Delta}{=} \mathcal{C}_{\mathsf{fa}} - \mathcal{C}_{\mathsf{ca}} \ . \quad (12)$$

Then, the probability of correct association is defined by the probability that $\Delta_{f,c} \geq 0$. **The aim of this article is to give closed-form expressions for this probability**.

Let be $\tilde{\varepsilon}_{\mathsf{com}}$ the vector of components, that vectors $\tilde{\varepsilon}_{\mathsf{ca}}$ and $\tilde{\varepsilon}_{\mathsf{fa}}$ have in common, and define $\tilde{\varepsilon}_l$ and fa_l as the complementary vectors, so that:

$$\tilde{\varepsilon}_{\mathsf{ca}} = \tilde{\varepsilon}_{\mathsf{com}} + \tilde{\varepsilon}_l \ \ , \ \ \tilde{\varepsilon}_{\mathsf{fa}} = \tilde{\varepsilon}_{\mathsf{com}} + \mathsf{fa}_l \ . \quad (13)$$

With these notations, the difference between the correct and wrong costs $\Delta_{f,c}$ can be written:

$$\begin{array}{lcl} \Delta_{f,c} & = & \mathsf{fa}_l^T \mathcal{M}\ \mathsf{fa}_l - (\tilde{\varepsilon}_l)^T \mathcal{M}(\tilde{\varepsilon}_l) \ , \\ & & -2\ (\tilde{\varepsilon}_l - \mathsf{fa}_l)^T \mathcal{M}(\tilde{\varepsilon}_{\mathsf{com}}) \ . \end{array} \quad (14)$$

Since the components of the vector $\tilde{\varepsilon}_{\mathsf{com}}$ are normally distributed and supposed independent, this vector is normal ($\tilde{\varepsilon}_{\mathsf{com}} \sim \mathcal{N}(\boldsymbol{O}, \Sigma_{\mathsf{com}})$), and similarly for $\tilde{\varepsilon}_l$ ($\tilde{\varepsilon}_l \sim \mathcal{N}(\boldsymbol{O}, \Sigma_l)$).

Assuming that the vector $\tilde{\varepsilon}_l$ is set to a **fixed** value $\mathbf{e}_l$, the law of the difference of costs $\mathcal{L}(\Delta_{f,c} | \tilde{\varepsilon}_l = \mathbf{e}_l)$ is normal with characteristics:

$$\mathcal{L}\left(\Delta_{f,c}|\,\tilde{\varepsilon}_l = \mathbf{e}_l\right) = \mathcal{N}\left[\mathsf{fa}_l^T \mathcal{M}\mathsf{fa}_l - (\mathbf{e}_l)^T\mathcal{M}\mathbf{e}_l,\, 4(\mathbf{e}_l - \mathsf{fa}_l)^T\Phi(\mathbf{e}_l - \mathsf{fa}_l)\right] , \tag{15}$$

where: $\Phi \triangleq \mathcal{M}\Sigma_{\mathsf{com}}\mathcal{M}^T$. Integrating this conditional density w.r.t. the Gaussian vector $\tilde{\varepsilon}_l$, yields:

$$P(\Delta_{f,c}(l) \geq 0) = \mathbb{E}_{\tilde{\varepsilon}_l}\left[\mathsf{erfc}\left(\frac{(\mathbf{e}_l)^T\mathcal{M}\mathbf{e}_l - \mathsf{fa}_l^T\mathcal{M}\,\mathsf{fa}_l}{2\sqrt{(\mathbf{e}_l - \mathsf{fa}_l)^T\Phi(\mathbf{e}_l - \mathsf{fa}_l)}}\right)\right] \tag{16}$$

Considering eq. 16, it is not surprising that it is the functional $\Psi(\mathbf{e}_l)$:

$$\Psi(\mathbf{e}_l) = \frac{(\mathbf{e}_l)^T\mathcal{M}\mathbf{e}_l - \mathsf{fa}_l^T\mathcal{M}\,\mathsf{fa}_l}{2\sqrt{(\mathbf{e}_l - \mathsf{fa}_l)^T\Phi(\mathbf{e}_l - \mathsf{fa}_l)}} , \tag{17}$$

which will play the fundamental role for analyzing the probability of correct association. So, the aim of the next subsection is to provide a simpler form of $\Psi(\mathbf{e}_l)$.

C. A closed-form for the $\Psi(\mathbf{e}_l)$ functional

The first step consists in calculating a closed form for the $\Psi(\mathbf{e}_l)$ numerator. Considering the special forms[3] of the vectors $\mathbf{e}_l$ and fa_l, only a closed form expression of the $\mathcal{M}_{l,l}$ 2×2 block matrix is required. Routine calculations yield:

$$\mathcal{M}_{l,l} = \frac{1}{(N+1)(N+2)}\left[1 - \frac{2\left(2N+1-6\,l+\frac{6\,l^2}{N}\right)}{(N+1)(N+2)}\right] I_2 , \tag{18}$$

so that:

$$\begin{aligned}(\mathbf{e}_l)^T\mathcal{M}\mathbf{e}_l - \mathsf{fa}_l^T\mathcal{M}\,\mathsf{fa}_l &= \left[1 - \frac{2\left(2N+1-6l+\frac{6l^2}{N}\right)}{(N+1)(N+2)}\right] , \\ &\quad \times \left(\|\mathbf{e}_l\|^2 - \|\mathsf{fa}_l\|^2\right) .\end{aligned} \tag{19}$$

In the second step, the $\Psi(\mathbf{e}_l)$ denominator is considered. First, it is worth recalling the form of the Φ matrix:

$$\begin{aligned}\Phi &= (I-\mathcal{H})\Sigma_{\mathsf{com}}(I-\mathcal{H}^T) , \\ &= \underbrace{\Sigma_{\mathsf{com}} - \Sigma_{\mathsf{com}}\mathcal{H}^T - \mathcal{H}\Sigma_{\mathsf{com}}}_{\Phi_1} + \mathcal{H}\Sigma_{\mathsf{com}}\mathcal{H}^T ,\end{aligned} \tag{20}$$

and noticing that the 2×2 block matrix $\Phi_1(l,l)$ is zero. Thus, we can restrict to the 2×2 block matrix of the $\mathcal{H}\Sigma_{\mathsf{com}}\mathcal{H}^T$ matrix. Straightforward calculations yield:

$$(N+1)^2(N+2)^2\,\mathcal{H}\Sigma_{\mathsf{com}}\mathcal{H}^T = \mathcal{X}\,\mathcal{C}\Sigma_{\mathsf{com}}\mathcal{C}^T\,\mathcal{X}^T ,$$

with:

$$\mathcal{C} = \begin{pmatrix} (4N+2)I_2 & \dots & (4N+2-6(k-1))I_2 & \dots \\ -\frac{6}{\delta}I_2 & \dots & -\frac{6}{\delta}(1-\frac{2(k-1)}{N})I_2 & \dots \end{pmatrix} \tag{21}$$

