The Conditionalizing Identity Management Bayesian Filter (*CIMBal*)

Qirong Ho Christopher Geyer

10/28/2008

CMU-RI-TR-08-47

April 2009

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Abstract

We present a large-scale data association tracker that can handle variable numbers of world objects and measurements. Large-scale data association problems arise in surveillance, wildlife monitoring, and applications of sensor networks. Several approaches have recently been proposed that represent the uncertainty in data association using a parameterized family of distributions on the set of permutations. Whereas these approaches were restricted to a fixed and known number of objects (and sometimes measurements), we generalize these approaches to varying numbers of objects and measurements. We also present a modification that allows one to focus on a set of objects of interest, while maintaining data association with all other objects that may be confused with these objects of interest. We justify the approach with an analysis and show experiments on a large-scale simulated tracking sequence.

Contents

1 Introduction

The problem of tracking and identifying multiple objects features prominently in a number of real-world applications, such as surveillance (Pasula, Russell, Ostland, and Ritov 1999), air traffic control (Kondor, Howard, and Jebara 2007), or counting animals in the wildlife (Betke, Hirsh, Bagchi, Hristov, Makris, and Kunz 2007). For example, consider that extremely high-resolution cameras—approaching 1 gigapixel—that could surveil every single vehicle in the Washington D.C. metropolitan area (Page 2007) could soon be deployed at high altitudes in military scenarios. Such sensors could eventually track tens of thousands of vehicles. Identity information is likely to be sparse and uncertain, such as "Vehicle X left address A," implying something about X's identity. Trackers that greedily estimate data association, or that defer it by maintaining multiple hypotheses, are infeasible with this number of objects and with this level of association ambiguity.

In this paper we improve upon recent statistics-based trackers, which have the advantage of non-combinatorially deferring data association decisions. We allow for a variable number of objects as well as a variable number of state measurements of these objects, and we consider a sub-problem where the user may be interested in focusing on a small set of objects within scenes that have a large number of confusing objects. First, we describe some related work.

1.1 Related Work

The *data association* problem in multi-object tracking is the task of maintaining the unique identity of an object while it is being tracked by one or more sensors. Eventually each measurement should be *associated* with or assigned to some object, and in some scenarios where clutter is prominent, measurements may not correspond to any object at all. We focus on non-batch methods to solve this problem, and we group methods into four categories: (1) *greedy trackers*: methods that greedily assign measurements to objects and maintain a single data association (DA) hypothesis immediately after a measurement is received, such as the *joint probabilistic data association* (JPDA) (Fortmann, Bar-Shalom, and Scheffe 1980) tracker; (2) *finite horizon* or *finite hypothesis trackers*: multi-hypothesis trackers that maintain a bounded number of possible hypotheses, which are constructed by branching at object crossings, e.g. the *multi-hypothesis tracker* (MHT) (Reid 1979); (3) *sampling-based trackers*: trackers based on Monte Carlo sampling methods that sample and weigh particles on the set of possible data associations, which typically also have a finite horizon, e.g. Dellaert *et al.* (2007) and Oh *et al.* (2005); and, (4) *statistics-based trackers*: methods that maintain a statistic that parameterizes a distribution on the set of all possible data associations.

We focus primarily on the last case, statistics-based trackers, which were the first to address large-scale data association problems, where objects number in the thousands. Statistics-based trackers provide a non-combinatorial way to recursively filter in the combinatorial state-space. Like the Kalman filter, statistics-based methods maintain a statistic that parameterizes distributions on the space of data associations. In the case of the Kalman filter, the covariance and mean parameterize Gaussians. Thus, instead of maintaining some number of hypotheses, statistics-based approaches update a data association statistic that encodes past identity readings; however, the statistic is not claimed to be sufficient. Shin *et al.* (2003) maintain an "identity belief matrix" that roughly constitutes a "sufficient" statistic of past measurements of data associations. Unlike the MHT, computations and data structures are independent of the measurements; and, unlike JDPA, the data associations can be deferred or recomputed using a statistic that "remembers" old measurements. Schmutisch *et al.*'s (2006) IMKF approximately performs a Bayesian update, and also diffuses the statistic according to a process model. Kondor *et al.* (2007) consider the connection between these approaches and the Fourier transform of the PMF on the set of data associations, which they assume to be the group of permutations. Huang *et al.* (2007) consider implementing Bayesian updates in the Fourier domain, as well as the effects of bandlimiting.

