# A Consistent Metric for Performance Evaluation of Multi-Object Filters

Dominic Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo

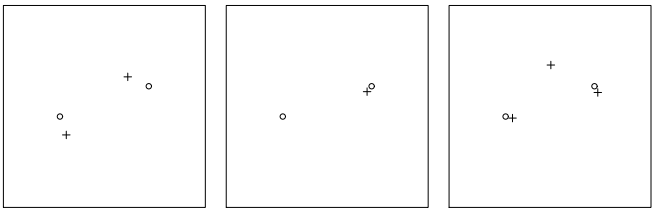
***Abstract*— The concept of a miss-distance, or error, between a reference quantity and its estimated/controlled value, plays a fundamental role in any filtering/control problem. Yet there is no satisfactory notion of a miss-distance in the well-established field of multi-object filtering. In this paper, we outline the inconsistencies of existing metrics in the context of multi-object miss-distances for performance evaluation. We then propose a new mathematically and intuitively consistent metric that addresses the drawbacks of current multi-object performance evaluation metrics.**

***Index Terms*— Miss-distance, Performance evaluation, Wasserstein distance, Multi-object systems, Multi-object filtering Multi-target tracking, Random sets, Point processes**

## I. Introduction

Dating back to the early 1970s, multi-object systems research was driven primarily by aerospace applications such as radar, sonar, guidance, navigation, and air traffic control (see [1] and references therein). Today, multi-object systems research is an established discipline [2]–[4] with a host of diverse application areas including computer vision [5], oceanography [6], [7], autonomous vehicle/robotics [8], remote sensing [9], and biomedical research [10]. A multi-object system is fundamentally different from a single-object system in that the system state, called a *multi-object state*, is a finite set of vectors rather than a single vector. Thus, not only the constituent vectors of the multi-object state evolve in time, but the number of these vectors also changes with time due to the appearance and disappearance of objects [4]. Filtering involves jointly estimating the number of constituent vectors and their values from the observation history, whereas control involves determining the control signals to achieve a certain objective, which in many applications is a function of the filtering performance.

The concept of a *miss-distance*, or error, between a reference quantity and its estimated/controlled value, plays a fundamental role in any filtering/control problem. For example, in single-object systems, widely used concepts of optimality, such as least-squares, expected value, and root-mean-square error, would be incomprehensible without the concept of a miss-distance. The miss-distance is indispensable in the formulation and evaluation of filtering and control algorithms.



(a) (b) (c)

Fig. 1. Hypothetical scenarios with ∘ representing true targets and + representing estimated targets. Which estimate is closest to the truth?

While the concept of a miss-distance is taken for granted in single-object systems, this has not been the case in the multi-object realm. Consider a tracking application where the number of targets is not known and is to be inferred along with the positions of the targets. Suppose that the true number of targets is two and consider three different estimates as shown in Fig. 1: (a) 2 points that are close to the true positions of the targets; (b) 1 point that almost coincides with the true position of one target; and (c) 3 points 2 of which almost coincide with the true positions of the targets. The question is which of these estimates is "closest" to the truth? A satisfactory multi-object miss-distance needs to capture the "difference" between two sets of vectors, namely the reference multi-object state and the estimated multi-object state, in a mathematically consistent yet physically meaningful manner. Thus, taking practical considerations into account, a multi-object miss-distance should:

- be a metric on the space of finite sets,
- have a natural (meaningful) physical interpretation,
- capture cardinality errors and state errors meaningfully,
- be easily computed.

At present, a satisfactory notion of a multi-object miss-distance is not available despite the abundance of multi-object filtering techniques and applications [2]–[4]. Works such as [11] discuss performance evaluation from a more applied viewpoint. The optimal assignment paradigm of Drummond and associates [12], [13] has an intuitively appealing physical interpretation, but it can only measure the distance between multi-object states of the same cardinality in a consistent manner. The Hausdorff distance rigorously accommodates multi-object states of different cardinalities. However, it is relatively insensitive to differences in cardinalities [14]. The first

This work is partly supported by the ARC Discovery Grant No. DP0878158

D. Schuhmacher is with the School of Mathematics and Statistics, The University of Western Australia, Crawley, WA 6009, Australia (email: dominic@maths.uwa.edu.au).

B.-T. Vo is with the Western Australian Telecommunications Research Institute, The University of Western Australia, Crawley, WA 6009, Australia (email: vob@watri.uwa.edu.au).

B.-N. Vo is with the Department of Electrical Engineering, The University of Melbourne, Melbourne, Vic. 3010, Australia (email: bv@ee.unimelb.edu.au).

rigorous theory of multi-object miss-distance was proposed by Hoffman and Mahler, based on a Wasserstein construction [14]. This miss-distance is more sensitive to differences in cardinalities than the Hausdorff distance. Moreover, it extends the optimal assignment paradigm of Drummond and hence inherits the physically intuitive interpretation when the two multi-object states have the same cardinality. However, it does not have a physically consistent interpretation when the multi-object states have different cardinalities, and suffers from a number of other serious limitations.

In this article we outline the drawbacks of the miss-distance proposed by Hoffman and Mahler in the context of multi-object filtering performance evaluation, and propose a mathematically consistent multi-object miss-distance that captures what we intuitively want to measure. Our approach is also based on a Wasserstein construction, and inherits the optimal assignment interpretation of the miss-distances proposed in [12], [14] when the multi-object states have the same cardinality. More importantly, our miss-distance can accommodate cardinality differences in a mathematically consistent and physically meaningful manner that eliminates any element of arbitrariness suffered by ad hoc extensions of Drummond's approach (as discussed in Subsection I.B of [14]). It also addresses many of the other drawbacks of previous approaches and can be easily computed. The miss distance we propose has a parameter which controls the relative emphasis of localization and cardinality errors. This feature provides users with enough flexibility to accommodate the goals of particular missions.

The paper is organized as follows. In Section II we give an overview of previous metrics used for multi-object systems and identify their strengths and weaknesses. Section III presents our proposed miss-distance, demonstrates how it solves the problems associated with the earlier metrics, and gives details about its computation. Section IV presents numerical studies that verify the salient features of the new miss-distance. Concluding remarks are given in Section V and mathematical details are given in the appendix.

## II. Previous metrics for multi-object systems

We begin with a more detailed analysis of the multi-object miss-distances mentioned in the introduction. From a mathematical point of view, the fundamental requirement to allow for consistent distance measuring is that our miss-distances are metrics on the space of finite sets of objects [14]. For completeness, we recall the definition of a metric. Let $\mathcal{X}$ be an arbitrary non-empty set. A function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+ = [0, \infty)$ is called a *metric* if it satisfies the following three axioms:

1) *(identity)* $d(x, y) = 0$ if and only if $x = y$;
2) *(symmetry)* $d(x, y) = d(y, x)$ for all $x, y \in \mathcal{X}$;
3) *(triangle inequality)* $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in \mathcal{X}$.

In the context of multi-object miss-distances, we fix a closed and bounded observation window $W \subset \mathbb{R}^N$, and choose $\mathcal{X}$ to be the set of finite subsets of $W$. In what follows, $d$ denotes always the metric used on $W$ (typically the Euclidean metric, $d(x, y) = \|x - y\|$) while for the various metrics considered on $\mathcal{X}$ appropriate indices are attached ($d_H$, $d_p$ or $\bar{d}_p^{(c)}$).

