[3] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Norwood, MA: Artech House, 2007.

[4] S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 5–18, 2004.

[5] D. Mušicki, R. Evans, and S. Stankovic, "Integrated probabilistic data association," *IEEE Trans. Autom. Control*, vol. 39, no. 6, pp. 1237–1241, 1994.

[6] D. Mušicki and R. Evans, "Joint integrated probabilistic data association: JIPDA," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 3, pp. 1093–1099, 2004.

[7] D. Mušicki, B. L. Scala, and R. J. Evens, "Multi-target tracking in clutter without measurement assignment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 877–896, 2008.

[8] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1152–1178, 2003.

[9] B. N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, 2005.

[10] B. N. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, 2006.

[11] B. N. Vo, B. T. Vo, N. T. Pham, and D. Suter, "Joint detection and estimation of multiple objects from image observations," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5129–5141, 2010.

[12] D. Lerro and Y. Bar-Shalom, "Automated tracking with target amplitude information," in *Proc. Amer. Control Conf.*, 1990, pp. 2875–2880.

[13] B. F. La Scala, "Viterbi data association tracking using amplitude information," presented at the 7th Int. Conf. Inf. Fusion, Stockholm, Sweden, Jun. 28–Jul. 1, 2004.

[14] D. Clark, B. Ristic, B. N. Vo, and B. T. Vo, "Bayesian multi-object filtering with amplitude feature likelihood for unknown object SNR," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 26–37, 2010.

[15] S. W. Yeom, T. Kirubarajan, and Y. Bar-Shalom, "Track segment association fine-step IMM and initialization with Doppler for improved track performance," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 1, pp. 293–309, 2004.

[16] X. Wang, D. Mušicki, R. Ellem, and F. Fletcher, "Efficient and enhanced multi-target tracking with Doppler measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 4, pp. 1400–1417, 2009.

[17] D. Mušicki, "Doppler-aided target tracking in heavy clutter," presented at the 13th Int. Conf. Inf. Fusion (FUSION), Edinburgh, U.K., Jul. 26–29, 2010.

[18] Y. Bar-Shalom and G. A. Watson, "Automatic track formation in clutter with a recursive algorithm," in *Proc. 28th Conf. Decision, Control*, Dec. 1989, pp. 1402–1408.

[19] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part 1: Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1363, 2003.

[20] M. Longbin, S. Xiaoquan, Z. Yizu, Z. S. Kang, and Y. Bar-Shalom, "Unbiased converted measurements for tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 3, pp. 1023–1027, 1998.

[21] D. Schuhmacher, B. T. Vo, and B. N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, 2008.

[22] B. Han, Y. Zhu, D. Comaniciu, and L. S. Davis, "Visual tracking by continuous density propagation in sequential Bayesian filtering framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 5, pp. 919–930, 2009.

[23] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1523–1532, 2007.

[24] B. Vo, B. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3533–3567, 2007.

A Metric for Performance Evaluation of Multi-Target Tracking Algorithms

Branko Ristic, Ba-Ngu Vo, Daniel Clark, and Ba-Tuong Vo

Abstract—**Performance evaluation of multi-target tracking algorithms is of great practical importance in the design, parameter optimization and comparison of tracking systems. The goal of performance evaluation is to measure the distance between two sets of tracks: the ground truth tracks and the set of estimated tracks. This paper proposes a mathematically rigorous metric for this purpose. The basis of the proposed distance measure is the recently formulated consistent metric for performance evaluation of multi-target *filters*, referred to as the OSPA metric. Multi-target filters sequentially estimate the number of targets and their position in the state space. The OSPA metric is therefore defined on the space of finite sets of vectors. The distinction between *filtering* and *tracking* is that tracking algorithms output *tracks* and a track represents a labeled temporal sequence of state estimates, associated with the same target. The metric proposed in this paper is therefore defined on the space of finite sets of tracks and incorporates the labeling error. Numerical examples demonstrate that the proposed metric behaves in a manner consistent with our expectations.**

Index Terms—**Estimation, performance evaluation, tracking.**

I. INTRODUCTION

Multi-target tracking refers to the sequential estimation of the number of targets and their states (positions, velocities, etc.) tagged by a unique label. Hence, the output of a tracking algorithm are *tracks*, where a track represents a labeled temporal sequence of state estimates, associated with the same target.

In evaluating the performance of a multi-target tracking algorithm the goal is to measure the distance between two sets of tracks: the set of ground truth tracks and the set of estimated tracks, produced by the tracking algorithm under evaluation. Performance evaluation of multi-target tracking algorithms is of great practical importance, with applications in tracking system design, parameter tuning and tracker comparisons (e.g., in tender evaluations). Consequently, the topic has been studied extensively; see, for example, [1], [2], [3, Ch. 13], [4], and [5].