Routine calculations then yield a simple expression for the 4×4 matrix $\mathcal{C}\Sigma_{\mathsf{com}}\mathcal{C}^T$:

$$\mathcal{C}\Sigma_{\mathsf{com}}\mathcal{C}^T = \frac{1}{(N+1)^2(N+2)^2}\begin{pmatrix} Q_1(l,N)I_2 & Q_2(l,N)I_2 \\ Q_2(l,N)I_2 & Q_3(l,N)I_2 \end{pmatrix} , \tag{22}$$

from which, we deduce finally:

$$\Phi_{l,l} = \frac{1}{(N+1)^2(N+2)^2}\,\left[Q_1(l,N) + 2l\,\delta Q_2(l,N) + l^2\,\delta^2 Q_3(l,N)\right] I_2 , \tag{23}$$

[3]These two vectors are made of zeros except for x and y l-th components

where the Q_1, Q_2 and Q_3 polynomials have the following expression:

$$\left|\begin{aligned} Q_1(l,N) &= 4N^3 - 50N^2 + N(48l-18) + l(24-36l) + 4 . \\ Q_2(l,N) &= -\tfrac{6}{\delta}\left[N^2 - 5N - 2 + 4l(1 + \tfrac{1}{N} - \tfrac{3l}{N})\right] \\ Q_3(l,N) &= \tfrac{36}{\delta^2}\left[\tfrac{N}{3} - 1 + \tfrac{2}{N}(\tfrac{1}{3} + 2\,l - \tfrac{2\,l}{N^2})\right] .\end{aligned}\right.$$

Gathering the numerator and denominator closed forms, we have just obtained a closed form expression for Ψ_l:

$$\begin{aligned}\Psi(e_l) &= \frac{\left[(N+1)(N+2) - 2(2N+1-6\,l+\frac{6\,l^2}{N})\right]}{2\left[Q_1(l,N)+2l\delta Q_2(l,N)+l^2\delta^2 Q_3(l,N)\right]}\left(\frac{\|\mathbf{e}_l\|^2 - \|\mathsf{fa}_l\|^2}{\|\mathbf{e}_l - \mathsf{fa}_l\|}\right) , \\ &\propto \frac{N^2}{2\left(N^3 - 3lN^2 + 3l^2N\right)^{1/2}}\left(\frac{\|\mathbf{e}_l\|^2 - \|\mathsf{fa}_l\|^2}{\|\mathbf{e}_l - \mathsf{fa}_l\|}\right) ,\ N \text{ great} ,\end{aligned} \tag{24}$$

Considering eq. 24 (last row), we can notice that the variations of $\Psi(e_l)$ as a function of l as a function of l are not very important. Actually, it is easily seen that $\frac{N^2}{2\,(N^3-3lN^2+3l^2N)^{1/2}}$ is varying between $\frac{\sqrt{N}}{2}$ and $\frac{\sqrt{N}}{4}$ as l varies between 0 and N. Now, the erfc function is quite flat for large values of N, which means that $P(\Delta_{f,c}(e_l) \geq 0)$ is almost independent of the value of l.

This closed-form of $\Psi(e_l)$ is instrumental for deriving a closed-form approximation of $P(\Delta_{f,c} \geq 0)$.

D. Multiple false associations

The previous calculations can be rather easily extended to **multiple** false associations. Let $\mathsf{FA}_K = (l_k)_{k=1}^K$, be the vector made by indices l_k of the (possible) false associations. A closed-form expression of the numerator of eq. 17 is:

$$\begin{aligned}&\mathbf{e}_K^T\mathcal{M}\mathbf{e}_K - \mathsf{FA}^T\mathcal{M}\mathsf{FA} = \\ &\sum_{k=1}^K\sum_{k'=1}^K\left(\mathbf{1}_{\{k=k'\}} - \frac{2(2N+1-3\,l_{k'}-3\,l_k+\frac{6\,l_k^2}{N})}{(N+1)(N+2)}\right) \\ &\times\left(\langle\mathbf{e}_{l_k},\mathbf{e}_{l'_k}\rangle - \langle\mathsf{fa_k},\mathsf{fa_{k'}}\rangle\right) \\ &= \sum_{k=1}^K\sum_{k'=1}^K\alpha_N(k,k')\left(\langle\mathbf{e}_{l_k},\mathbf{e}_{l'_k}\rangle - \langle\mathsf{fa_k},\mathsf{fa_{k'}}\rangle\right) .\end{aligned} \tag{25}$$

Similarly, for the denominator D_{Ψ_K} of $\Psi_{\mathsf{FA_K}}$, we have:

$$D_{\Psi_K} = 2\sqrt{\sum_{k=1}^K\sum_{k'=1}^K\theta(l_k,l_{k'})\,\langle\mathbf{e}_{l_k} - \mathsf{fa}_{l_k},\mathbf{e}_{l_{k'}} - \mathsf{fa}_{l_{k'}}\rangle} ,$$

with:

$$\begin{aligned}&(N+1)^2(N+2)^2\,\theta(l_k,l_{k'}) = \\ &\left[Q_1^*(\mathsf{FA}_K,N) + (l_k + l_{k'})\,Q_2^*(\mathsf{FA}_K,N) + l_k l_{k'}\,Q_3^*(\mathsf{FA}_K,N)\right] \\ &-\left(\alpha_N(l_k,l_{k'}) + \alpha_N(l_k,l_{k'})\right)(N+1)^2(N+2)^2.\end{aligned} \tag{26}$$

The polynoms Q_1^*, Q_2^* and Q_3^* stand as follows:

$$\begin{aligned}Q_1^*(\mathsf{FA}_K,N) &= \textstyle\sum_{l=0,l\notin\mathsf{FA_K}}^N(4N+2-6l)^2 , \\ Q_2^*(\mathsf{FA}_K,N) &= -\tfrac{6}{\delta}\left[\textstyle\sum_{l=0,l\notin\mathsf{FA_K}}^N(4N+2-6l)(1-\tfrac{2\,l}{N})\right] , \\ Q_3^*(\mathsf{FA}_K,N) &= \tfrac{36}{\delta^2}\left[\textstyle\sum_{l=0,l\notin\mathsf{FA_K}}^N(1-\tfrac{2\,l}{N})^2\right] .\end{aligned}$$