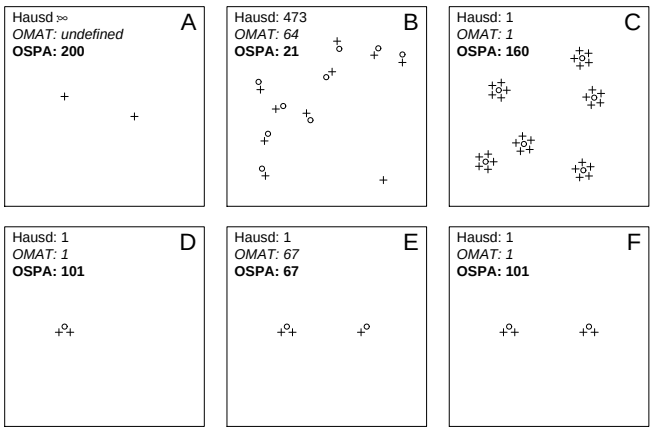



Fig. 2. Six scenarios in a $1000 \times 1000m^2$ window to illustrate the strengths and weaknesses of the various metrics. ∘ actual objects, + estimates. Parameters are $p = 1$ (OMAT and OSPA) and $c = 200$ (OSPA). Scenarios depict A) two false estimates; B) an outlier false estimate among several accurate estimates; C) multiple estimates per object; D–F) a comparison of balanced and unbalanced allocations of estimates to objects. Note that Scenarios B to F are an artistic impression rather than an exact rendering of the situation in the sense that the smallest distances have been considerably inflated for better viewing (from 1 meter, which was assumed in the computations, to 40 meters in the scale of the pictures).

### A. Hausdorff metric

For finite non-empty subsets $X$ and $Y$ of $W$, define

$$d_H(X, Y) = \max\Big\{\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y)\Big\}. \quad (1)$$

The function $d_H$ is called the *Hausdorff metric*. A proof that $d_H$ is a metric is given in Subsection VII-A of [14], and Section III of the same paper discusses some of its merits and difficulties in the context of multi-object filtering. We emphasize here only the most important points.

First of all, it should be noted that, from a practical point of view, the Hausdorff metric is traditionally used as a measure of dissimilarity between binary images, for which it is well suited both for theoretical and for intuitive reasons (it typically gives a good idea of the difference in the optical impressions a human would get from two images). However, it is very insensitive to differing cardinalities of finite sets, as can be seen from Scenarios C to F in Fig. 2, which is not desirable for a performance measure of multi-object filters. Furthermore, it penalizes outliers heavily (see Scenario B of Fig. 2), and cannot be reasonably defined if one of the sets is empty (Scenario A of Fig. 2) though some authors set it to $\infty$ in this case.

From a mathematical point of view, the Hausdorff metric has a long tradition in stochastic geometry for its theoretical merits. It generates the standard topology considered on the set of closed subsets of $W$ (see [15], p.3 or equivalently p.12), which is used to define random sets. In the context of finite set statistics (FISST) this topology is usually called the Matheron topology [4].

### B. Optimal Mass Transfer (OMAT) metric

In 2004, Hoffman and Mahler [14] introduced a new metric to overcome some of the problems of the Hausdorff metric in the context of multi-object filtering performance evaluation.

For $1 \leq p < \infty$ and finite non-empty subsets $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$ of $W$, define

$$\begin{aligned} d_p(X,Y) &:= \min_C \left( \sum_{i=1}^{m} \sum_{j=1}^{n} C_{i,j} d(x_i, y_j)^p \right)^{1/p}, \\ d_\infty(X,Y) &:= \min_C \max_{1 \leq i \leq m, 1 \leq j \leq n} \tilde{C}_{i,j} d(x_i, y_j), \end{aligned} \tag{2}$$

where the minima are taken over all $m \times n$ transportation matrices $C = (C_{i,j})$ and where $\tilde{C}_{i,j} = 1$ if $C_{i,j} \neq 0$ and $\tilde{C}_{i,j} = 0$ otherwise. An $m \times n$ matrix $C$ is a transportation matrix if all of its entries are non-negative and if

$$\sum_{j=1}^{n} C_{i,j} = \frac{1}{m} \text{ for } 1 \leq i \leq m, \sum_{i=1}^{m} C_{i,j} = \frac{1}{n} \text{ for } 1 \leq j \leq n.$$

We call the function $d_p$ the *OMAT metric* of order $p$, where OMAT stands for Optimal MAss Transfer (compare the remark about the "earth mover's distance" after Equation (6) in [14]). Hoffman and Mahler used the term Wasserstein metric because the above definition yields the $p$-th order Wasserstein metric between the empirical distributions of the point patterns $X$ and $Y$. Since the metric we propose in the next section is also based on a Wasserstein construction, we refrain from using this term in relation to the OMAT metric to avoid potential confusion.

The merits of the OMAT metric are that it partly fixes the undesirable cardinality behaviour of the Hausdorff metric (see Scenario E of Fig. 2) and that it can cope with the outlier problem by the introduction of the parameter $p$ (Scenario B of Fig. 2). Regarding this second point, it should be mentioned that there are also generalizations of the Hausdorff metric that avoid this problem [16].

On the downside, the OMAT metric entails a host of problems; the most important of these are outlined as follows:

*1) Inconsistency of the metric:* Consider Scenarios D to F in Fig. 2. Each of these examples involves a cardinality error, and from an intuitive point of view we would say that the estimation errors are roughly equal, but that E is probably a bit better than F and (arguably) somewhat better than D. While the Hausdorff distance is the same in all three examples (undesirably small due to the cardinality problem, but at least consistent), the OMAT metric actually depends on how well balanced the numbers of estimate points are among the actual objects. It therefore assigns a much larger distance in Scenario E than in the other two cases, and thus ranks the scenario that we intuitively prefer (E) as by far the worst among the three.

This series of examples further reveals that the OMAT metric is not always better than the Hausdorff metric at detecting different cardinalities. This becomes even more obvious in Scenario C of Fig. 2, where the cardinality of the estimated point pattern is quite far from the truth. However, since the estimates are perfectly balanced among the ground truth objects, the OMAT metric does not detect this. While the scenario depicted is certainly an extreme one, it is in essence not unrealistic (see our comparison of the MeMBer and PHD filters in Subsection IV-B).

*2) Contrived construction for differing cardinalities:* Decomposing individual objects into small parts does not seem very attractive from an intuitive point of view and often makes the resulting distances hard to interpret. While visual perception sometimes tells us that a reasonable matching of estimated objects to ground truths would involve splitting up of unit masses (e.g. Scenario E in Fig. 2, where we would naturally match the two estimates on the left with the one ground truth object next to them), the OMAT metric does not cater for such natural structure (and it is in fact hard for any metric to do so). Instead, it tends to assign partial masses between the two sets of objects in a complicated way, which may be difficult to comprehend for a human observer. For example the Hausdorff distance in Scenario B is clearly the distance from the isolated filter estimate to its closest ground truth object, whereas the optimal mass transfer for $d_1$ or $d_2$ is not so obvious.

*3) Geometry dependent behaviour:* In [14], Hoffman and Mahler describe what they call the *geometry dependence* of the OMAT metric, stating that a multi-object filter should be more heavily penalized for misestimating cardinality when the objects are far apart than when they are closely spaced. The rationale behind this statement is that it is harder to estimate the number of objects when they are closely spaced. Under the assumption that a multi-object filter consistently misestimates the number of objects, the authors argue that the magnitude of $d_\infty$ approximately equals the diameter of the ground truth (and that $d_p$ shows a similar dependence on the diameter for smaller $p$ albeit to a lesser degree).

In contrast, we do not think that geometry dependence is a desirable property. Ideally, an objective performance metric should depend on as few specific features of the considered test case as possible in order to allow direct comparisons between different scenarios (or different stages within the same scenario). The need for reinterpretation of the OMAT distances according to the geometry of the ground truth as described in [14] means that we have to transform these distances in order to arrive at error quantities that are comparable with one another. Such quantities, however, do not in general satisfy the metric axioms any more and other desirable properties of the OMAT metric might be lost as well.