The general performance evaluation methodology includes three steps: 1) the creation of a scenario of a certain level of difficulty; 2) running of the multi-target tracking algorithm under assessment; and 3) assignment of a performance measure (score) as a measure of the distance between the ground truth and the tracker output. In order to estimate the expected performance, the score is typically averaged

over independent Monte Carlo runs. The randomness in each run is caused by the measurements that feed the tracker (a consequence of measurement noise and false or missed detections).

In order to carry out the third step above (assessment), the standard approach to tracker performance evaluation is to first assign the tracker output to the ground truth [1]. Once the assignment is made, a list of operator defined measures of effectiveness (MoEs) are computed. These MoEs describe various aspects of tracking performance, such as [3, Ch. 13]: *timeliness* (e.g., track initiation delay, track overshoot), *track accuracy* (e.g., positional, heading, velocity error), *continuity* (e.g., track fragmentation, track labeling swaps), and *false tracks* (their count and duration), to name a few.

The problem a practitioner faces is twofold. The first is how to choose the relevant MoEs and the second is how to combine them into a single score (which is typically required for parameter tuning or in tender evaluations). The choice of relevant MoEs is far from clear, since the traditional MoEs listed above are fairly arbitrary with various authors arguing in favour of different ones. Combination of MoEs is even more questionable from the theoretical point of view, since the adopted MoEs can be correlated. To illustrate this point, note that a reduction in the track initiation delay typically increases the number of false tracks (conversely, the reduction of the number of false tracks leads to an increase in the track initiation delay). Another important issue is the "transitive" property of the tracker score resulting from the combination of MoEs. This property can be explained as follows. Consider two trackers, A and B, and suppose that tracker A output is "close" (as indicated by the combined MoE) to the ground truth. If tracker B output is "close" to that of tracker A, can we then make the conclusion that tracker B output is also "close" to the ground truth? The answer is positive only if the combined MoE possesses the "transitive" property. It is far from clear if any combination of traditional MoEs would satisfy this property.

In order to overcome the aforementioned shortcomings of traditional MoEs, it is necessary to define a mathematically rigorous metric for measuring the distance between two sets of tracks. First of all, such a metric would naturally satisfy the transitive property discussed above, by the virtue of the *triangle inequality* axiom. Furthermore, a metric also has the necessary property that the estimated state is the same as the true state, if and only if the distance between them is zero. These properties are not guaranteed by other MoEs. The development of a metric is also important if we wish to determine the underlying theoretical performance of an algorithm, such as its asymptotic convergence and consistency properties, i.e., whether in the long run the algorithm does converge to the true state. Without a metric, these notions have no meaning.

There are a few known metrics for measuring the distance between any two sets. The starting point is the Hausdorff metric, whose performance was found to be unsatisfactory due to its insensitivity to the difference in the number of elements in the set (cardinality). Hoffman and Mahler [6] then proposed a new metric for sets, based on Wasserstein distance [7], which partly fixed the undesirable cardinality behavior of the Hausdorff metric. Schuhmacher *et al.* [8], however, subsequently demonstrated a number of shortcomings of the Hoffman–Mahler metric and proposed a new consistent metric for sets, referred to as the *optimal subpattern assignment* (OSPA) metric. The OSPA metric is shown to eliminate all of the shortcomings of the Hoffman–Mahler metric. Schuhmacher *et al.* [8] demonstrated the OSPA metric in the context of performance evaluation of multi-target *filtering algorithms*. These algorithms sequentially estimate the number of currently existing targets and their position in the state space. The OSPA metric, being a metric between two sets of points in

the state space is an appropriate metric for this application: it incorporates both the cardinality error and the spacial distance of points.

For target tracking, however, we require a metric on the space of finite sets of tracks, where a track has been defined as a labeled temporal sequence. The tracks are typically of unequal length in time. The paper proposes an adaptation of the standard OSPA metric for this purpose. The resulting metric measures the distance on the joint state-space—target-label domain. In addition, for the purpose of assigning the tracker output to the ground truth, the global OSPA distance between the pairs of tracks is minimized. The resulting metric is mathematically rigorous and consistent. Preliminary formulation of the OSPA metric for tracks have been announced in [9].

The paper is organized as follows. Section II formally introduces the problem and proposes the conceptual solution based on the OPSA metric. Section III discusses the base distance to be used in OSPA metric for tracks. Numerical examples are presented in Section IV, while the contribution is summarized in Section V.

II. PROBLEM AND ITS CONCEPTUAL SOLUTION

A. Problem Specification

Our goal is to define a metric on the space of finite sets of objects called *tracks*, defined over discrete-time support points $\tau = (t_1, t_2, \ldots, t_K)$. A track X on τ is a labeled sequence of length K:

$$X = (X_1, X_2, \ldots, X_K) \tag{1}$$

where X_k, $k = 1, \ldots, K$, is either an empty set or a singleton whose element is $(\ell, \mathbf{x}_k)$. Here $\ell \in \mathbb{N}$ is the track label, which does not change with time, while $\mathbf{x}_k$ is the time evolving state vector, a point in an N-dimensional *state space* $\mathbf{W} \colon \mathbf{x}_k \in \mathbf{W} \subseteq \mathbb{R}^N$. The state space typically involves target position and velocity in the Cartesian coordinates, but may also include other target features, such as amplitude, size, shape, or similar. For convenience let us introduce an indicator e_k, which takes the value of one, if the track exists at time t_k and zero otherwise. Then we can write

$$X_k = \begin{cases} \emptyset, & \text{if } e_k = 0 \\ \{(\ell, \mathbf{x}_k)\}, & \text{if } e_k = 1. \end{cases} \tag{2}$$