Finally, we have obtained the following closed-form of $\Psi_{\mathsf{FA_K}}$:

$$\Psi_{\mathsf{FA}_K} = \frac{\sum_{k=1}^{K}\sum_{k'=1}^{K} \alpha_N(k,k')\left(\langle \mathbf{e}_{l_k}, \mathbf{e}_{l'_k}\rangle - \langle \mathsf{fa_k}, \mathsf{fa_{k'}}\rangle\right)}{2\sqrt{\sum_{k=1}^{K}\sum_{k'=1}^{K} \theta(l_k, l_{k'})\,\langle \mathbf{e}_{l_k} - \mathsf{fa}_k, \mathbf{e}_{l_{k'}} - \mathsf{fa}_{k'}\rangle}} \; . \tag{27}$$

Again, this expression is remarkably simple.

E. Diffusive Target

Up to now, the target modelling was deterministic (see eq. 3). However, this assumption is not realistic, especially if the duration of the scenario is great. Actually, denoting X_k the 2-dimensional vector made of target position at time period k, and V_k its velocity vector. The following modelling (discrete Orenstein-Uhlenbeck) is considered:

$$\begin{cases} X_k &= X_{k-1} + V_k \\ V_k &= V_{k-1} + A_k \end{cases} \tag{28}$$

where A_k is a white noise with variance σ_a^2. The diffusive target scenario is depicted in 2. However, notice that opposite to (Kalman) filtering our aim is restricted to the effect analysis of this target trajectory randomization within the common regression framework.

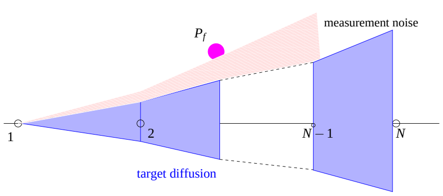

Figure 2. *The scenario for a diffusive target*

Integrating, an equivalent expression for the above target model is:

$$\begin{aligned} X_k &= X_0 + kV_0 + \sum_{j=1}^{k} (k-j+1)\,A_j \;, \\ &= X_0 + k\,V_0 + W_k \;. \end{aligned} \tag{29}$$

We do have w_k a gaussian noise with variance $\frac{k(k+1)(2k+1)}{6}\sigma_a^2$. We can then follow the same way as for the previous model, and thus we have a few modifications for Σ_{COM} and then Ψ. Errors are now heteroscedastic, and then the changes are:

$$\begin{aligned} d_j &= \tfrac{j(j+1)(2j+1)}{6} \;, \\ \tilde{Q}_1(l,N) &= \textstyle\sum_{j=0, j\neq l}^{N} d_j^2 (4N+2-6j)^2 \;, \\ \tilde{Q}_2(l,N) &= -\tfrac{6}{\delta}\left[\textstyle\sum_{j=0,j\neq l}^{N} d_j^2 (4N+2-6j)(1-\tfrac{2j}{N})\right] \;, \\ \tilde{Q}_3(l,N) &= \tfrac{36}{\delta^2}\left[\textstyle\sum_{j=0,j\neq l}^{N} d_j^2 (1-\tfrac{2j}{N})^2\right] \;, \\ e\tilde{D}(l,N,\delta) &= \text{Extra-diag terms equivalent to } \tilde{Q}_1(l,N) \end{aligned} \tag{30}$$

The functional Ψ becomes:

$$\Psi = \frac{\left[(N+1)(N+2)-2(2N+1-6l+\frac{6l^2}{N})\right]\left(\|e_l\|^2 - \|\mathsf{fa_l}\|^2\right)}{2\sigma_a\,\|e_l - \mathsf{fa_l}\| \cdot \sqrt{\left[\tilde{\mathsf{Q}}_1(\mathsf{l,N}) + 2\mathsf{l}\delta\tilde{\mathsf{Q}}_2(\mathsf{l,N}) + \mathsf{l}^2\delta^2\tilde{\mathsf{Q}}_3(\mathsf{l,N}) + \mathsf{e}\tilde{\mathsf{D}}(\mathsf{l,N},\delta)\right]}} \tag{31}$$

There is a fundamenatl difference between this expression and the previous one. The greater N is, the smaller Ψ begins. In fact, Ψ is equivalent to $\frac{1}{\sigma_a N^3}$. And then, if N is great, the probability of correct association becomes close to 0.5.

F. An extension to radar measurements

Up to now, we assumed that (Cartesian) ε_x and ε_y measurement errors were independent. Actually, this is not true for an important context like the radar one. Actually, the aim of an active localization device is to estimate the range (say r) and bearing (say θ) of a target. It is also quite reasonable to assume that range and bearing measurements are uncorrelated. However, even under this assumption, ε_x and ε_y are correlated [3]. Thus the (2×2) I matrix must be replaced by the A matrix [3]:

$$A = \begin{pmatrix} r^2\sigma_\theta^2 \sin^2(\theta) + \sigma_r^2\cos^2(\theta) & (\sigma_r^2 - r^2\sigma_\theta^2)\sin(\theta)\cos(\theta) \\ (\sigma_r^2 - r^2\sigma_\theta^2)\sin(\theta)\cos(\theta) & r^2\sigma_\theta^2\cos^2(\theta) + \sigma_r^2\sin^2(\theta) \end{pmatrix} \tag{32}$$

where r and θ are the range and bearing of the target, while σ_r^2 and σ_θ^2 are corresponding variances. Note, that no bias is considered since it is assumed that a preliminary debiaising step has been applied at the measurement level [3]. Since the Σ_{com} matrix plays a fundamental role in our calculation, we have to modify it accordingly. The matrix Σ_{com} then becomes block-diagonal, i.e.:

$$\Sigma_{\mathsf{com}} = \text{block-diag}\left(A, \cdots, A, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, A, \cdots, A\right) \; . \tag{33}$$

Quite similarly to the uncorrelated measurements case, we obtain:

$$\mathcal{C}\Sigma_{\mathsf{com}}\mathcal{C} = \frac{1}{(N+1)^2(N+2)^2}\begin{pmatrix} Q_1(l,N)A & Q_2(l,N)A \\ Q_2(l,N)A & Q_3(l,N)A \end{pmatrix} \; . \tag{34}$$

So, the only change is that $\|\mathbf{e}_l - \mathsf{fa}_l\|^2$ is replaced by $(\mathbf{e}_l - \mathsf{fa}_l)^T A (\mathbf{e}_l - \mathsf{fa}_l)$. The numerator is left unchanged since it does not involve Σ_{com}.