Moreover, the stated dependence of $d_\infty$ on the diameter of the ground truth is a strong simplification. Consider for example a ground truth of $n$ objects that are arranged as a regular $n$-gon and a filter estimate that is perfect except that one of the objects is missing. In this case the $d_\infty$-distance will be equal to the side length of the polygon, which for large $n$ is very different from its diameter. In general, the order of magnitude of $d_p$ depends substantially on the number of objects as well as on more specific information about the geometry of the ground truth, which makes an objective performance evaluation based on the OMAT metric all the more difficult. Thus, in the case where the miss-distance versus time curve increases with time, it is virtually impossible to tell whether a multi-object filter diverges or performs well (compare Fig. 11). Even if this curve decreases to a small constant value, it cannot be inferred that the filter performance is good.

*4) Undefined if cardinality is zero:* The distance $d_p(X,Y)$ is not defined if one of the two point patterns is empty, and in fact there is no natural extension for this case. However, when evaluating multi-object filter performance, the situation where there are no objects present but we obtain false estimates for whatever reason, is quite common and should be captured by a reasonable metric.

*5) Incompatibility with mathematical theory:* The topology generated by the OMAT metric does not conform with any well-established topology from the appropriate fields of mathematics such as stochastic geometry or point process theory. While this poses no direct problem for the practical use nor for defining random events that involve the metric (compare the technical note at the end of Section IV in [14]), it still makes certain theoretical statements difficult. Most notably, it is not possible to describe the convergence of point patterns or the convergence in distribution of random point patterns (which are both well-defined concepts in stochastic geometry and point process theory) based on the OMAT metric.

## III. Optimal Subpattern Assignment (OSPA) Metric

In this section, we introduce our new performance metric, which is still based on a Wasserstein construction, but completely eliminates most of the aforementioned problems of the OMAT metric. This metric was introduced only very recently in [17] to address certain problems in point process theory and spatial statistics.

Denote by $d^{(c)}(x,y) := \min\big(c, d(x,y)\big)$ the distance[1] between $x, y \in W$ cut off at $c > 0$, and by $\Pi_k$ the set of permutations on $\{1,2,\ldots,k\}$ for any $k \in \mathbb{N} = \{1,2,\ldots\}$. For $1 \leq p < \infty$, $c > 0$, and arbitrary finite subsets $X = \{x_1,\ldots,x_m\}$ and $Y = \{y_1,\ldots,y_n\}$ of $W$, where $m, n \in \mathbb{N}_0 = \{0,1,2,\ldots\}$, define

$$\bar{d}_p^{(c)}(X,Y) := \left(\frac{1}{n}\left(\min_{\pi\in\Pi_n}\sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})^p + c^p(n-m)\right)\right)^{1/p} \tag{3}$$

if $m \leq n$, and $\bar{d}_p^{(c)}(X,Y) := \bar{d}_p^{(c)}(Y,X)$ if $m > n$; moreover,

$$\bar{d}_\infty^{(c)}(X,Y) := \begin{cases} \min\limits_{\pi\in\Pi_n} \max\limits_{1\leq i\leq n} d^{(c)}(x_i, y_{\pi(i)}) & \text{if } m = n \\ c & \text{if } m \neq n; \end{cases} \tag{4}$$

in either case set the distance to zero if $m = n = 0$. We call the function $\bar{d}_p^{(c)}$ the *OSPA metric* of order $p$ with cut-off $c$, where OSPA stands for Optimal SubPattern Assignment.

It is by no means obvious that $\bar{d}_p^{(c)}$ is indeed a metric. This has been proved for the case $p = 1$ and $c = 1$ in [17]. We give a proof for the general case in the appendix.

From a practical point of view, for $p < \infty$ and assuming that $m \leq n$, the OSPA distance between two point patterns $X$ and $Y$ is obtained by going through the following three steps:

1) find the $m$ point subpattern (subset consisting of $m$ elements) of $Y$ that is closest to $X$ in terms of the $p$-th order Wasserstein metric (= $p$-th order OMAT metric), which results in an optimal point assignment (see Fig. 3);
2) for each point $y_j$ of $Y$, let $\alpha_j$ be the cut-off value $c$ if there is no point assigned to it or else let $\alpha_j$ be the minimum of $c$ and the distance to its assigned point in $X$;
3) compute the $p$-th order average $\left(\frac{1}{n}\sum_{j=1}^{n}\alpha_j^p\right)^{1/p}$ of the numbers $\alpha_1,\ldots,\alpha_n$.

[1] An arbitrary metric $d^{(c)}$ with values in $[0,c]$ could be used; in particular, we can choose $d^{(c)}(x,y) = t(d(x,y))$, where $t : [0,\infty) \to [0,c]$ is any transformation that is non-decreasing, subadditive (i.e. $t(u+v) \leq t(u)+t(v)$ for all $u, v \geq 0$), and satisfies $t(u) = 0$ if and only if $u = 0$.

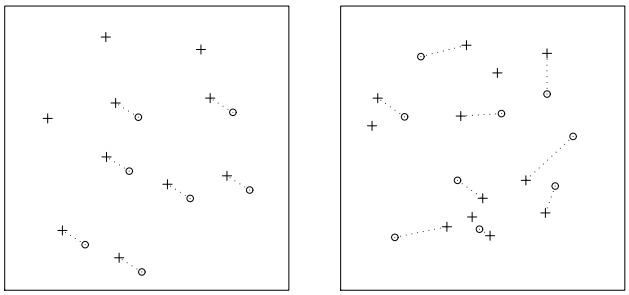


Fig. 3. Two examples of an optimal subpattern assignment (for $p = 1$ and $c = 200$). In the scenario on the left, the optimal assignment is immediately clear from inspection; on the right the situation is more complicated and has been obtained using the Hungarian method (see Subsection III-C).

*Example:* In order to illustrate the three-step instruction, we compute the OSPA distance between the two point patterns in the left panel of Fig. 3, assuming the window to be $1000 \times 1000m^2$, $p = 1$ and $c = 200m$. The optimal point assignment is obtained by inspection (clearly, for any other assignment, the average length of the dotted lines would be larger). Going through each point + of $Y$ (the point pattern with the larger cardinality), we obtain that three of the $\alpha_j$s are equal to the cutoff value $200m$ and seven of the $\alpha_j$s are equal to the length of a dotted line, which is $90m$. Therefore, we have $\bar{d}_1^{(c)}(X,Y) = \frac{1}{10}(3\cdot 200 + 7\cdot 90)m = 123m$.

*Remark:* A slightly different way of constructing the OSPA metric is the following: we fill up the point pattern $X$, which has the smaller cardinality $m$, with $n - m$ "dummy" points located at distance $\geq c$ of any points in $Y$ (we typically need to extend our window $W$ in order to do that), and then calculate the $p$-th order Wasserstein metric between the resulting point patterns. While the three-step instruction is more suitable for practical considerations, this second interpretation is helpful for theoretical reasoning. It furthermore provides us with an efficient way of computing $\bar{d}_p^{(c)}$ as will be seen in Subsection III-C.

### A. Interpreting $p$ and $c$

The order parameter $p$ plays a role similar to the one it has in the OMAT metric (compare the discussion after Equation (14) in [14]). As $p$ increases and $c$ remains fixed, the metric $\bar{d}_p^{(c)}$ becomes ever more unforgiving to "outlier" estimates that are not close to any objects of the ground truth. This happens because the $p$-th order average assigns more weight to exceptionally large values among the $\alpha_j$ the larger $p$ is (compare also Equation (7) in the appendix). Note, however, that in the case of the OSPA metric this effect is somewhat

mitigated by the fact that distances between points are cut off at $c$, which is also the penalty that a point gets when it is deemed "unassignable". Using Hölder's inequality it can be shown that OSPA distances are ordered according to the value of $p$, i.e.

$$\bar{d}_{p_1}^{(c)}(X,Y) \leq \bar{d}_{p_2}^{(c)}(X,Y) \quad \text{for } 1 \leq p_1 < p_2 \leq \infty \text{ and } c > 0.$$

The cut-off parameter $c$ determines the weighting of how the metric penalizes cardinality errors as opposed to localization errors. For $p = 1$ it is exactly the penalty given to any false or missing point estimate. It thus represents the threshold at which, with respect to the local situation, we no longer distinguish between whether two points of different point patterns are paired together or whether one of the points remains unassigned while the other can make an exact match. If they were any closer together, we would prefer pairing the two points; if they were further apart, we would prefer to say that they are unrelated. From the definition of the OSPA metric it is clear that

$$\bar{d}_p^{(c)}(X,Y) \in [0,c] \text{ for any } c > 0 \tag{5}$$

and that

$$\bar{d}_p^{(c_1)}(X,Y) \leq \bar{d}_p^{(c_2)}(X,Y) \text{ for } 0 < c_1 < c_2 < \infty.$$

Equation (5) gives us some "calibration" for $\bar{d}_p^{(c)}$; that is, we may now also consider how the metric performs in relation to the worst possible distance $c$. For more applied considerations and guidelines on how to choose $c$ in concrete situations, see the corresponding discussion in Subsection III-D.