1.2 Our Contribution

Despite the improvements made by the IMKF, using it for real-time tracking is still made difficult by the fact that an $\mathrm{O}(n^3)$ algorithm — the Hungarian algorithm — needs to be used per time step. The IMKF was also restricted to a constant number of objects N and measurements M, where $N = M$. Our contribution is an improved algorithm called the *C*onditioning *I*dentity *M*anagement *Ba*yesian Fi*l*ter (*CIMBal*), which builds upon the IMKF in the following ways:

- We generalize the IMKF model to allow for varying numbers of objects and measurements.
- We show that eliminating a row and a column from the identity belief matrix (the maintained "sufficient" statistic in the IMKF) is equivalent to conditioning upon the hypothesis that the measurement associated with the column belongs to the object associated with the row, thus providing a rigorous justification for pruning the matrix.
- We permit situations in which only a small subset of the objects are really of interest, and use this assumption to minimize the size of the identity belief matrix.

These improvements not only allow the CIMBal to be applied to a wider range of scenarios than the IMKF can handle, but also reduce the runtime per time step to $\mathrm{O}(z^3)$ where typically $z \ll N$, provided that the subset of objects of interest is small.

2 Problem Formulation and Notation

We suppose that there are n objects in the world whose continuous state are of interest. We use $\mathbb{S}_n$ to denote the group of permutations on a set of size n, which we represent using permutation matrices. $\mathcal{N}(0, \Sigma)$ denotes a Gaussian distribution with mean μ and covariance Σ. We make the following assumptions about objects, identities and targets/measurements:

1. *Data association model at time t*: Let $\mathcal{W}$ be a set of trackable world objects. We assume that there are N *identities*, and an injective mapping $id : [N] \rightarrow \mathcal{W}$. We shall use the term *labeled object* to denote any object with an identity associated to it by id. We make the following assumptions:

   (a) At each time step t, a tracker provides M_t state *measurements*. Measurements represent readings of the state of world objects, and we say that a measurement "corresponds" to a world object if and only if the measurement was taken from that world object. We allow situations in which some measurements do not correspond to any labeled object, or even to any world object at all ("ghost" measurements).

   (b) There exists a *true* correspondence from measurements to world objects, given by a mapping $m_t : [M_t] \rightarrow \mathcal{W} \cup \text{Ø}$. $m_t(j) = \text{Ø}$ denotes that measurement j does not correspond to any labeled world object.

   We also assume that both M_t and m_t are time-dependent, as denoted by the subscripted t. Furthermore, we require that m_t be injective for all t, i.e. our model does not permit two or more measurements to correspond to the same world object. Even if there is good reason to suppose that multiple measurements correspond to a single world object, e.g. the tracker produces duplicate measurements, we will nevertheless assume that only one measurement can be truly associated with that world object.

   To illustrate our model with an example, $\mathcal{W}$ might be the set of cars in the world, id the mapping from license plate numbers to cars, and M_t the number of cars the tracker "sees" — this might include false positives, or cases where the same car is mistakenly detected twice or more. The goal is to estimate $id^{-1} \circ m_t$, so as to track the continuous state of the world objects by using their corresponding measurements.