IV. Closed-form approximations of the probability of correct association

For the sake of simplicity, the error measurement components $\tilde{\varepsilon}_{x,l}$ and $\tilde{\varepsilon}_{y,l}$ will be simply denoted as x and y. We have now to deal with convenient approximations of the association cost difference $\Delta_{f,c} \triangleq \mathcal{C}_{\mathsf{fa}} - \mathcal{C}_{\mathsf{ca}}$. We restrict us to a single outlier measurement. At this point, it is worth recalling that it is *conditionally* distributed as a normal density (see eq. 15):

$$\mathcal{N}\left[\mathsf{fa}_l^T \mathcal{M}\mathsf{fa}_l - (\mathbf{e}_l)^T\mathcal{M}\mathbf{e}_l, 4(\mathbf{e}_l - \mathsf{fa}_l)^T\Phi(\mathbf{e}_l - \mathsf{fa}_l)\right] \tag{35}$$

This section will be divided in three subsections corresponding to the main steps of the development. We will now turn toward the results of section III-C.

A. Approximating the normal density by a sum of indicator functions

A first step will consist in approximating $\mathcal{L}(\Delta_{f,c} | \tilde{\varepsilon}_l = \mathbf{e}_l)$ by a sum of n indicator functions. Thus considering a "3σ" support of this approximation centered on the mean m of this normal density, i.e. $[m-3\sigma, m+3\sigma]$ leads to:

$$\mathcal{L}\left(\Delta_{f,c} \middle| \tilde{\boldsymbol{\varepsilon}}_l = \mathbf{e}_l\right) \simeq \sum_{i=1}^{n} \frac{\gamma_i}{6\frac{i}{n}\mathsf{den}(x,y)}\ \varphi_i(x,y)\ ,$$

where:

$$\varphi_i(x,y) \triangleq \mathbf{1}_{\Delta_{f,c}\in[b^i_{\mathsf{inf}}(x,y)\ ,\ b^i_{\mathsf{sup}}(x,y)]}\ ,\ \mathbf{e}_l = (x,y)^T\ . \tag{36}$$

This means that the supports of these n indicator functions vary from $[-3\,\frac{\sigma}{n}, 3\,\frac{\sigma}{n}]$, to $[-3\sigma, 3\sigma]$, whose parameters are defined by:

$$\begin{aligned}
\mathsf{den}(x,y) &= 2\sqrt{(\mathbf{e}_l - \mathsf{fa}_l)^T\Phi(\mathbf{e}_l - \mathsf{fa}_l)}\ ,\\
&= 2\sqrt{\beta_N\ [(x)^2 + (y+\lambda)^2\ ]}\ ,\\
b^i_{\mathsf{sup}}(x,y) &= \mathsf{fa}^T\mathcal{M}\mathsf{fa} - (\mathbf{e}_l)^T\mathcal{M}(\mathbf{e}_l) + \frac{3i}{n}\ \mathsf{den}\ ,\\
&= \alpha_N(x^2+y^2-\lambda^2) + \frac{3i}{n}\ \mathsf{den}(x,y)\\
b^i_{\mathsf{inf}}(x,y) &= \mathsf{fa}^T\mathcal{M}\mathsf{fa} - (\mathbf{e}_l)^T\mathcal{M}(\mathbf{e}_l) - \frac{3i}{n}\ \mathsf{den}\ ,\\
&= \alpha_N(x^2+y^2-\lambda^2) - \frac{3i}{n}\ \mathsf{den}(x,y)\ ,
\end{aligned} \tag{37}$$

where the scalar parameters $\alpha_N(l)$ and $\beta_N(l)$ are:

$$\alpha_N(l) = [\tfrac{2(2N+1-6l+6\frac{l^2}{N})}{(N+1)(N+2)} - 1]\ ,$$

$$\beta_N(l) = \tfrac{Q_1(l,N)+2l\delta Q_2(l,N)+l^2\delta^2 Q_3(l,N)}{(N+1)^2(N+2)^2}\ .$$

For instance, for $l = N$, we have more simply ($l = N$):

$$\alpha_N = \tfrac{N(1-N)}{(N+1)(N+2)} \approx -1\ ,$$

$$\beta_N = \tfrac{4N^3+226N^2-66N+4}{(N+1)^2(N+2)^2} \approx \tfrac{4}{N}\ (N \gg 1)\ .$$

The definition and meaning of the φ_i functions are represented on fig. 3. With these definitions, we thus have the

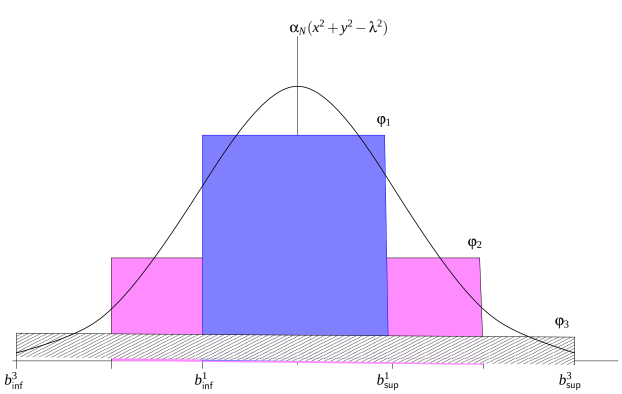

Figure 3. *The approximation scheme: the φ_i functions*

following approximation:

$$P\left(\Delta_{f,c} \geq 0 \middle| \tilde{\boldsymbol{\varepsilon}}_l = \mathbf{e}_l\right) =$$

$$\sum_{i=1}^{n}\gamma_i\left[\frac{b^i_{\mathsf{sup}}}{2\frac{3i}{n}\mathsf{den}}\ \mathbf{1}_{b^i_{\mathsf{sup}}\geq 0} - \frac{b^i_{\mathsf{inf}}}{2\frac{3i}{n}\mathsf{den}}\ \mathbf{1}_{b^i_{\mathsf{inf}}\geq 0}\right] =$$

$$\sum_{i=1}^{n}\frac{\gamma_i}{2}\left(\mathbf{1}_{b^i_{\mathsf{sup}}(x,y)\geq 0} + \mathbf{1}_{b^i_{\mathsf{inf}}(x,y)\geq 0}\right)$$

$$+\tfrac{\alpha_N\ n}{12\sqrt{\beta_N}}\ \tfrac{(x^2+y^2-\lambda^2)}{\sqrt{((x-\lambda)^2+y^2)}}\ \sum_{i=1}^{n}\frac{\gamma_i}{i}\ \left(\mathbf{1}_{b^i_{\mathsf{sup}}(x,y)\geq 0} - \mathbf{1}_{b^i_{\mathsf{inf}}(x,y)\geq 0}\right)\ .$$