### B. Solving the problems of the earlier metrics

We briefly present the solutions that the OSPA metric has to offer with regard to the five problems identified in Subsection II-B.

*1) Consistency:* The consistency problem of Scenarios D to F in Fig. 2 is resolved; see the distances given in the same. The OSPA metric penalizes relative differences in cardinality in an impartial way by introducing an additive component on top of the average distance in the optimal subpattern assignment. In this way differences in cardinalities cannot go unnoticed (for a reasonably high cut-off $c$; see the discussion in Subsection III-A) in cases where an "unfortunate" positioning of the points occurs (see Scenarios C, D and F). Furthermore, our investigations have revealed no other serious inconsistencies.

*2) Intuitive construction:* In Section I-B of [14], the authors voice their concern that the process of assigning objects (without splitting up their masses) between two point patterns with different cardinalities by declaring either "false" or "missing" points may violate the metric axioms and can easily become arbitrary, which might then favour one multi-object filtering algorithm over another for spurious reasons. Our construction, however, eliminates any element of arbitrariness by providing an objective and intuitively reasonable criterion for the assignment, while at the same time observing the metric axioms. It is furthermore conceptually easier than splitting up individual objects into many small parts (compare for example the left hand side of Fig. 3).

*3) Geometry dependence:* For a given cut-off and a given value of $p$, the $\bar{d}_p^{(c)}$-distance does not substantially depend on the size of the ground truth pattern (see the scenario in Subsection IV-B, where the targets are moving radially outwards). This is because extra points are essentially penalized according to the cut-off $c$ rather than according to their distance from other points in the pattern. While we recommend basing the choice of $c$ only on the most fundamental a priori requirements of the test case, such as the size of the observation window, sensor accuracy, and possibly expected number of targets, more subtle adaptations may be justified in special situations, and the OSPA metric leaves this possibility open.

*4) Cardinality zero:* The OSPA metric is defined between any two point patterns. It is equal to $c$ (i.e. maximal) if one of the two patterns is empty and the other is non-empty. While sometimes it would be desirable to have different values according to the cardinality of the non-empty pattern, this is only a minor inconvenience.

*5) Compatibility with mathematical theory:* Section 2 of [17] contains several theoretical properties of the metric $\bar{d}_1^{(1)}$, which can all be adapted for $\bar{d}_p^{(c)}$. Technical discussions of these properties are, for the large part, beyond the scope of this article. Nonetheless, they show that $\bar{d}_p^{(c)}$ is compatible with the mathematical theory. One aspect that we would like to stress is that it can be shown that $\bar{d}_p^{(c)}$ generates the vague topology on the space of finite point patterns on $W$, which is the standard topology used in point process theory (see [18], Section 15.7, or [19], Section 1.1 and Appendix A)[2]. While this is not the same as the Matheron topology, it is in our opinion even better suited as a theoretical base for performance evaluation in multi-object filtering because it takes appropriate care of differing cardinalities. The most important difference between the two topologies is that, if $X_n$ consists of just two points that merge into one as $n \to \infty$, then $X_n$ converges with respect to the Matheron topology against a point pattern that has just one point (e.g. $\{-\frac{1}{n}, \frac{1}{n}\} \to \{0\}$ for $W = [-1,1] \subset \mathbb{R}$), but does not converge with respect to the vague topology. However, it converges against a point pattern that has a double point at one location (at 0 in the above example) if we choose to admit such patterns for theoretical considerations.

### C. Computation of the OSPA metric

The OSPA metric can be computed efficiently by using the Hungarian method for optimal point assignment (see [21], Section 11.2). For $p < \infty$ and two point patterns $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$ with $m \leq n$, we use the distance matrix $D = (D_{i,j})_{1 \leq i,j \leq n}$, where $D_{i,j} = d^{(c)}(x_i, y_j)^p$ if $1 \leq i \leq m$ and $1 \leq j \leq n$, and $D_{i,j} = c^p$ otherwise. This corresponds to the introduction of $n - m$ "dummy" points at distance $\geq c$ of any points in $Y$ that was described earlier.

The Hungarian method is known to have an asymptotic complexity that is cubic in the dimension of the distance matrix, so that we obtain $O\big(\max(m,n)^3\big)$ for computing the

[2]Some authors, such as [20] (see Section A2.3), prefer to use the weak topology. However, we are not concerned with such fine distinctions here because for our choice of $W$ the two topologies coincide.

OSPA distance. While we are not aware of a method for computing the OMAT distance that is proven to be of a theoretical complexity that is just as good, we remark that, for practical purposes, the OSPA and OMAT metrics can both be computed with similar ease and efficiency (see Section V in [14] for methods of computing the OMAT metric).

### D. Use of the OSPA metric in multi-object estimation

In the context of multi-object performance evaluation, we can interpret the OSPA-distance as a $p$-th order "per-object" error[3]. This error is comprised of two components each separately accounting for "localization" and "cardinality" errors. Precisely, for $p < \infty$ these components are given by

$$\begin{aligned} \bar{e}^{(c)}_{p,\text{loc}}(X,Y) &:= \left( \frac{1}{n} \cdot \min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})^p \right)^{1/p}, \\ \bar{e}^{(c)}_{p,\text{card}}(X,Y) &:= \left( \frac{c^p(n-m)}{n} \right)^{1/p} \end{aligned} \tag{6}$$

if $m \leq n$, and $\bar{e}^{(c)}_{p,\text{loc}}(X,Y) := \bar{e}^{(c)}_{p,\text{loc}}(Y,X)$, $\bar{e}^{(c)}_{p,\text{card}}(X,Y) := \bar{e}^{(c)}_{p,\text{card}}(Y,X)$ if $m > n$. They can thus be interpreted as contributions due to localization only (within the optimal subpattern assignment) and cardinality only (penalized at maximal distance). Note however, that the functions $\bar{e}^{(c)}_{p,\text{loc}}$ and $\bar{e}^{(c)}_{p,\text{card}}$ are themselves *not* metrics on the space of finite subsets. Nonetheless, $\bar{e}^{(c)}_{p,\text{loc}}$ can be considered a metric on the space of finite subsets with fixed cardinality (i.e. once the optimal assignment is determined), and $\bar{e}^{(c)}_{p,\text{card}}$ can be considered a metric on the space of non-negative integers (i.e. using the cardinalities of the sets only). We also remark that the decomposition of the OSPA metric into separate components is usually not necessary for performance evaluation, but may provide valuable additional information.

As discussed in Subsection III-A, the value of $p$ determines the sensitivity of $\bar{d}^{(c)}_p$ to outlier estimates. We emphasize here two important choices. For $p = 1$, we have the benefit that the OSPA-metric measures a first order per-object error and that the sum of the localization and cardinality components equals the total metric, which facilitates a direct interpretation of the metric and its components. However, $p = 2$ is usually a more practical choice since it yields smooth distance curves and is traditionally encountered in other metrics that use a $p$-th order average construction. In subsequent parts of this paper we will not focus on the effect of choosing different values of $p$ and only consider $p = 2$ for convenience.