2. *World object and measurement model*: Let $x_t^i \in \mathbb{R}^d$ represent the state of (labeled) world object $id(i)$. We assume that the change of state of all world objects is object-independent and obeys a linear model:

$$x_{t+1}^i = A\, x_t^i + w_t^i \tag{1}$$

   where A is the state transition matrix and $w_t^i \sim \mathcal{N}(0, \Sigma_w)$ for some covariance matrix Σ_w. If $id(i) = m_t(j)$, then $z_t^j \in \mathbb{R}^f$ is the measurement corresponding to world object $id(i)$, and follows the linear model:

$$z_t^j = H x_t^i + v_t^j \tag{2}$$

   where H is the observation matrix, and $v_t^j \sim \mathcal{N}(0, \Sigma_v)$ for some covariance matrix Σ_v.

3. *Data association transition in the tracker*: Consider the measurement mappings m_t and m_{t+1}. We make no prior assumptions about the relationship between $m_t(j)$ and $m_{t+1}(k)$ for any j, k. For example, the tracker might assign different indices to measurements taken at different times but corresponding to the

same labeled world object, i.e. $m_t(j) = id(i) = m_t(k)$ for some j, k. Or, perhaps the tracker might begin taking measurements for some labeled world object that hitherto lacked corresponding measurements, i.e. $\exists i(\nexists j(m_t(j) = id(i)) \wedge \exists k(m_{t+1}(k) = id(i)))$. It might also be the case that the tracker stops taking measurements of a labeled world object, i.e. $\exists i(\exists j(m_t(j) = id(i)) \wedge \nexists k(m_{t+1}(k) = id(i)))$.

In order to characterize the transition from m_t to m_{t+1}, we define an $M_{t+1} \times M_t$ association transition matrix Q:

$$Q^t_{j,i} = \begin{cases} 1 & m_{t+1}(j) = m_t(i) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

There are some specifics about Q worth mentioning:

- Row j of Q being zero implies that either measurement z^j_{t+1} corresponds to a labeled world object lacking a corresponding measurement at time t, or z^j_{t+1} does not correspond to a labeled world object at all.
- Column i of Q being zero either implies that measurement z^i_t's corresponding labeled world object is no longer being measured at time $t+1$, or z^i_t did not correspond to a labeled world object to start with.
- The injectivity of m_t for all t restricts each row and column to contain no more than one '1'.

We assume that the tracker also provides a distribution ψ_t over possible transition association matrices Q^t. ψ_t is conditioned on the state of the tracker, represented by T_t. $\psi_t(Q)$ is the probability that Q represents the *true association transition* between m_t and m_{t+1}. **In other words, we are assuming that the tracker can give us probabilistic information ψ_t about how it thinks measurements relate to each other between timesteps.** Notice that ψ_t does not encode the world object identities corresponding to each measurement; the mechanism used to communicate such identity information will be described shortly.

4. *Conditional independence of transitions*: We make the assumption that the true association transition (given by some matrix Q^t) is conditionally independent of $\{x^i_s\}$ for all i and $s \leq t$, given T_t. That is to say the trajectories and dynamics of all world objects, if not encoded in T_t, are assumed to not reveal extra information about the evolution of the data association. The IMKF operates under this assumption, whereas the JPDA and MHT do not. Should one wish to make use of such extra information, there is always the option to process ψ_t before it is passed to the CIMBal.

5. *Identity readings*: We assume the tracker obtains intermittent world object identity information about its measurements, which we shall term *identity readings*. These identity readings are represented by an $N_t \times M_t$ matrix ω_t that encodes log-identity information of the form $\ln \mathrm{pr}(x^i_t \to z^j_t | T_t)$, where $x^i_t \to z^j_t$ represents the event that $id_t(i) = m_t(j)$. This will be further elaborated under the CIMBal's "Update Step".

Now that we have laid the definitions, we can state the problem goal:

> **To estimate the true data association from measurements to identities at each time step, using the tracker-provided data z_t^i, ψ_t, and ω_t.**

This enables us to estimate the continuous state of the world objects corresponding to the identities being tracked.