Our aim is to define a metric at one particular time instant t_k, $k = 1, \ldots, K$. Let the set of all tracks on τ at specific time t_k be denoted $\mathbb{X}_k$ (i.e., $X_k \in \mathbb{X}_k$). Furthermore, let the set of finite subsets of $\mathbb{X}_k$ be denoted by $\mathcal{X}_k$. We need to define a metric space $(\mathcal{X}_k, D)$, where function $D \colon \mathcal{X}_k \times \mathcal{X}_k \to \mathbb{R}_+ = [0, \infty)$ is a metric that satisfies the following three axioms for all $\mathfrak{X}_k, \mathfrak{Y}_k, \mathfrak{Z}_k \in \mathcal{X}_k$ [10]:

1) *identity:* $D(\mathfrak{X}_k, \mathfrak{Y}_k) = 0$ if and only if $\mathfrak{X}_k = \mathfrak{Y}_k$;
2) *symmetry:* $D(\mathfrak{X}_k, \mathfrak{Y}_k) = D(\mathfrak{Y}_k, \mathfrak{X}_k)$;
3) *triangle inequality:* $D(\mathfrak{X}_k, \mathfrak{Y}_k) \leq D(\mathfrak{X}_k, \mathfrak{Z}_k) + D(\mathfrak{Z}_k, \mathfrak{Y}_k)$.

If, as a matter of convention, we denote by $\mathfrak{X}_k \in \mathcal{X}_k$ the set of tracks representing the ground truth at t_k and by $\mathfrak{Y}_k \in \mathcal{X}_k$ the set of estimated tracks produced at t_k by the algorithm under assessment, metric $D(\mathfrak{X}, \mathfrak{Y}_k)$ should quantify the overall estimation error of the tracking algorithm at time t_k. In a mathematically rigorous manner, it should combine various aspects of tracking performance (e.g., timeliness, track accuracy, continuity, data association, false tracks, etc.) into a single metric.

B. Conceptual Solution

The conceptual solution for proposed the multi-target tracking metric is based on the OSPA metric. The OSPA metric on $\mathcal{X}_k$ [8] is a distance between any two sets of objects. In our case the objects are tracks at t_k and the two sets are

$$\mathfrak{X}_k = \{(\ell_1, \mathbf{x}_{k,1}), \ldots, (\ell_m, \mathbf{x}_{k,m})\} \tag{3}$$

$$\mathfrak{Y}_k = \{(s_1, \mathbf{y}_{k,1}), \ldots, (s_n, \mathbf{y}_{k,n})\}. \tag{4}$$

According to the convention, $\mathfrak{X}_k \in \mathcal{X}_k$ and $\mathfrak{Y}_k \in \mathcal{X}_k$ are the existing ground truth tracks and tracker produced tracks at t_k, respectively. The cardinalities of sets $\mathfrak{X}_k$ and $\mathfrak{Y}_k$, m and n respectively, depend on k due to (2). For $m \leq n$, the OSPA distance between $\mathfrak{X}_k$ and $\mathfrak{Y}_k$ is defined as

$$
D_{p,c}(\mathfrak{X}_k, \mathfrak{Y}_k)
$$
$$
= \left[\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} \left(d_c\left(\tilde{\mathbf{x}}_{k,i}, \tilde{\mathbf{y}}_{k,\pi(i)}\right) \right)^p + (n-m) \cdot c^p \right) \right]^{\frac{1}{p}} \tag{5}
$$

where $\tilde{\mathbf{x}}_{k,i} \equiv (\ell_i, \mathbf{x}_{k,i})$, $\tilde{\mathbf{y}}_{k,\pi(i)} \equiv (s_{\pi(i)}, \mathbf{y}_{k,\pi(i)})$ and
- $d_c(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \min(c, d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}))$ is the *cutoff distance* between two tracks at t_k, with $c > 0$ being the cutoff parameter;
- $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is the *base distance* between two tracks at t_k;
- Π_n represents the set of permutations of length m with elements taken from $\{1, 2, \ldots, n\}$;
- $1 \leq p < \infty$ is the OSPA metric order parameter.

For the case $m > n$, the definition is $D_{p,c}(\mathfrak{X}, \mathfrak{Y}) = D_{p,c}(\mathfrak{Y}_k, \mathfrak{X}_k)$. If both $\mathfrak{X}_k$ and $\mathfrak{Y}_k$ are empty sets (i.e., $m = n = 0$), the distance is zero. The proof that OSPA distance is indeed a metric, including a number of arguments that it complies with intuition, is given in [8].