As alluded to in Subsection III-A, the value of the cut-off $c$ determines the relative weighting of the cardinality error component against the localization error component as parts of the total error. Smaller values of $c$ tend to emphasize localization errors and make the metric mostly insensitive to cardinality errors, whereas larger values of $c$ predominantly indicate cardinality errors and ignore localization errors. The following guidelines can be used for choosing the cut-off value $c$. A value of $c$ which corresponds to the magnitude of a typical localization error can be considered small and has the effect of emphasizing localization errors. A value of $c$ which corresponds to the maximal distance between objects can be considered large and has the effect of emphasizing cardinality errors. Any value of $c$ significantly larger than a typical localization error, but significantly smaller than the maximal distance between objects, can be considered moderate and maintains a balance between the two components. Ultimately, the value of $c$ should be chosen to answer the question with what distance ("how many meters") the designer wants to penalize a false or missing estimate, which in concrete application narrows down the reasonable choices significantly. For our experiments in the next section, we consider each case of a small, a moderate and a large value of $c$.

[3]Strictly speaking, the "per-object" error is a "per-estimated-object" error when the cardinality is overestimated and a "per-true-object" error when the cardinality is underestimated.

## IV. Experiments

We demonstrate the proposed OSPA metric with two multi-target tracking applications. In the first, we employ a typical scenario to illustrate the use and interpretation of the OSPA metric and compare with the OMAT metric. In the second, we employ a more extreme scenario to illustrate how the OMAT metric is unable to distinguish good and bad performance due to its inconsistencies and geometry dependence, and then illustrate how the OSPA metric resolves these drawbacks.

In both scenarios, we use the following dynamical and observation model. The duration of both scenarios is $K = 100s$. The observation window is the square $W = [-1000, 1000]m \times [-1000, 1000]m$. The individual target states are 4-D vectors $x_k = [\, p_{x,k}, p_{y,k}, \dot{p}_{x,k}, \dot{p}_{y,k} \,]^T$ of $x$ and $y$ position and velocity. Individual target motions follow a linear Gaussian constant velocity motion model with a sampling period of $\Delta = 1s$ and process noise standard deviation of $\sigma_\nu = 3m/s^2$ on both the $x$ and $y$ directions. Target originated measurements are 2-D vectors $z_k = [\, m_{x,k}, m_{y,k} \,]^T$ of $x$ and $y$ position. Individual coordinate measurements are corrupted by additive Gaussian noise with noise standard deviation of $\sigma_\varepsilon = 10m$ on both the $x$ and $y$ coordinates. The probability of survival is $p_{S,k} = 0.99$ and the probability of detection is $p_{D,k} = 0.98$ in Scenario 1 and $p_{D,k} = 0.90$ in Scenario 2. Clutter follows a uniform Poisson process over the observation window with intensity $\lambda_{c,k} = 1.25 \times 10^{-5} m^{-2}$ on $W$ in Scenario 1 and $\lambda_{c,k} = 2.50 \times 10^{-6} m^{-2}$ on $W$ in Scenario 2. Note that when the metrics are computed, we take only the positional components of the state vector.

### A. Scenario 1

Consider a time-varying number of targets (max of 10) moving with non-constant velocity as shown in Fig. 4. Note that targets 1–6 are born at time $k = 1$, targets 7–10 are born at time $k = 30$, and targets 3–4 die at time $k = 80$. The starting and stopping positions for each track are labelled with a circle and triangle respectively. The multi-target filters used to test the new metric are the Gaussian mixture PHD [22] and Gaussian mixture CPHD [23] filters. Although it is not our primary intention here to examine these filters in detail, we note that the former is known to have much higher variance on its estimate of the number of targets than the latter.

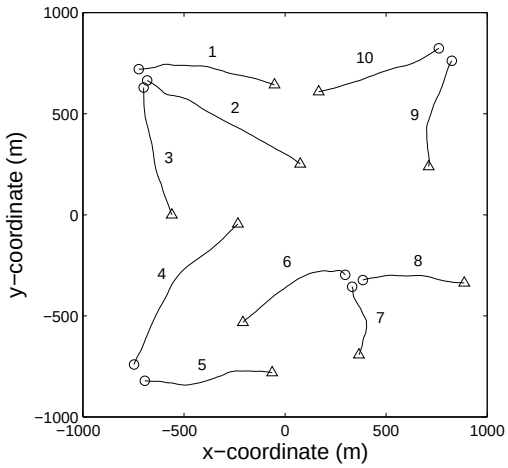



Fig. 4. True tracks in the $xy$ plane. Targets move with non-constant velocity along the paths shown. Start/Stop positions are shown with ○/△.

To capture the average performance, we run 1000 Monte Carlo (MC) trials for each filter with the same target tracks but independently generated measurements. Fig. 5 shows the mean and standard deviation of the estimated cardinality distribution. These results confirm that both filters provide unbiased estimates of the multi-target cardinality. It can also be seen that the PHD filter has a significantly larger variance on the estimated cardinality, but has a faster response to changes in cardinality than the CPHD filter.

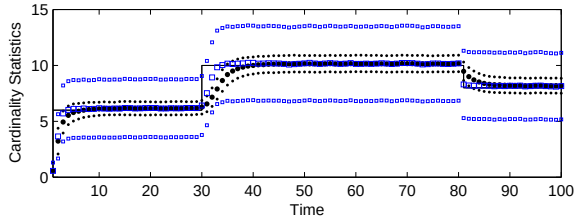



Fig. 5. True cardinality (solid line) shown versus estimated mean cardinality and corresponding standard deviation for the CPHD filter (•/•) and similarly for the PHD filter (▫/▫). Both filters are unbiased, and the CPHD filter has a lower variance but slower response to changes than the PHD filter.

In Fig. 6, the MC average of the OMAT distance for $p = 2$ is shown. Although these numerical results are consistent with the results of Fig. 5, it is not at all clear how the numerical values produced by the OMAT metric should be interpreted, especially in terms of localization and cardinality error. These results demonstrate the contrived nature of the OMAT metric in general as discussed in Subsection II-B. In contrast, the OSPA metric proposed in Section III has a much more natural and intuitive interpretation.

We illustrate the effect of varying the cut-off parameter $c$ in Fig. 7, which shows the MC average of the OSPA distance for $p = 2$ and $c = 30m, 100m, 2828m$. Fig. 7(a) shows the results for $c = 30m$ (in fact $c = 3\sigma_\varepsilon$) which suggests that on average the CPHD and PHD filters are making an error of roughly $11m$ and $12m$ respectively per target. It can be seen that the CPHD filter appears to perform slightly better than the PHD filter due to the disparity in the variance of their cardinality estimates. However, in this case the actual magnitude of the difference of roughly $1m$ per target is relatively small, since the choice of a small cut-off effectively tells the metric not to penalize cardinality errors harshly and hence to reflect mainly localization errors.

Fig. 7(b) shows the results for $c = 100m$ where the CPHD and PHD filters make average errors of roughly $20m$ and $26m$ respectively per target. Both the magnitude of the errors themselves and the magnitude of the difference between the errors of the two filters have increased, due to the choice of a moderate cut-off, which has the effect of balancing between the localization and cardinality errors.

Fig. 7(c) shows the results for $c = 2828m$ (the maximum possible distance, i.e. the diagonal of the observation window), which suggests that the CPHD and PHD filters are making errors of $300m$ and $500m$ respectively per target. In this case the metric is predominantly reporting a cardinality penalty due to the choice of a very high cut-off value relative to the magnitude of a typical localization error.

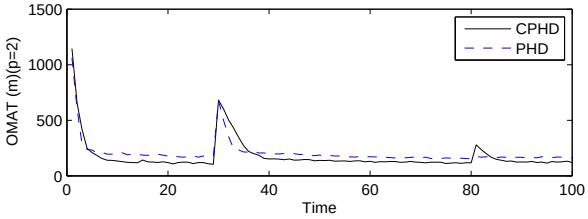



Fig. 6. 1000 MC run average OMAT distance versus time. The numerical values of the OMAT distances are difficult to interpret in terms of a physical distance.