3 Parameterizing Data Association PMFs

For each time step t, the IMKF and CIMBal maintain a point mass function (PMF) α over the set of possible data associations from measurements at time t to identities (the set of which does not change with time). We will first describe these data associations, and then proceed to define α.

3.1 Representing Data Associations

Data associations are represented by the set of $m \times n$ matrices $\mathbb{A}_{m,n}$ defined by:

$$\mathbb{A}_{m,n} \equiv \begin{cases} \{P \in \{0,1\}^{m\times n} : \forall i, \sum_j p_{i,j} = 1\} & m \le n \\ \{P \in \{0,1\}^{m\times n} : \forall i, \sum_j p_{j,i} = 1\} & m > n \end{cases} \tag{3}$$

where m is the number of measurements, and n is the number of identities. There are several points to note about such matrices:

- $|\mathbb{A}_{m,n}| = \frac{\max(m,n)!}{|m-n|!}$.
- No column or row contains more than one 1, but there are a total of $\min(m, n)$ 1's in the matrix.
- In the special case where $m = n$, then $\mathbb{A}_{n,n}$ is the set of permutation matrices, thus representing $\mathbb{S}_n$.
- For a given time step t, the matrices $P \in \mathbb{A}_{M_t,N}$ are interpreted to represent the possible data associations between measurements and identities. In any such matrix P, row j represents measurement z_t^j, while column i represents identity i (and thus world object $id(i)$). P is defined to associate measurement j with identity i if and only if $P_{j,i} = 1$.

It should be noted that the $M_{t+1} \times M_t$ *association transition matrices* Q mentioned earlier differ from these *data association matrices* $P \in \mathbb{A}_{m,n}$, in that the former are permitted to contain fewer than $\min(m, n)$ 1's. More importantly, the matrices Q represent *transitions between successive measurements*, whereas the matrices $P \in \mathbb{A}_{m,n}$ represent *associations from measurements to identities*.

3.2 Data Association PMFs and Information-Form Matrices

The PMF α represents our beliefs about the true association between measurements and identities. Initially, any uncertainty given by α is due to our lack of prior knowledge about the identities of observed measurements. Mass in α is consolidated when we are given identity readings. Mass in α is diffused when confusion in the tracker occurs (e.g., two tracks of measurements cross), and this confusion is modeled as a distribution over association transitions (i.e. ψ_t).

Analogous to the Gaussian parametrization for continuous random variables, we define a family of PMFs[1] on $\mathbb{A}_{m,n}$ using $n \times m$ real matrices Ω, as follows:

$$\alpha(P; \Omega) \equiv \frac{\exp \operatorname{tr} P\,\Omega}{\sum_{R \in \mathbb{A}_{m,n}} \exp \operatorname{tr} R\,\Omega}, \tag{4}$$

where the denominator normalizes α so that it sums to 1 over $\mathbb{A}_{m,n}$. In the special case of $m = n$, Schumitsch *et al.* termed any such matrix Ω an *information-form matrix*; our definition thus extends their original formulation to general association matrices $\mathbb{A}_{m,n}$. If a random variable matrix P is distributed according to $\alpha(\cdot; \Omega)$, we write $P \sim \mathcal{A}(\Omega)$. The IMKF and CIMBal force the approximation $P \sim \mathcal{A}(\Omega)$, where P is a random variable representing the true data association. All Bayesian filtering operations—namely prediction and update—are performed directly on Ω.

The most likely data association $\hat{P}$ of α is given by

$$\hat{P} = \operatorname*{arg\,max}_{P \in \mathbb{A}_{m,n}} \alpha(P) \tag{5}$$

This can be solved using the *Hungarian algorithm* on Ω (Papadimitriou and Steiglitz 1998), which chooses the list of $\min(n, m)$ elements of the matrix, at most one from each column and each row, such that their sum is maximized. The Hungarian algorithm has runtime $O(\max(n, m)^3)$. For temporally coherent matrices, Mills-Tettey *et al.* (2007) give a dynamic Hungarian algorithm, though with the same worst-case runtime.