The $\{\gamma_i\}$ coefficients are obtained as the solution of an optimization problem (see Appendix A). At least they must satisfy to the following conditions:

$$\begin{aligned}
\sum_{i=1}^{n}\gamma_i &= 1\ ,\\
\sum_{i=1}^{n}\frac{\gamma_i}{6\frac{i}{n}\mathsf{den}} &= \frac{1}{\mathsf{den}\sqrt{2\pi}}e^0\ .
\end{aligned} \tag{38}$$

We stress that these $\{\gamma_i\}$ coefficients are considered as *fixed* whatever the value of the $\mathbf{e}_l$ vector. So, integrating over the possible values of the $\mathbf{e}_l$ vector, we obtain:

$$\begin{aligned}
P(\Delta_{f,c} \geq 0) &= \int_{\mathbb{R}^2} P\left(\Delta_{f,c} \geq 0 \middle| \tilde{\boldsymbol{\varepsilon}}_l = \mathbf{e}_l\right)\ dx\,dy\ ,\\
&= \sum_{i=1}^{n}\frac{\gamma_i}{2}A_i + \frac{\alpha_N}{\sqrt{\beta_N}}\frac{n}{12}\sum_{i=1}^{n}\frac{\gamma_i}{i}\ B_i\ ,
\end{aligned} \tag{39}$$

where:

$$A_i = \int_{\mathbb{R}^2}\mathcal{N}_{(0,1)}(x,y)\left[\mathbf{1}_{f(x,y)\geq -\frac{6i\sqrt{\beta_N}}{n\,\alpha_N}} + \mathbf{1}_{f(x,y)\geq \frac{6i\sqrt{\beta_N}}{n\,\alpha_N}}\right]\ dxdy,$$

$$B_i = \int_{\mathbb{R}^2}\mathcal{N}_{(0,1)}(x,y)f(x,y)\left[\mathbf{1}_{f(x,y)\geq -\frac{6i\sqrt{\beta_N}}{n\,\alpha_N}} - \mathbf{1}_{f(x,y)\geq \frac{6i\sqrt{\beta_N}}{n\,\alpha_N}}\right]dxdy\ ,$$

and:

$$f(x,y) = \tfrac{x^2+y^2-\lambda^2}{\sqrt{x+(y+\lambda)^2}}\ . \tag{40}$$

For reasons which will clearly appear soon, it is worth to rewrite the A_i and B_i integrals:

$$\begin{aligned}
B_i &= \int_{-\frac{6i\sqrt{\beta_N}}{n\,\alpha_N}\leq f(x,y)\leq \frac{6i\sqrt{\beta_N}}{n\,\alpha_N}}\mathcal{N}_{(0,1)}(x,y)\ f(x,y)\ dxdy\ ,\\
A_i &= \int_{-\frac{6i\sqrt{\beta_N}}{n\,\alpha_N}\leq f(x,y)\leq \frac{6i\sqrt{\beta_N}}{n\,\alpha_N}}\mathcal{N}_{(0,1)}(x,y)\ \ dxdy;,\\
&+\ 2\int_{f(x,y)\leq \frac{-6i\sqrt{\beta_N}}{n\,\alpha_N}}\mathcal{N}_{(0,1)}(x,y)\ \ dxdy\ .
\end{aligned} \tag{41}$$

So, now the problem we have to face is to calculate the B_i integrals.

B. *Approximating the B_i integrals*

It is clear that deriving a general closed-form expression for B_i (or A_i) is hopeless[4]. However, an accurate closed-form approximation can be obtained thanks to the

[4] There does not exist a primitive function of $\mathcal{N}_{(0,1)}(x,y)\ f(x,y)$ and the integral is **implicitely** defined

following remark. When the scan number N becomes great, then the ratio $\rho = \frac{\sqrt{\beta_N}}{\alpha_N}$ is close to zero. Now the numerator of the f function is zeroed on a circle (equation $x^2 + y^2 = \lambda^2$). This leads us to consider the following parametrization of the (x, y)-plane.

$$\| \; x = (-\lambda + \varepsilon)\ \sin(\theta)\ , y = (-\lambda + \varepsilon)\ \cos(\theta)\ . \tag{42}$$

The function $f(x, y)$ is then changed in a $f(\varepsilon, \theta)$ function defined below (see Appendix B), which leads to the following changes for the integral:

$$\left| \begin{array}{l} f(\varepsilon,\theta) = \frac{-\varepsilon(2\lambda-\varepsilon)}{\sqrt{4\lambda\sin^2(\theta/2)(\lambda-\varepsilon)+\varepsilon^2}} \\ \text{moreover:} \\ \exp\left(-\frac{x^2+y^2}{2}\right) = \exp\left(-\frac{(\lambda-\varepsilon)^2}{2}\right)\ , \\ dxdy = |-\lambda+\varepsilon|\ d\varepsilon d\theta\ . \end{array} \right. \tag{43}$$

Now, since we are considering only the small values of the f function (num.$(f) = -\varepsilon(2\lambda-\varepsilon)$), it is quite legitimate [5] to restrict our analysis to small values of ε. More precisely, we assume $\varepsilon \ll \lambda$. Now, the **second** order expansion of the $f(\varepsilon, \theta)$ is:

$$f(\varepsilon,\theta) \stackrel{2}{=} \frac{-\varepsilon}{|\sin(\theta/2)|}\ . \tag{44}$$

Practically, this is rather important since the integration domain which was previously implicitely defined is now **explicitely** defined; i.e.:

$$\left\{ \begin{array}{l} \frac{-3i|\sin(\theta/2)|\sqrt{\beta_N}}{n\alpha_N} \le \varepsilon \le \frac{3i|\sin(\theta/2)|\sqrt{\beta_N}}{n\alpha_N}\ , \\ 0 \le \theta \le \pi\ . \end{array} \right. \tag{45}$$