The choice of parameters c and p follows the guidelines of the original OSPA metric [8]. The cutoff parameter c determines the relative weighting given to the cardinality error component against the base distance error component. Larger values of c tend to emphasize cardinality errors and vice versa. Note that the cardinality error is a consequence of an unequal number of tracks in $\mathfrak{X}$ and $\mathfrak{Y}$, meaning that c can be interpreted as a measure of penalty assigned to missed or false tracks. The order parameter, p, controls the penalty assigned to "outlier" estimates (that are not close to any of the ground truth tracks). A higher value of p increases sensitivity to outliers.

The OSPA metric is increasingly being used as a performance measure in estimation and has already made an impact in the development of new filters and estimators, such as the Set JPDA filter [11] and the minimum mean OSPA estimator [12].

In order to apply the OSPA metric for tracks described above, we need to define the base distance $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ between tracks $\tilde{\mathbf{x}} \equiv (\ell, \mathbf{x})$ and $\tilde{\mathbf{y}} \equiv (s, \mathbf{y})$. The base distance must be a proper metric (i.e., it needs to satisfy the three axioms) and will be discussed next.

III. THE BASE DISTANCE AND LABELING OF ESTIMATED TRACKS

A. The Base Distance Between Two Labeled Vectors

The base distance $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ is a metric the on space $\mathbb{N} \times \mathbb{R}^N$ and is defined as

$$d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \left(d(\mathbf{x}, \mathbf{y})^{p'} + d(\ell, s)^{p'} \right)^{\frac{1}{p'}} \tag{6}$$
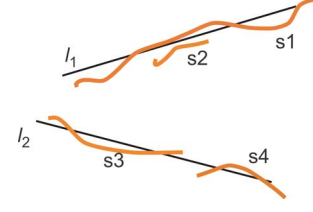

Fig. 1. Example with two true tracks labeled as ℓ_1 and ℓ_2 (thin black lines) and four estimated tracks, labeled as $s_1, \ldots, s_4$.

where

$1 \leq p' < \infty$ is the base distance order parameter;

$d(\mathbf{x}, \mathbf{y})$ the localization base distance, a metric on $\mathbb{R}^N$, typically adopted as the p'-norm: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_{p'}$;

$d(\ell, s)$ the labeling error, a metric on $\mathbb{N}$, adopted as follows:

$$d(s, t) = \alpha\, \bar{\delta}[s, t] \tag{7}$$

where $\bar{\delta}[i, j]$ is the complement of the Kronecker delta, that is $\bar{\delta}[i, j] = 0$ if $i = j$ and $\bar{\delta}[i, j] = 1$ otherwise.

Parameter $\alpha \in [0, c]$ in (7) controls the penalty assigned to the labeling error $d(s, t)$ interpreted relative to the localization distance $d(\mathbf{x}, \mathbf{y})$. The case $\alpha = 0$ assigns no penalty and $\alpha = c$ assigns the maximum penalty.

It is straightforward to verify that the base distance $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ defined by (6) satisfies the axioms of a metric. To prove identity and symmetry is trivial. To prove the triangular inequality, let $\tilde{\mathbf{x}} = (\ell, \mathbf{x})$, $\tilde{\mathbf{y}} = (s, \mathbf{y})$, $\tilde{\mathbf{z}} = (u, \mathbf{z})$ be any three labeled vectors in $\mathbb{N} \times \mathbb{R}^N$. Then we have

$$
\begin{aligned}
d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})^{p'} &= \|\mathbf{x} - \mathbf{y}\|_{p'}^{p'} + \alpha^{p'}\, \bar{\delta}[\ell, s] \\
&\leq \|\mathbf{x} - \mathbf{z}\|_{p'}^{p'} + \|\mathbf{z} - \mathbf{y}\|_{p'}^{p'} + \alpha^{p'} \left(\bar{\delta}[\ell, u] + \bar{\delta}[u, s] \right) \\
&= d(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})^{p'} + d(\tilde{\mathbf{z}}, \tilde{\mathbf{y}})^{p'}.
\end{aligned}
$$

Having defined the base distance (6), we are still not in the position to apply the OSPA metric for tracks. What remains is to determine the labels of estimated tracks.

B. Labeling of Estimated Tracks

In order to explain why this step is necessary, consider an illustrative example shown in Fig. 1. There are two true tracks, labeled as ℓ_1 and ℓ_2, (shown in thin black lines) and four estimated tracks labeled as $s_1 \ldots s_4$ (thick orange lines). In order to apply the OSPA metric in the framework described above, it is necessary to assign the labels of true tracks to some of the estimated tracks (in this example ℓ_1 and ℓ_2) in a globally optimum manner. For the case shown in Fig. 1, it would be desirable to assign ℓ_1 to track s_1, because estimated track s_1 is the best approximation of ℓ_1. For the same reason, we should assign ℓ_2 to s_3 and also make sure that the other two estimated tracks are assigned labels which are different from both ℓ_1 and ℓ_2. Recall from the introduction that the assignment of the tracker output to the ground truth is required in all standard tracking evaluation schemes [1].