In both the IMKF and CIMBal, the information-form matrix Ω is time-dependent; hence the PMF $\alpha(P; \Omega)$ is also time-dependent. We will thus use the shorthand $\alpha_t(P) \equiv \alpha(P; \Omega_t)$, where Ω_t refers to Ω at time t.

4 The Identity Management Kalman Filter (*IMKF*)

The *identity management Kalman filter* (IMKF) described in (Schumitsch, Thrun, Guibas, and Olukotun 2006) is essentially an approximate Bayesian filter on the set of possible data associations. The IMKF assumes a fixed and equal number of measurements and identities, i.e., $N_t = M_t = n$; moreover a one-to-one correspondence between measurements and identities is also assumed (i.e. no "ghost" or duplicate measurements).

[1] When Ω is square, it can be shown that Ω is a linear transformation of a matrix of coefficients; these coefficients come from the Fourier transform of $\ln \alpha$ at the first two representations of the group $\mathbb{S}_n$, where α indicates any distribution over $\mathbb{S}_n$ (not just those parameterized by Ω). This curious property lends the IMKF and CIMBal some connection to the Fourier transform-based tracker of Kondor *et al.* (2007).

Hence, the *true* data association P_t is simply the product of Q_t's (association transition matrices):

$$P_t = Q_t \cdot Q_{t-1} \cdots Q_1$$

The goal, then, is to estimate P_t for each time step. The IMKF estimates P_t using the most likely data association at time t. A naive first step towards determining the latter would be to maintain a distribution over every element of $\mathbb{S}_n$. This, of course, becomes infeasible beyond a small number of targets, since the storage alone is $O(n!)$. Instead, Schumitsch *et al.* assume that $P_t \sim \mathcal{A}(\Omega_t)$, where Ω_t is an $n \times n$ information-form matrix. Therefore, the most likely data association, and thus the estimate of P_t, is simply $\hat{P}_t = \arg\max_{P \in \mathbb{A}_{n,n}} \alpha_t(P)$ as given in the previous section. The IMKF directly modifies the information-form matrix Ω_t by employing certain operations that approximate diffusion and measurement updates. The CIMBal uses variants of these same operations when performing the corresponding updates, the details of which will be given later.

With regards to the continuous state, Schumitsch *et al.* assume that P_t and the x_t^i's are conditionally independent given T_t, thus decoupling the problem of tracking the x_t^i's from estimating P_t. Once the estimate $\hat{P}_t$ has been obtained, the IMKF reduces the continuous tracking problem to that of tracking the state of world objects/identities whose corresponding measurements have been re-ordered according to $\hat{P}_t$. Both the IMKF and the CIMBal employ Kalman filters to perform state tracking of objects, using $\hat{P}_t$ to determine the objects' corresponding measurements.

5 The Conditioning Identity Management Bayesian Filter (*CIMBal*)

5.1 Motivation

Our goal is to improve the IMKF to handle varying numbers of measurements with time, and to reduce computation when only a few identities are of interest to the user. The result is the *Conditioning Identity Management Bayesian Filter* or *CIMBal*, whose purpose is otherwise the same as the IMKF's: to estimate the most likely data association matrix, so as to track the continuous states of some set of world objects using Kalman filters. An overview of the CIMBal is given in Algorithm 1.

In this section, we introduce the key ideas of the CIMBal. The details of the CIMBal's operations will be given in the following section.