The accuracy of this approximation illustrated by fig. 4. In

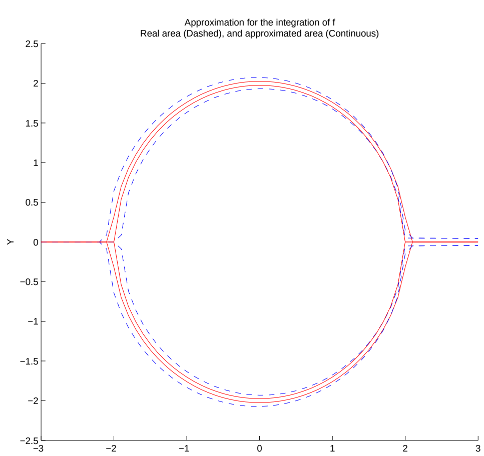

Figure 4. *The $f(x, y)$ function and its approximation (real: dashed; approximation: continuous)*

[5]Actually, there are two roots to the equation num.$f(\varepsilon) = 0$, $\varepsilon = 0$ and $\varepsilon = 2\lambda$. However, both are represented by a unique transformation (see eq. 43)

a second step, we consider a second order expansion of the integrand $F(\varepsilon, \theta)$ (i.e. $f(\varepsilon, \theta)\mathcal{N}(\varepsilon, \theta)\,|J(\varepsilon, \theta)|$), yielding:

$$F(\varepsilon,\theta) \stackrel{2}{=} -\lambda\varepsilon\frac{e^{-\lambda^2/2}}{|\sin(\theta/2)|} + \frac{(1-2\lambda^2)}{2\,|\sin(\theta/2)|}e^{-\lambda^2/2}\ \varepsilon^2\ . \tag{46}$$

Considering on the first hand the effect of changing ε into $-\varepsilon$ for this 2-nd order expansion and the integration domain on the second one, the effect of the ε term is zero, so that:

$$\begin{aligned} B_i &= \frac{1}{2\pi}\ \int_\theta\int_\varepsilon \frac{(1-2\,\lambda^2)}{2|\sin(\theta/2)|}\ e^{-\frac{(\lambda)^2}{2}}\ \varepsilon^2 d\lambda d\epsilon \\ &= \frac{(1-2\lambda^2)}{2\pi}\ e^{-\frac{(\lambda)^2}{2}}\ \frac{\beta_i^3}{3}\ \int_\theta (\sin(\theta/2))^2\, d\theta\ , \end{aligned} \tag{47}$$

where $\beta_i = \frac{3i\sqrt{\beta_N}}{n\alpha_N}$. Thus, a very simple closed-form approximation has been obtained

$$\frac{\alpha_N}{\sqrt{\beta_N}}\frac{n}{12}\sum_i^n\frac{\gamma_i}{i}B_i = \frac{3(1-2\lambda^2)e^{-\lambda^2/2}}{32\ n^2}\frac{\beta_N}{\alpha_N^2}\sum_i^n i^2\ \gamma_i\ . \tag{48}$$

C. *Approximating the A_i integrals*

We have now to turn toward the A_i. First, we remark that:

$$\begin{aligned} &\mathbf{1}_{f(x,y)\le-\beta_i} + \mathbf{1}_{f(x,y)\le\beta_i} = \\ &\qquad \mathbf{1}_{-\beta_i\le f(x,y)\le\beta_i} + 2\left(\mathbf{1}_{f(x,y)\le 0} - \mathbf{1}_{-\beta_i\le f(x,y)\le 0}\right)\ , \end{aligned} \tag{49}$$

so that, we have:

$$\begin{aligned} A_i &= 2\underbrace{\int_{\mathbb{R}^2}\mathcal{N}(0,1)(\mathbf{1}_{f(x,y)\le 0} - \mathbf{1}_{-\beta_i\le f(x,y)\le 0})dxdy}_{A_{i,1}}\ , \\ &\quad + \underbrace{\int_{\mathbb{R}^2}\mathcal{N}(0,1)\mathbf{1}_{-\beta_i\le f(x,y)\le\beta_i}dxdy}_{A_{i,2}}\ . \end{aligned}$$

We use the same change of variable (see eq. 43) as previously. For the $A_{i,1}$ integral the normal density is integrated over the (ε, θ) domain $[0, 2\lambda] \times [0, 2\pi]$; while for the $A_{i,2}$ integral it is $[0, \beta_i\,|\sin(\theta/2)|] \times [0, 2\pi]$. We thus have:

$$\begin{aligned} A_{i,1} &= \frac{1}{\pi}\int_0^{2\pi}[e^{-(\lambda-\varepsilon)^2/2}]_0^\lambda - [e^{-(\lambda-\varepsilon)^2/2}]_\lambda^{2\lambda}d\theta\ , \\ &\quad + \frac{1}{\pi}\int_0^\pi[e^{-(\lambda-\varepsilon)^2/2}]_0^{\beta_i|\sin(\theta/2)|}d\theta\ , \\ &\simeq 2 - e^{-\lambda^2/2}[2 + 2\beta_i - \frac{(\lambda^2-1)}{4}\beta_i^2] \end{aligned} \tag{50}$$

For the $A_{i,2}$ integral, we proceed in the same way that for B_i,i.e.:

$$\begin{aligned} A_{i,2} &= \frac{1}{2\pi}\int_0^{2\pi}\left[e^{-\frac{(\lambda-\beta_i|\sin(\theta/2)|)^2}{2}} - e^{-\frac{(\lambda+\beta_i|\sin(\theta/2)|)^2}{2}}\right]d\theta\ , \\ &\simeq \frac{2\lambda e^{-\lambda^2/2}}{\pi}\beta_i\ . \end{aligned} \tag{51}$$

Again, we have used the hypothesis $\beta_i \ll 1$ to obtain an accurate expansion of the $A_{i,2}$ integrand $g(\lambda, \beta)$, i.e:

$$g(\lambda, \beta) = 2\lambda\beta e^{-\lambda^2/2} + \mathbf{o}(\beta^3) \ .$$

Gathering the above results, we have just obtained a closed form approximation of the A_i term:

$$A_i = \frac{-2\pi + (2\lambda - 2\pi)\beta_i + \frac{\pi}{4}(\lambda^2 - 1)\beta^2}{\pi} e^{-\lambda^2/2} \ , \qquad (52)$$

D. The closed-form approximation of $P(\Delta_{f,c} \geq 0)$

We are now in position to present the aim of this paper, i.e. a closed-form approximation of $P(\Delta_{f,c} \geq 0)$:

$$\boxed{P(\Delta_{f,c} \geq 0) = 1 + (a + b\,\lambda + c\,\lambda^2) e^{-\frac{\lambda^2}{2}}}$$

with:

$$\left| \begin{array}{lcl} a & = & -\dfrac{\left(1 + \displaystyle\sum_i \frac{\gamma_i}{i} \frac{\sqrt{\beta_N}}{\alpha_N}\right) + \frac{66\pi}{32n^2} \frac{\beta_N}{\alpha_N^2} \displaystyle\sum_i i^2\, \gamma_i}{2\pi} , \\ b & = & \dfrac{\frac{6}{n} \frac{\sqrt{\beta_N}}{\alpha_N} \displaystyle\sum_i i\, \gamma_i}{2\pi} , \\ c & = & \frac{15}{16n^2} \, \frac{\beta_N}{\alpha_N^2} \displaystyle\sum_i i^2\, \gamma_i \ . \end{array} \right. \qquad (53)$$

This formula is quite simple and relevant. We can notice also that $P(\Delta_{f,c} \geq 0)$ is independent of the kinematic scenario parameters, since it involves only the ratio λ/σ (denoted here λ)., and the scan number N (via α_N and β_N).