The experiment can be further analyzed by examining the "localization" and "cardinality" components of the OSPA metric given in (6). The MC average of these two components in this scenario for $c = 100m$ is shown in Fig. 8. In terms of localization error, the two filters are penalized in roughly equal amounts, consistent with the standard deviation of the measurement noise. On closer inspection of the localization errors, the PHD filter appears to perform slightly better. This can be attributed to the fact that the optimal subpattern assignment gives a filter a slight advantage for the localization component when its cardinality estimate is wrong: if there are too many estimates, only the best ones are used; if there are too few, they are interpreted as estimates for the closest subset of true targets. In terms of cardinality error, the PHD filter is penalized much more heavily than the CPHD filter as a result of the former having a much larger variance on its cardinality estimate than the latter. Specifically on cardinality errors, the CPHD filter on average has half the penalty of the PHD filter at instants when there are no cardinality changes (see Fig. 5), but the CPHD filter is penalized for a longer duration and/or more severely than the PHD filter at instants where there are cardinality changes (see Fig. 5 at $k = 30s, 80s$) as a result of being slower to respond to the change.

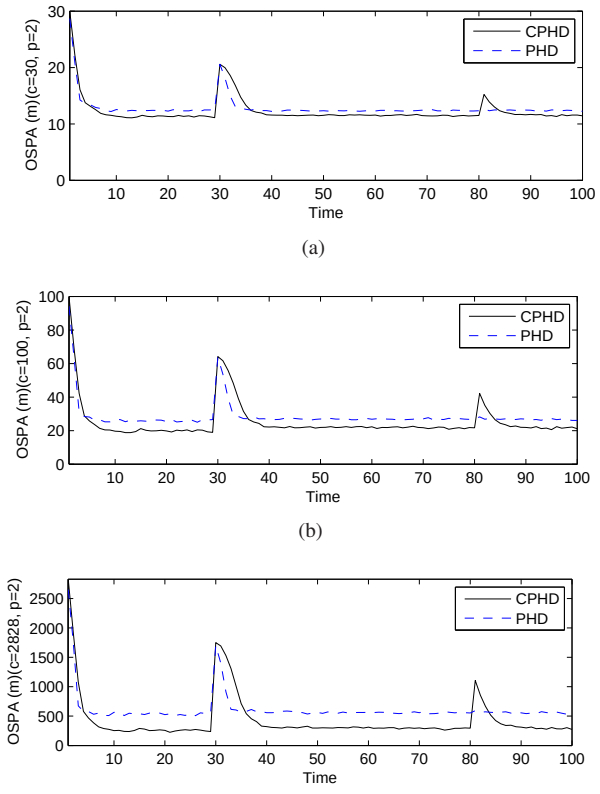



Fig. 7. 1000 MC run average OSPA distance versus time measuring the "per-target" error for various cut-off values. (a) $c = 30m$ emphasizing localization errors; (b) $c = 100m$ balancing between localization and cardinality errors; (c) $c = 2828m$ emphasizing cardinality errors.

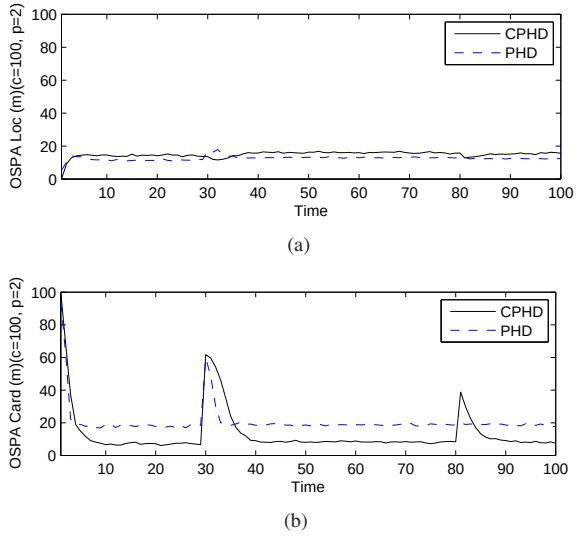



Fig. 8. 1000 MC run average OSPA components. (a) localization component versus time, where both filters have roughly similar errors; (b) cardinality component versus time, where the CPHD filter performs significantly better than the PHD filter due to the large difference in the variance of their cardinality estimates, except at instants of cardinality changes.

### *B. Scenario 2*

Consider 10 targets moving radially outwards with constant velocity as shown in Fig. 9. The multi-target filters used to test the new metric are the Gaussian mixture MeMBer filter [4] and the Gaussian mixture PHD filter [22]. We compare these two filters to demonstrate the inability of the OMAT metric to adequately indicate good and bad performance, and also to illustrate how the proposed OSPA metric resolves these inconsistencies.

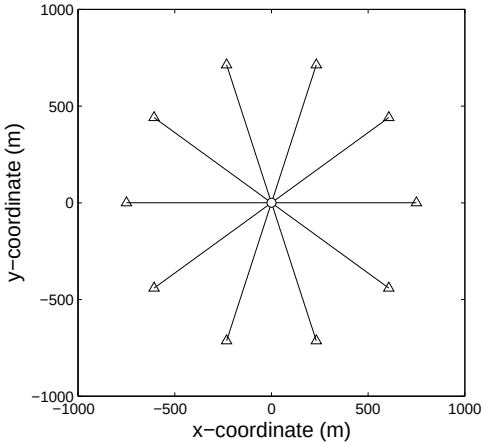



Fig. 9. True tracks in the $xy$ plane. Targets move with constant velocity radially outwards. Start/Stop positions are shown with ○/△.

Again, we perform 1000 Monte Carlo (MC) runs of each filter with the same target tracks but independently generated measurements. We observed that the MeMBer filter is biased in its estimate of the multi-target cardinality, and that the bias is systematic in the sense that the filter tends to output as estimates both its time-predicted and measurement-updated tracks for each target present. This observation will be pertinent to the interpretation of our results. In fact, the MC average mean and standard deviation of the estimated cardinality distribution, shown in Fig. 10 indicate that the MeMBer filter overestimates the cardinality by a factor of approximately 2, whereas the PHD filter is unbiased in this respect.

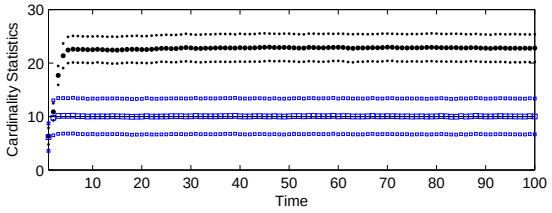



Fig. 10. True cardinality (solid line) shown versus estimated mean cardinality and corresponding standard deviation for the MeMBer filter (•/•) and similarly for the PHD filter (□/□). The MeMBer filter is biased but the PHD filter is unbiased.

In Fig. 11, the MC average of the OMAT distance for $p = 2$ is shown. We can see at first glance that the values of the OMAT distance increase with time. These results are

not an indication of filter divergence, but are actually caused by the geometry dependence of the OMAT metric reflecting the outward radial travel of the targets. Moreover, the relative positions of the two curves suggest that the MeMBer filter performs about equally well as the PHD filter (even marginally better, in fact), in spite of the former committing significant cardinality errors and the latter being unbiased. However, these results are not an accurate evaluation of filter performance since the situation here resembles that of Scenario C in Fig. 2 and is actually caused by the OMAT metric being insensitive to multiple estimates concentrated around a single true point. Furthermore, when cardinality errors arise in this scenario such that it resembles that of Scenarios D–F in Fig. 2, the OMAT metric becomes yet more difficult to interpret. We are thus faced with a practical situation where the OMAT metric is not a reliable measure of filter performance.