5.2 (Potential) Identities of Interest and Relevant Measurements

The biggest difference between the CIMBal and the IMKF is that the CIMBal maintains its data association PMF α_t only over a subset of identities and a subset of measurements, both of which may change with time. This feature allows the CIMBal to reduce computation when only a subset of identities are of interest to the user. Let $\mathcal{I}d_0 \subset [N]$ denote the *identities of interest* whose continuous states the user wishes to track (recall

Algorithm 1 Outline of *CIMBal* filter

procedure *CIMBal*
    Initialize the filter ▷ (see **Initialization**)
    for $t = 1, 2, \ldots$ **do**
        $\Omega_t^- \leftarrow \mathrm{Diffuse}(\Omega_{t-1}^+, \psi_t)$ ▷ (see **Diffusion Step**)
        $\Omega_t^+ \leftarrow \mathrm{Update}(\Omega_t^-, \omega_t)$ ▷ (see **Update Step**)
        $\Omega_t^+ \leftarrow \mathrm{Prune}(\Omega_t^+)$ ▷ (see **Pruning Step**)
        $\hat{P} \leftarrow \mathrm{Associate}(\Omega_t^+)$ ▷ (see **Association Step**)
        Reorder measurements using $\hat{P}$.
        Use reordered Z_t to update the Kalman filters K.
    end for
end procedure

that N is the total number of identities). The states of all other identities are assumed to be of no interest to the user, and are therefore not tracked.

Initially, the CIMBal maintains its data association PMF α_t over $\mathcal{I}d_0$ and the $|\mathcal{I}d_0|$ measurements they correspond to. As time passes however, the CIMBal expands α_t to associate increasingly larger sets of identities with increasingly larger sets of measurements. This is because measurements initially associated with identities in $\mathcal{I}d_0$ may over time cross paths or otherwise become confused with other measurements. This leads to data association ambiguity, in particular ambiguity over which measurements should be used to update the states of $\mathcal{I}d_0$. In order to incorporate this ambiguity into α_t, the CIMBal must "expand" it to include the confounding identities and measurements.

Let us now introduce nomenclature to refer to the identities and measurements maintained by the CIMBal. For a given time t,

- Call the set of identities maintained in α_t the *potential identities of interest* at time t or $\mathcal{I}d_t \subset [N]$.

- Call the set of measurements maintained in α_t the *relevant measurements* at time t or $\mathcal{M}_t \subset [M_t]$ (recall that M_t is the total number of measurements the tracker provides at time t).

There are two restrictions on $\mathcal{I}d_t$ and $\mathcal{M}_t$:

1. $\mathcal{I}d_0 \subset \mathcal{I}d_t$, i.e. the CIMBal always keeps track of the original identities of interest.

2. $|\mathcal{M}_t| \geq |\mathcal{I}d_t|$, i.e. there must be at least as many relevant measurements as potential identities of interest.

These definitions allow us to restate the CIMBal's data association PMF α_t as a *partial data association PMF* between potential identities of interest $\mathcal{I}d_t$ and relevant measurements $\mathcal{M}_t$.

5.3 PMFs over Partial Data Associations

Give that the CIMBal is a tracker over partial data associations, how should its PMF α_t be defined? Recall that α_t is parameterized by an information-form matrix Ω_t. Rather than defining Ω_t as a full $N \times M_t$ matrix — full in the sense that it encodes association information about each of the N targets and M_t measurements — the CIMBal defines Ω_t as a $|\mathcal{I}d_t| \times |\mathcal{M}_t|$ matrix that encodes information only about potential identities of interest and relevant measurements.