E. The case of a random λ

Up to now, it was assumed that the parameter λ was deterministic. However, it is more realistic to model this seducing measurement by a normal density $\mathcal{N}(\lambda_0, \sigma_0)$. Let $\bar{\Delta}_{f,c}$ be the (extended) cost difference for this λ modelling, conditioning on λ, we then have:

$$\begin{array}{lcl} P(\bar{\Delta}_{f,c} \geq 0) & = & \mathbb{E}_\lambda \left[P(\Delta_{f,c} \geq 0) | \lambda \right] \ , \\ \text{with:} & & \\ P(\Delta_{f,c} \geq 0) & = & 1 + (a + b\lambda + c\lambda^2) e^{-\lambda^2/2} \ . \end{array} \qquad (54)$$

Performing straightforward calculations, we obtain:

$$P(\bar{\Delta}_{f,c} \geq 0) = 1 + \frac{1}{\sqrt{\sigma_0^2 + 1}} \left[a + b\bar{\lambda}_0 + c(\bar{\lambda}_0^2 + s_0^2) \right] e^{-\frac{\lambda_0^2}{2(\sigma_0^2 + 1)}} \ ,$$

where:

$$\bar{\lambda}_0 = \frac{1}{\sigma_0^2 + 1} \lambda_0 \qquad , \qquad s_0^2 = \frac{\sigma_0^2}{\sigma_0^2 + 1} \qquad (55)$$

V. Simulation Results

Once we get the main result (eq. 53) we have to test the accuracy of our approximations. For doing that, we just have to consider the variations of the two dimensioning parameters (λ and N). For the first one (λ), the number of scans (N) is a fixed value ($N = 20$ and $N = 40$). Then, we compare the exact value of $P(\Delta_{f,c} \geq 0)$ and its approximation as given by eq. 53, for increasing values of the λ parameter. Note that λ represents in fact the ratio λ/σ where λ is the distance between the exact target position and the position of the "false" target, while σ is the observation noise standard deviation. The result is displayed on fig. 5. We can see that our approximation

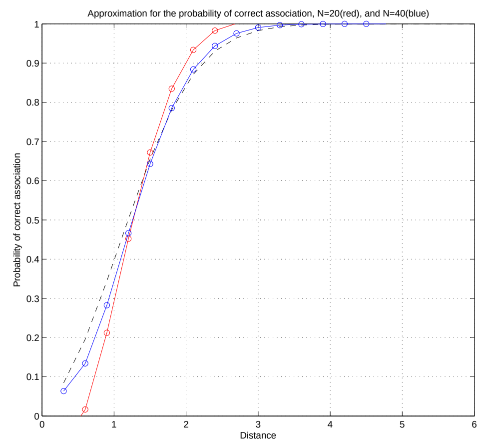

Figure 5. *The probability of correct association (dashed) and approximated (in red: $N = 20$, in blue $N = 40$), versus λ.*

is quite good, in general, but is better as N increases. This is not surprising, especially if we remind that our approximations were based on the fact that the integration bounds b_i were small, meaning that N was great. Thus, it remains to analyze the effect of the N parameter. This is shown in figure 6.

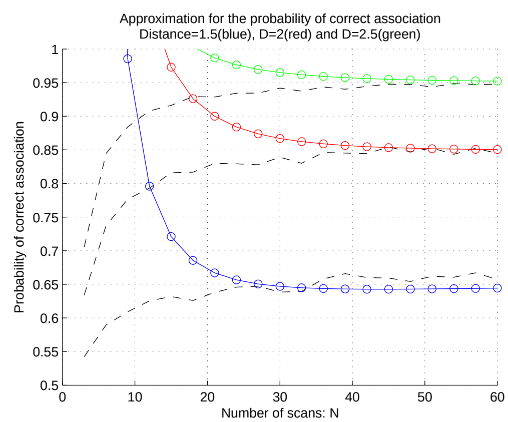

Figure 6. *The probability of correct association (exact: dashed) and approximated (continuous) for various values of λ: in blue $\lambda = 1.5$, in red $\lambda = 2.$, in green $\lambda = 2.5$.*

Results are restricted to fixed values of λ, that is equal to 1.5, 2 and 2.5 because they are the most interesting values, representing the more common association problem. We can see that for a number of scan greater than 30, the approximation is very good. The difference is less than 0.05, which is quite satisfactory. Moreover, for greater values of N, exact values and approximations cannot be distinguished.

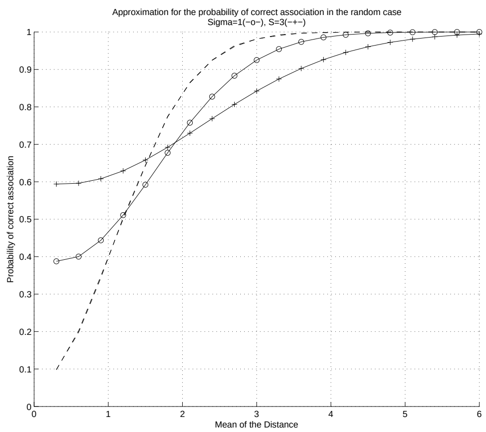

Figure 7. *The probability of correct association for a random λ. Dashed: deterministic λ, continuous: random λ.*

Finally, we present the results for a random λ on fig. 7 (see subsection IV-E). The values of $P(\bar{\Delta}_{f,c} \geq 0)$ are plotted on the y-axis, versus the mean value of λ (λ_0), for two values of the σ_0 parameters (1 and 3). Not surprisingly, the effect of this randomization is far to be negligible.