In Fig. 12, the MC average of the OSPA distance for $p = 2$ and $c = 100m$ is shown. It can be seen that after the initial settle-in phase the curves stabilize to an average error of $75m$ and $35m$ per target for the MeMBer and PHD filters respectively. These results confirm that the OSPA metric is not affected by the geometry of the scenario, and that it penalizes the MeMBer filter for cardinality errors appropriately compared to the PHD filter. Examining the "localization" and "cardinality" components of the OSPA metric given in Fig. 13 further verifies our observations. On localization errors, the PHD filter settles to an error consistent with the standard deviation of the measurement noise, whilst the MeMBer filter achieves an unexpectedly lower error as a direct result of always double-guessing the true location of each target. In fact, in terms of localization errors only, the MeMBer filter is at a significant advantage in this scenario since for each target present, this filter consistently outputs its time-predicted track as one of its states estimates, which is likely to be very accurate considering the perfectly straight line paths of the targets and the relatively low probability of detection encountered here. On cardinality errors, the results indicate that the MeMBer filter is penalized much more than the PHD filter for always overestimating the cardinality.

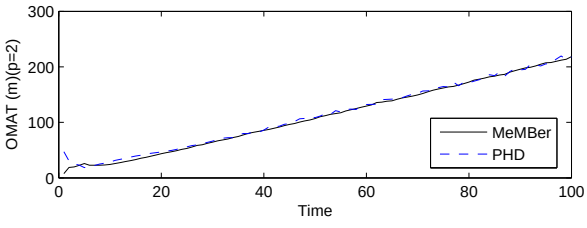



Fig. 11. 1000 MC run average OMAT distance versus time. The distances increase with time due to the geometry dependence of the metric. Also, the MeMBer filter appears to perform equally well as the PHD filter due to inconsistencies of the metric which result in its inability to distinguish good and bad performance.

## V. Conclusion

This paper has identified a number of serious limitations in previous miss-distances used in multi-object systems, including inconsistent behaviour and the lack of a meaningful physical interpretation if the cardinalities of the two finite sets under consideration differ. We have subsequently introduced a new miss-distance, called the optimal subpattern assignment (OSPA) metric, that overcomes these limitations. It allows for a natural physical interpretation even if the cardinalities of two sets are not the same, without exhibiting any elements of arbitrariness inherent in ad hoc assignment approaches.

The OSPA metric has two adjustable parameters $p \in [1, \infty]$ and $c > 0$ that have meaningful interpretations as outlier sensitivity and cardinality penalty, respectively. If for a particular application it is crucial to estimate the number of objects correctly, we can cater for this by choosing a large value for $c$, whereas, if exact position estimates are important and cardinality errors are almost negligible, we would rather choose a small value for $c$. The transparent interpretations for $p$ and $c$ also mean that grossly unsuitable parameter choices are easily seen, which provides a certain degree of protection against abuse from biasing the metric to favour specific scenarios for spurious reasons.

The various advantages of the OSPA metric have been illustrated in two extensive simulation studies.

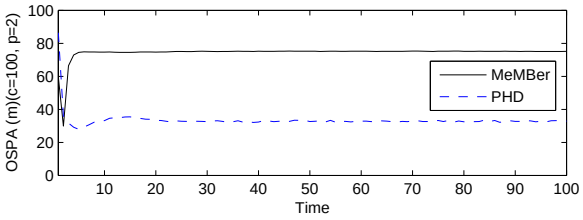
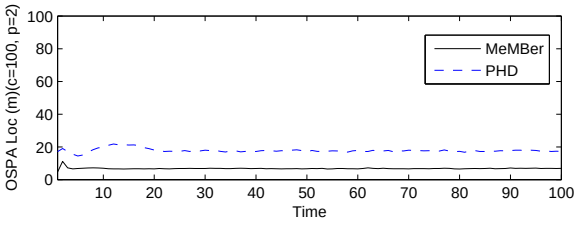
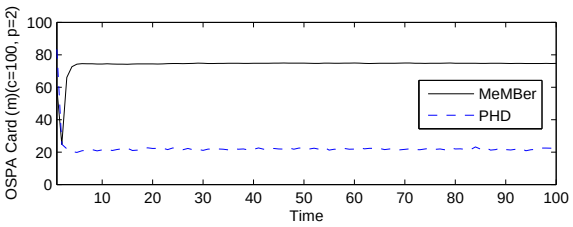



Fig. 12. 1000 MC run average OSPA distance versus time. The OSPA metric does not exhibit any geometry dependence and appropriately penalizes cardinality errors in this scenario.




(a)




(b)

Fig. 13. 1000 MC run average OSPA components. (a) localization component versus time, where the MeMBer filter performs unexpectedly well due to double-guessing, but the PHD filter performs as expected; (b) cardinality component versus time, yielding appropriate penalties for both filters.

While closer investigation of the metric reveals no serious weaknesses, there are several minor issues worth mentioning. First, it would sometimes be desirable to have separate control over the threshold at which physical distances in the state space are cut off and the penalty given to cardinality errors, which in our definition are both determined by the single parameter $c$ (cf. the two occurences of $c$ in Equation (3), first in the role of a distance threshold, then in the role of a cardinality penalty). However, choosing a distance threshold that is larger than the cardinality penalty in the current definition does not usually result in a metric. Secondly, if one of the point patterns is empty, the OSPA metric is insensitive to the cardinality of the non-empty point pattern, which is slightly inconvenient. Lastly, our proposed metric is formulated only on closed and bounded subsets, although this is not a problem in most practical situations.

## APPENDIX
## PROOF THAT $\bar{d}_p^{(c)}$ IS A METRIC

Fix $p \in [1,\infty]$ and $c > 0$. From the definition of $\bar{d}_p^{(c)}$ it is clear that $\bar{d}_p^{(c)}(X,Y) \geq 0$ and that $\bar{d}_p^{(c)}$ satisfies the identity and symmetry properties. Let $X = \{x_1,\ldots,x_l\}$, $Y = \{y_1,\ldots,y_m\}$, and $Z = \{z_1,\ldots,z_n\}$, where $l,m,n \in \mathbb{N}_0$. We show $\bar{d}_p^{(c)}(X,Y) \leq \bar{d}_p^{(c)}(X,Z) + \bar{d}_p^{(c)}(Z,Y)$ (triangle inequality) for $p < \infty$. Since this inequality is symmetric in $X$ and $Y$, we may assume without loss of generality that $l \leq m$. We furthermore assume that at most one of the cardinalities $l,m,n$ is zero because otherwise the inequality is clearly satisfied. Introduce two sequences $(u_i)_{i\in\mathbb{N}}$ and $(v_i)_{i\in\mathbb{N}}$ of pairwise distinct points in $\mathbb{R}^N \setminus W$ (the complement of $W$ in $\mathbb{R}^N$) such that

$$d(u_i,x) \geq c,\ d(v_j,x) \geq c,\ \text{and } d(u_i,v_j) \geq c$$

for all $x \in W$ and all $i,j \in \mathbb{N}$.[4]