A simple way to carry out the assignment task is by using one of the existing two-dimensional assignment algorithms (Munkres, JVC,

auction, etc.) [1, p. 342]. Suppose the sets of ground truth tracks and estimated tracks are $\{X^{(1)}, \ldots, X^{(L)}\}$ and $\{Y^{(1)}, \ldots, Y^{(R)}\}$, respectively. Using the existence indicators for tracks, a track $X^{(\ell)}$ at time k, denoted $X_k^{(\ell)}$, is defined as in (2). An optimal global assignment λ^* can be determined as follows ($L \leq R$):

$$\lambda^* = \arg \min_{\lambda \in \Lambda_R} \sum_{\ell=1}^{L} \sum_{k=1}^{K} \left[e_k^{\ell} e_k^{\lambda(\ell)} \min(\Delta, \| \mathbf{x}_k^{\ell} - \mathbf{y}_k^{\lambda(\ell)} \|_2) \right.$$
$$\left. + (1 - e^{\ell})_k e_k^{\lambda(\ell)} \Delta + e_k^{\ell} (1 - e_k^{\lambda(\ell)}) \Delta \right] \quad (8)$$

where

Λ_R the set of permutations of length L with elements taken from $\{1, 2, \ldots, R\}$;

Δ the penalty assigned to instantaneous miss or false track;

$e_k^{\ell}, e_k^{\lambda(\ell)}$ existence indicators for true and λ-assigned estimated tracks.

The case $L > R$ is a trivial modification of (8).

A close inspection of (8) reveals that the term in the square brackets is in fact the OSPA distance between $X_k^{(\ell)}$ and $Y_k^{(\lambda(\ell))}$, with $p = 2$ and $c = \Delta$ (recall from (2) that these two sets are either empty or singletons, hence the form in (8)). Thus, the assignment of estimated tracks to ground-truth tracks is carried out by minimizing the global OSPA distance between the pairs of tracks, accumulated over the discrete-time points τ. Parameter Δ controls the penalty assigned when at time t_k one of the tracks exists and the other does not. A high value of Δ will favour the assignment of longer duration estimated tracks to the true tracks. If an estimated track has been assigned by λ^* to a true track with label ℓ, then its label is set to ℓ too. Estimated tracks which remain unassigned according to λ^*, are given labels different from all the true track labels.

The basic steps in the computation of the proposed OSPA based metric for tracks are given in Algorithm 1.

Algorithm 1: Computation Steps of the Proposed Metric

1: **function** OSPA-T $(\{X^{(1)}, \ldots, X^{(L)}\}, \{Y^{(1)}, \ldots, Y^{(R)}\})$
2:      % Label the estimated tracks
3:      For $j = 1, \ldots, R$, $\mathrm{Label}[Y^{(j)}] = I$ (initial value, different from all true track labels)
4:      Find λ^*, the globally optimum assignment of tracks $\{X^{(1)}, \ldots, X^{(L)}\}$ to $\{Y^{(1)}, \ldots, Y^{(R)}\}$
5:      For $i = 1, \ldots, L$, $\mathrm{Label}[Y^{(\lambda^*(i))}] = \mathrm{Label}[X^{(i)}]$
6:      % Compute the distance
7:      For $k = 1, \ldots, K$
8:          Form the labeled sets $\mathfrak{X}_k$ and $\mathfrak{Y}_k$ at t_k, see (3) and (4)
9:          Compute the OSPA distance at t_k using (5) with the base distance (6)
10: **end function**

IV. NUMERICAL EXAMPLES

The proposed OSPA based metric for tracks is demonstrated using a scenario consisting of five crossing targets, shown in Fig. 2. The targets are moving in the x–y plane, from left to right. The total observation interval is $K = 120$ time-steps long. The targets cross (effectively are on top of each other) at $k = 71$. Their duration varies: the start-time indexes are $k = 3, 5, 7, 9, 10$; the termination-time indexes are $k = 101, 103, 105, 108, 109$. The sensor is placed at $(200, 200)$ m
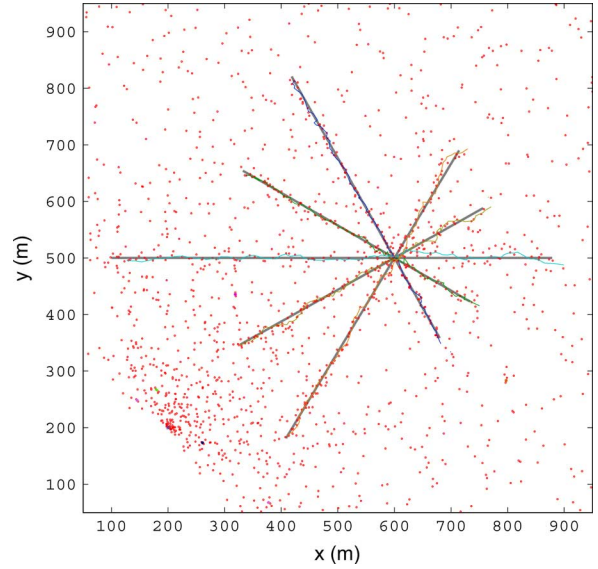