VI. CONCLUSION

Deriving accurate closed-form approximations of the probability of correct association is of fundamental importance for understanding the behavior of data association algorithms. However, though numerous association algorithms are available, performance analysis is rarely considered from an analytical point of view. Actually, this is not too surprising when we consider the difficulties we have to face even in the simplistic framework of linear regression.
So, the main contribution of this paper is to show that such derivations are possible. This has been achieved via elementary though rigourous derivations, developed in a unique framework. Multiple extensions and applications render it quite attractive.

VII. APPENDIX A

This appendix deals with the calculation of the coefficients γ_i for the least square criterion. Denoting φ_i ($i = 1, \cdots, n$) the functions defined by $\varphi_i \stackrel{\Delta}{=} \frac{n}{6i \ \mathsf{den}} \mathbf{1}_{[b^i_{\mathsf{inf}}, b^i_{\mathsf{sup}}]}$, the coefficients γ_i are the solutions of the following optimization problem:

$$\min_{\gamma_i} \|g - \sum_{i=1}^{n} \gamma_i \varphi_i\|_2^2 \ , \tag{56}$$

where g is the normal density given by eq. 15, and $\| - \|_2$ is the L^2 norm. It is the known that the γ_i are the solutions of the following linear system:

$$\begin{cases} \gamma_1 \|\varphi_1\|_2^2 + \gamma_2 \langle \varphi_2, \varphi_1 \rangle + \cdots + \gamma_n \langle \varphi_n, \varphi_1 \rangle & = \quad \langle g, \varphi_1 \rangle \ , \\ \vdots & \\ \gamma_1 \langle \varphi_1, \varphi_n \rangle + \gamma_2 \langle \varphi_2, \varphi_n \rangle + \cdots + \gamma_n \|\varphi_n\|_2^2 & = \quad \langle g, \varphi_n \rangle \ . \end{cases} \tag{57}$$

The norms $\|\varphi_i\|_2^2$, as well as the scalar products $\langle \varphi_i, \varphi_j \rangle$ are straightforwardly calculated , yielding:

$$\langle \varphi_i, \varphi_j \rangle = \frac{n}{6 \ \inf(i,j)} \ \frac{1}{\mathsf{den}} \ .$$

and solving the linear system:

$$\begin{aligned} &\sum_i^n \gamma_i = \langle g, \mathbf{1}_{[b^1_{\mathsf{inf}}, b^1_{\mathsf{sup}}]} \rangle \\ &\gamma_i = i(i-1) \langle g, \varphi_{i-1} - \varphi_i \rangle - i(i+1) \langle g, \varphi_i - \varphi_{i+1} \rangle \ . \end{aligned} \tag{58}$$

VIII. APPENDIX B

The aim of this appendix is to provide a "geometric" presentation of eq. 43. We have to consider the functional:

$$\begin{aligned} &f(\mathbf{e}) = \frac{\|\mathbf{e}\|^2 - \|\Lambda\|^2}{\|\mathbf{e} - \Lambda\|} \ . \\ &\text{with the following parametrization:} \\ &\mathbf{e} = (1+\varepsilon) \ \tilde{\mathcal{R}} \Lambda \ , \ \varepsilon \in \mathbb{R} \ . \end{aligned} \tag{59}$$

In eq. 59, $\tilde{\mathcal{R}}$ is an isometry. We can restrict to positive isometry, so that we can factorize $\tilde{\mathcal{R}}$ as:

$$\tilde{\mathcal{R}} = \mathcal{R}_{\varphi}^T \mathcal{R}_{\theta} \mathcal{R}_{\varphi} \ . \tag{60}$$

Then, the numerator of $f(\mathbf{e})$ is simply $(2\varepsilon + \varepsilon^2) \|\Lambda\|^2$, while for the denominator we have:

$$\begin{aligned} \|\mathbf{e} - \Lambda\|^2 \quad &= \quad \|(\tilde{\mathcal{R}} - \mathsf{Id})\Lambda\|^2 \ , \\ &+ \quad \varepsilon \langle (\tilde{\mathcal{R}} - \mathsf{Id}), \tilde{\mathcal{R}} \Lambda \rangle + \varepsilon^2 \ \|\Lambda\|^2 \ . \\ \|(\tilde{\mathcal{R}} - \mathsf{Id})\Lambda\|^2 \quad &= \quad \|\underbrace{(\mathcal{R}_{\theta/2} - \mathcal{R}_{-\theta/2})}_{\mathcal{M}_{\theta/2}} \mathcal{R}_{\varphi} \Lambda\|^2 \ , \\ &= \quad 4\lambda^2 (\sin\theta/2)^2 (\sin\varphi)^2 \ , \\ \langle (\tilde{\mathcal{R}} - \mathsf{Id}), \tilde{\mathcal{R}} \Lambda \rangle \quad &= \quad \langle \mathcal{M}_{\theta/2}(\mathcal{R}_{\varphi}\Lambda), \mathcal{R}_{\theta/2}(\mathcal{R}_{\varphi}\Lambda) \rangle \ , \\ &= \quad -4\lambda^2 (\sin\theta/2)^2 (\sin\varphi)^2 \ . \end{aligned} \tag{61}$$

REFERENCES

[1] A. ANTONIADIS, J. BERRUYER and R. CARMONA, *Rgression non-linaire et applications*. Economica, Paris, 1992.

[2] F. BAVENCOFF, J.-M. VANPEPERSTRAETE and J.-P. LE CADRE, Performance Analysis of Optimal Data Association within a Linear Regression Framework. 8-th Conf. on Information Fusion, July 2005, pp. 683–690.

[3] Y. BAR-SHALOM and D. LERRO, Tracking with Debiased Consistent Converted Measurements versus EKF. IEEE *Trans. on* AES, vol. 29, no. 3, July 1993 1994, pp. 1015–1022.

[4] K.-C. CHANG, S. MORI and C.Y. CHONG, Performance Evaluation of Track Initiation in Dense Target Environment. IEEE *Trans. on Aut. Control*, vol. ac 29, no. 7, July 1984, pp. 585–594.

[5] K.-C. CHANG and Y. BAR-SHALOM, Joint Probabilistic Data Association for Multitarget Tracking with Possibly Unresolved Measurements and Maneuvers. IEEE *Trans. on* AES, vol. 30, no. 1, Jan. 1994, pp. 213–219.

[6] S. MORI, K.-C. CHANG and C.Y. CHONG, Performance Analysis of Optimal Data Association with Applications to Multiple Target Tracking. Y. Bar-Shalom (edt), *Multitarget Multisensor Tracking Advance Applications II*. Artech House, Norwood, MA.