*Case 1 ($l \leq m \leq n$):* Let $x_{l+i} := u_i$ for $1 \leq i \leq n-l$ and $y_{m+j} := v_j$ for $1 \leq j \leq n-m$. Choose $\sigma,\tau \in \Pi_n$ such that $\sum_{i=1}^n d^{(c)}(x_i,z_{\sigma(i)})^p = \min_{\pi\in\Pi_n}\sum_{i=1}^n d^{(c)}(x_i,z_{\pi(i)})^p$ and $\sum_{i=1}^n d^{(c)}(z_i,y_{\tau(i)})^p = \min_{\pi\in\Pi_n}\sum_{i=1}^n d^{(c)}(z_i,y_{\pi(i)})^p$. We then obtain

$$\begin{aligned}
&\bar{d}_p^{(c)}(X,Y)\\
&= \left(\frac{1}{m}\min_{\pi\in\Pi_m}\sum_{i=1}^m d^{(c)}(x_i,y_{\pi(i)})^p\right)^{1/p}\\
&\leq \left(\frac{1}{n}\min_{\pi\in\Pi_n}\sum_{i=1}^n d^{(c)}(x_i,y_{\pi(i)})^p\right)^{1/p}\\
&\leq \left(\frac{1}{n}\sum_{i=1}^n\Big(d^{(c)}(x_i,z_{\sigma(i)}) + d^{(c)}(z_{\sigma(i)},y_{\tau(\sigma(i))})\Big)^p\right)^{1/p}\\
&\leq \left(\frac{1}{n}\sum_{i=1}^n d^{(c)}(x_i,z_{\sigma(i)})^p\right)^{1/p}\\
&\qquad + \left(\frac{1}{n}\sum_{i=1}^n d^{(c)}(z_{\sigma(i)},y_{\tau(\sigma(i))})^p\right)^{1/p}\\
&= \bar{d}_p^{(c)}(X,Z) + \bar{d}_p^{(c)}(Z,Y),
\end{aligned}$$

[4]There is often (e.g. for the Euclidean metric) a natural extension of the metric $d$ from $W$ to $\mathbb{R}^N$ that makes such a choice possible. Otherwise, extend $d$ by setting $d(x,y) := \max\big(c, \operatorname{diam}(W)\big)$ if $x,y \in \mathbb{R}^N$ are distinct and not both in $W$, where $\operatorname{diam}(W) = \max\{d(x,y);\ x,y \in W\}$.

using the fact that $a \leq c^p m$ implies $\frac{a}{m} \leq \frac{a+c^p(n-m)}{n}$ for the second relation, and Minkowski's inequality for the fourth relation.

*Case 2 ($l,n \leq m$):* Let $x_{m-i+1} := u_i$ for $1 \leq i \leq m-l$ and $z_{m-j+1} := u_j$ for $1 \leq j \leq m-n$ (note that this implies $x_i = z_i$ for $\max(l,n) \leq i \leq m$). Choose $\sigma,\tau \in \Pi_m$ such that $\sum_{i=1}^m d^{(c)}(x_i,z_{\sigma(i)})^p = \min_{\pi\in\Pi_{l\vee n}}\sum_{i=1}^{l\vee n} d^{(c)}(x_i,z_{\pi(i)})^p$ and $\sum_{i=1}^m d^{(c)}(z_i,y_{\tau(i)})^p = \min_{\pi\in\Pi_m}\sum_{i=1}^m d^{(c)}(z_i,y_{\pi(i)})^p$, where $l \vee n = \max(l,n)$. We then have

$$\begin{aligned}
&\bar{d}_p^{(c)}(X,Y)\\
&= \left(\frac{1}{m}\min_{\pi\in\Pi_m}\sum_{i=1}^m d^{(c)}(x_i,y_{\pi(i)})^p\right)^{1/p}\\
&\leq \left(\frac{1}{m}\sum_{i=1}^m\Big(d^{(c)}(x_i,z_{\sigma(i)}) + d^{(c)}(z_{\sigma(i)},y_{\tau(\sigma(i))})\Big)^p\right)^{1/p}\\
&\leq \left(\frac{1}{m}\sum_{i=1}^m d^{(c)}(x_i,z_{\sigma(i)})^p\right)^{1/p}\\
&\qquad + \left(\frac{1}{m}\sum_{i=1}^m d^{(c)}(z_{\sigma(i)},y_{\tau(\sigma(i))})^p\right)^{1/p}\\
&\leq \bar{d}_p^{(c)}(X,Z) + \bar{d}_p^{(c)}(Z,Y),
\end{aligned}$$

where the third relation follows by Minkowski's inequality.

Since we have $l \leq m$, combining cases 1 and 2 yields the triangle inequality for $p < \infty$. By

$$\lim_{p\to\infty}\left(\frac{1}{m}\sum_{i=1}^m \alpha_i^p\right)^{1/p} = \max_{1\leq i\leq m}\alpha_i \tag{7}$$

for $\alpha_i \geq 0$, it can be seen that $\lim_{p\to\infty}\bar{d}_p^{(c)}(X,Y) = \bar{d}_\infty^{(c)}(X,Y)$. The triangle inequality for $p = \infty$ follows then by letting $p \to \infty$ in the triangle inequality for $p < \infty$. □

## REFERENCES


[1] M. L. Skolnik, *Radar handbook, 2nd Ed.* McGraw-Hill, 1990.

[2] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. Academic Press, San Diego, 1988.

[3] S. Blackman, *Multiple Target Tracking with Radar Applications*. Artech House, Norwood, 1986.

[4] R. Mahler, *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[5] I. J. Cox and S. L. Hingorani, "An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 2, pp. 138–150, Feb. 1996.

[6] D. M. Lane, M. J. Chantler, and D. Dai, "Robust tracking of multiple objects in sector-scan sonar image sequences using optical flow motion estimation," *IEEE J. Ocean. Eng.*, vol. 23, no. 1, pp. 31–46, Jan. 1998.

[7] D. M. Kocak, N. d. V. Lobo, and E. A. Widder, "Computer vision techniques for quantifying, tracking, and identifying bioluminescent plankton," *IEEE J. Ocean. Eng.*, vol. 24, no. 1, pp. 81–95, Jan. 1999.

[8] H. Durrant-Whyte and T. Bailey, "Simultaneous localisation and mapping (SLAM): Part I, the essential algorithms," *IEEE Robotics and Automation Magazine*, vol. 13, no. 2, pp. 99–110, Jun. 2006.

[9] U. Spagnolini and V. Rampa, "Multitarget detection/tracking for monostatic ground penetrating radar: application to pavement profiling," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 383–394, Jan. 1999.

[10] B. Hammarberg, C. Forster, and E. Torebjork, "Parameter estimation of human nerve C-fibers using matched filtering and multiple hypothesis tracking," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 4, pp. 329–336, Apr. 2002.

[11] O. E. Drummond, "Methodologies for performance evaluation of multitarget multisensor tracking," in *O. E. Drummond (ed.) Signal and Data Processing of Small Targets, Proc. SPIE*, vol. 3809, pp. 355–369, 1999.

[12] O. E. Drummond and B. E. Fridling, "Ambiguities in evaluating performance of multiple target tracking algorithms," in *O. E. Drummond (ed.) Signal and Data Processing of Small Targets, Proc. SPIE*, vol. 1698, pp. 326–337, 1992.
[13] R. L. Rothrock and O. E. Drummond, "Performance metrics for multiple-sensor, multiple-target tracking," in *O. E. Drummond (ed.) Signal and Data Processing of Small Targets, Proc. SPIE*, vol. 4048, pp. 521–531, 2000.
[14] J. Hoffman and R. Mahler, "Multitarget miss distance via optimal assignment," *IEEE Trans. Sys., Man, and Cybernetics-Part A*, vol. 34, no. 3, pp. 327–336, 2004.
[15] G. Matheron, *Random Sets and Integral Geometry*. Wiley, 1975.
[16] A. Baddeley, "Errors in binary images and an $L^p$ version of the Hausdorff metric," *Nieuw Archief voor Wiskunde*, vol. 10, pp. 157–183, 1992.
[17] D. Schuhmacher and A. Xia, "A new metric between point process distributions," *Submitted*, Aug. 2007, available online at http://arxiv.org/abs/0708.2777.
[18] O. Kallenberg, *Random measures, 4th Ed.* Akademie-Verlag, 1986.
[19] A. Karr, *Point processes and their statistical inference, 2nd Ed.* Marcel Dekker Inc., 1991.
[20] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes Vol. I, 2nd Ed.* Springer, 2003.
[21] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization; algorithms and complexity*. Dover Publications, 1998.
[22] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. SP.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.
[23] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic implementations of the cardinalized probability hypothesis density filter," *IEEE Trans. SP.*, vol. 55, no. 7, pp. 3553–3567, July 2007.