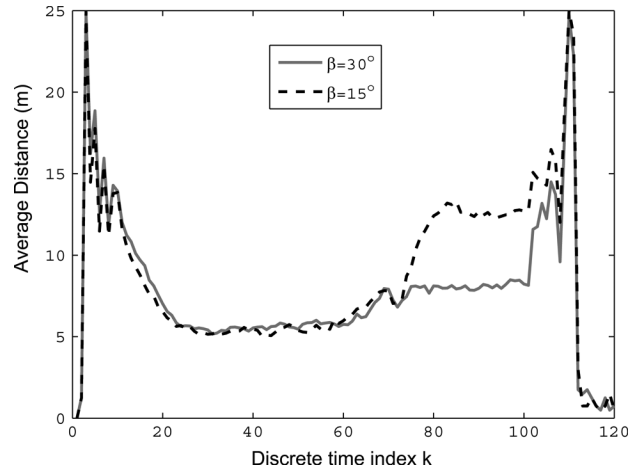Fig. 2. The testing scenario consisting of five crossing targets (crossing angle between the tracks is 30°).

Fig. 3. Average OSPA distance for tracks: different crossing angles β case.

and measures the range and bearing, with zero-mean white Gaussian noise. The measurement noise standard deviations are $\sigma_r = 3$ m and $\sigma_\theta = 1°$. The multi-target tracking algorithm used in demonstration is based on the Bernoulli/JoTT filter [13] and includes the linear multi-target (LM) data association [14] (details omitted, being irrelevant for the demonstration of the metric).

The numerical simulations have been carried out for four cases, listed below. Unless otherwise stated, the parameters of the proposed metric were set to: $\Delta = 100$ m, $\alpha = 25$ m, $c = 25$ m, and $p = p' = 1$. It was found that the metric is insensitive to the global assignment parameter Δ, as long as $\Delta > 10$ m. The choice of α will be discussed in the fourth case. The state space consisted of target positions in the x–y plane, only. In all the figures shown below, the curves of the proposed metric were obtained by averaging over 100 Monte Carlo runs.

1) Variation of the Crossing Angle: Fig. 3 shows the proposed metric for two different scenarios: the solid line corresponds to the crossing angle of $\beta = 30°$; the dashed line to the crossing angle of $\beta = 15°$. In both cases the probability of detection was $P_d = 0.9$ and the average number of clutter points (false detection) per scan
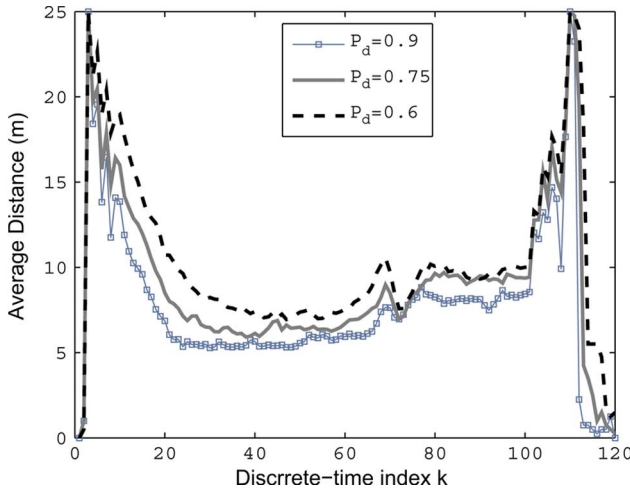
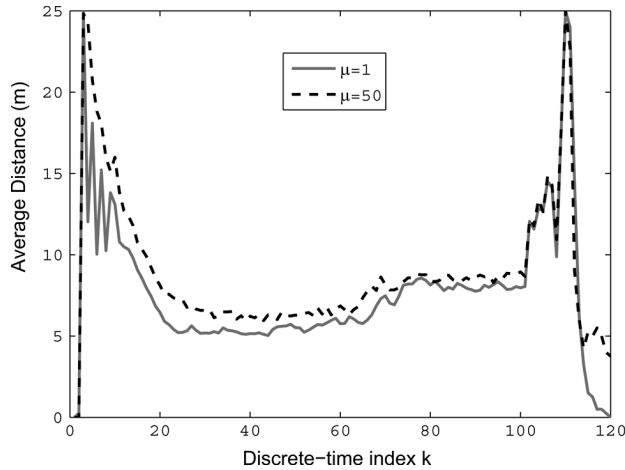Fig. 4. Average OSPA distance for tracks: different values of probability of detection P_d.

Fig. 6. Average OSPA distance for tracks: the effect of α parameter which penalizes the labeling error.

Fig. 5. Average OSPA distance for tracks: different values of the average number of false detections per scan μ.

was $\mu = 10$. The general trend in the OSPA distance for tracks in this scenario is that it has high values initially and at the end of the scenario. This is due to track initiation delay and the track termination delay. Furthermore, we observe that the OSPA metric for the two different scenarios does not change until the target crossing event (after $k = 71$). After this point of time, the distance between the ground truth and the tracker output becomes much higher for the smaller crossing angle β scenario. The result is in accordance with our intuition. There is no reason for the difference in the metric before the crossing, as the tracks are well separated. After the crossing, when some tracks are swapped or broken due to the difficulties in measurement-to-track association, the small β angle scenario results in the higher OSPA distance for tracks: this is to be expected as small β case is much more difficult for tracking.

2) Variation of the Probability of Detection: Fig. 4 shows the proposed metric for $P_d = 0.9$ (thin solid line with square markers), $P_d = 0.75$ (solid thick line) and $P_d = 0.6$ (dashed line). In all cases $\beta = 30°$ and $\mu = 10$. We observe that the higher value of P_d results in the smaller value of OSPA distance between the ground truth tracks and estimated tracks, during the entire observation interval. Again this is in
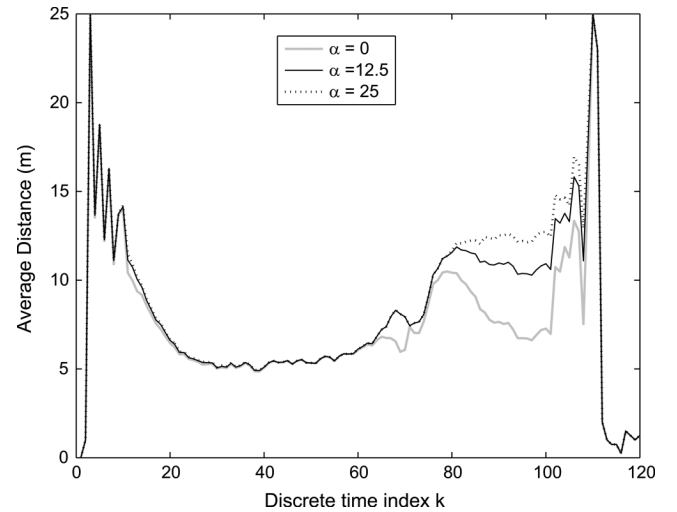
accordance with our intuition. The track initiation and track termination are delayed at lower P_d (as reflected in the results). In addition, the track estimation accuracy is worse at lower P_d throughout the observation interval.

3) Variation of the Density of Clutter: Fig. 5 compares the proposed metric at $\mu = 1$ (solid line) and $\mu = 50$ (dashed line). In both cases $P_d = 0.9$ and $\beta = 30°$. In accordance with intuition, the higher the value of μ, the greater the OSPA distance between the ground truth and estimated tracks. This is due to the increased probability of false tracks and more difficult data association at higher values of μ.

4) Variation of Parameter α: The parameters of the simulation setup were $P_d = 0.9$, $\beta = 15°$ and $\mu = 10$. Fig. 6 shows the proposed metric using three different values of the metric parameter α: $\alpha = 0$, $\alpha = \frac{c}{2} = 12.5$ m and $\alpha = c = 12$ m. Note that for $\alpha = 0$, the proposed metric for tracks is equivalent to the standard OSPA metric [8]. In the considered tracking scenario, the labeling error is expected to appear after the crossing of targets, when track labels are likely to swap. Fig. 6 demonstrates this effect and the importance of including the labeling error in the metric.

V. Summary

The paper introduced a mathematically rigorous metric for measuring the distance between two sets of tracks. Typically one set of tracks is the ground truth, while the other is produced by a tracker. The proposed metric is developed for the purpose of performance evaluation of tracking algorithms, with important implications to tracker design, parameter optimization and comparison.

The proposed metric captures all aspects of the difference between the two sets of tracks: the cardinality error, the localization error and the labeling error. The inclusion of the labeling error is the novelty in relation to the standard OSPA metric [8]. The relative weighting of the three contributing errors to the proposed metric for tracks can be controlled by the appropriate selection of the metric parameters. Numerical analysis demonstrates that the OSPA metric for tracks behaves in agreement with our intuition, in the manner consistent with our expectations. The Matlab source code for the computation of OSPA metric for tracks is available at http://randomsets.eps.hw.ac.uk/index. html.

Future work will investigate alternative choices for the localization base distance introduced in (6), based on probability density functions.

REFERENCES

[1] B. E. Fridling and O. E. Drummond, "Performance evaluation methods for multiple target tracking algorithms," in *Proc. SPIE, Signal Data Process. Small Targets*, 1991, vol. 1481, pp. 371–383.

[2] B. Ristic, "A tool for track-while-scan algorithm evaluation," in *Proc. Inf., Decision, Control (IDC)*, Adelaide, Australia, Feb. 1999, pp. 105–110.

[3] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.

[4] S. B. Colegrove, L. M. Davis, and S. J. Davey, "Performance assessment of tracking systems," in *Proc. Int. Symp. Signal Process. Appl. (ISSPA)*, Gold Coast, Australia, Aug. 1996, pp. 188–191.

[5] R. L. Rothrock and O. E. Drummond, "Performance metrics for multi-sensor, multi-target tracking," in *Proc. SPIE, Signal Data Process. Small Targets*, 2000, vol. 4048, pp. 521–531.

[6] J. R. Hoffman and R. P. S. Mahler, "Multitarget miss distance via optimal assignment," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 34, no. 3, pp. 327–336, May 2004.

[7] L. Rueshendorff, "Wasserstein (Vasershtein) metric," in *Encyclopaedia of Mathematics*, M. Hazewinkel, Ed. New York: Springer, 2001.

[8] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[9] B. Ristic, B.-N. Vo, and D. Clark, "Performance evaluation of multi-target tracking using the OSPA metric," presented at the 13th Int. Conf. Inf. Fusion, Edinburgh, U.K., Jul. 2010.

[10] E. Kreyszig, *Introductory Functional Analysis with Applications*. New York: Wiley, 1989.

[11] L. Svensson, D. Svensson, and P. Willett, "Set JPDA algorithm for tracking unordered sets of targets," presented at the 12th Int. Conf. Inf. Fusion, Seattle, WA, Jul. 2009.

[12] M. Guerriero, L. Svensson, D. Svensson, and P. Willett, "Shooting two birds with two bullets: How to find minimum mean OSPA estimates," presented at the 13th Int. Conf. Inf. Fusion, Edinburgh, U.K., Jul. 2010.

[13] R. Mahler, *Statistical Multisource Multitarget Information Fusion*. Norwood, MA: Artech House, 2007.

[14] D. Musicki and B. L. Scala, "Multi-target tracking in clutter without measurement assignment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 877–896, Jul. 2008.

A Fast Algorithm for Nonunitary Joint Diagonalization and Its Application to Blind Source Separation

Xian-Feng Xu, Da-Zheng Feng, *Member, IEEE*, and Wei Xing Zheng, *Senior Member, IEEE*

Abstract—A fast algorithm, named Complex-Valued Fast Frobenius DIAGonalization (CVFFDIAG), is proposed for seeking the nonunitary approximate joint diagonalizer of a given set of complex-valued target matrices. It adopts a multiplicative update to minimize the Frobenius-norm formulation of the approximate joint diagonalization problem. At each of multiplicative iterations, a strictly diagonally dominant updated matrix is obtained. This scheme ensures the invertibility of the diagonalizer. The CVFFDIAG relaxes several constraints on the target matrices and thus has much general applications. Furthermore, the special approximation of the cost function, the ingenious utilization of some structures and the adequate notation of concerned variables lead to the high computational efficiency of the proposed algorithm. Numerical simulations are conducted to illustrate good performances of the CVFFDIAG.

Index Terms—Blind source separation (BSS), CVFFDIAG (Complex-Valued Fast Frobenius DIA gonalization), Frobenius-norm formulation, multiplicative update, nonunitary approximate joint diagonalization.

I. INTRODUCTION

The problem of approximate joint diagonalization (AJD) of a set of target matrices is frequently encountered and widely used in the field of multivariate statistical signal processing (see, e.g., [1]–[6]). For AJD, it is desired to fit a set of $N \times N$ square matrices $\mathscr{C} = \{\mathbf{C}^k, \ k = 1, \ldots, K\}$ defined by common structured matrices of the form $\mathbf{C}^k = \mathbf{A}\mathbf{\Lambda}^k\mathbf{A}^H$ or $\mathbf{V}\mathbf{C}^k\mathbf{V}^H = \mathbf{\Lambda}^k$. Here $\mathbf{\Lambda}^k = \mathrm{diag}[\lambda_1^k, \ldots, \lambda_N^k]$ $(k = 1, \ldots, K)$ denote unknown diagonal matrices and $\mathbf{A}$ denotes an unknown $N \times N$ matrix with $\mathbf{V}$ being its inverse. In the context of blind source separation (BSS), which is one of the most prominent applications of AJD, $\mathbf{A}$ is often termed as the mixing matrix while $\mathbf{V}$ is often termed as the demixing matrix. We will adopt these terms in this paper.

Various AJD algorithms have already been proposed till now (see, e.g., [1]–[8]). They mainly differ in how the notion of diagonality is defined, what concrete methods are developed for solving the corresponding cost function and what diverse restrictions are imposed on $\mathscr{C}$ and $\mathbf{V}$ (or $\mathbf{A}$). The notion of diagonality was explicitly introduced in [5] and [7] and the concrete methods for solving the corresponding cost function may be referred to those above-mentioned papers on AJD. Here, we pay much attention to the diverse constraints which are imposed by those AJD algorithms on $\mathscr{C}$ and $\mathbf{V}$ (or $\mathbf{A}$). Some of the earlier AJD techniques, like SOBI [1] or JADE [6] for instance, require the