To make the relationship between Ω_t and these identities and measurements explicit, we define the bijections $\mathrm{R2I}_t : [|\mathcal{I}d_t|] \mapsto \mathcal{I}d_t$ and $\mathrm{C2M}_t : [|\mathcal{M}_t|] \mapsto \mathcal{M}_t$ that give the correspondences from row indices to potential identities of interest and column indices to relevant measurements, respectively. For example, row i of Ω_t corresponds to potential identity of interest $\mathrm{R2I}_t(i)$, while column j corresponds to relevant measurement $\mathrm{C2M}_t(j)$. According to these definitions, α_t is a PMF on $\mathbb{A}_{|\mathcal{M}_t|,|\mathcal{I}d_t|}$ given by:

$$\alpha_t(P) \equiv \alpha(P;\Omega_t) \equiv \frac{\exp \operatorname{tr} P\,\Omega_t}{\sum_{R\in\mathbb{A}_{|\mathcal{M}_t|,|\mathcal{I}d_t|}} \exp \operatorname{tr} R\,\Omega_t}, \tag{6}$$

where data association matrix P associates relevant measurement $\mathrm{C2M}_t(j)$ with potential identity of interest $\mathrm{R2I}_t(i)$ when $P_{j,i} = 1$. Thus, $\alpha_t(P)$ gives the probability that P is the true partial data association at time t.

For the sake of convenience, we shall write N'_t for $|\mathcal{I}d_t|$ and M'_t for $|\mathcal{M}_t|$ from now on.

5.4 Relationship between Partial Data Association PMFs and Full PMFs

A partial data association PMF $\alpha_t(P;\Omega_t)$ maintained by the CIMBal can be viewed as the result of conditioning some full data association PMF (i.e. over N identities and M_t measurements), where the conditioning is on the event

$$E \equiv \{P \in \mathbb{A}_{M_t,N} \mid \forall i \in \mathcal{I}d_t \forall j \notin \mathcal{M}_t (P_{j,i} = 0) \wedge \forall i \notin \mathcal{I}d_t \forall j \in \mathcal{M}_t (P_{j,i} = 0)\}.$$

In other words, the CIMBal's PMF $\alpha_t(P;\Omega_t)$ is the result of conditioning upon the event[2] that non-potential identities of interest are only associated with non-relevant measurements. To see why this is true, we start by considering the following *conditioning proposition*:

> *Conditioning data associations P on $x_t^i \to z_t^j$ is equivalent to deleting Ω_t's i-th row and j-th column.*

[2]For certain combinations of $N, M_t, |\mathcal{I}d_t|$ and $|\mathcal{M}_t|$, this event E will be empty since no legal data association matrices fit the conditions. As an example, consider $N = 5$, $M_t = 5$, $|\mathcal{I}d_t| = 2$ and $|\mathcal{M}_t| = 3$. This issue can be effectively avoided, by assuming the existence of "fake" identities and measurements that are not in $\mathcal{I}d_t$ and $\mathcal{M}_t$ respectively. For instance, in the previous example, expanding $M_t = 6$ by assuming one fake measurement causes $|E|$ to grow from 0 to $3! \cdot \frac{3!}{(3-2)!} = 36$.

For suppose that $P \sim \mathcal{A}(\Omega_t)$ and $P' \sim \mathcal{A}(\Omega'_t)$, where Ω_t is $N \times M_t$, and Ω'_t equals Ω_t with its i-th row and j-th column removed. Supposing that D_i is such that AD_i removes A's i-th column, then we can write $\Omega'_t = D_i^T \Omega D_j$. Then,

$$\mathrm{pr}(P|x_t^i \to z_t^j) = \begin{cases} \mathrm{pr}(P' = D_j^T P D_i) & \text{if } P_{j,i} = 1 \\ 0 & \text{otherwise} \end{cases} .$$

This implies that the conditional distribution of $P|x_t^i \to z_t^j$ can be exactly represented by a distribution on $\mathbb{A}_{M_t-1,N-1}$, because the values of the former's support are exactly given by the distribution generated by $\mathcal{A}(\Omega'_t)$.

Proof Outline of Conditioning Proposition:

- Let Ω'_t be Ω_t with row i and column j deleted. The resultant PMF over associations in $\mathbb{A}_{M_t-1,N-1}$ is

  $$\alpha'_t(P') \quad \equiv \quad \frac{\exp \mathrm{tr}\, P'\Omega'_t}{\sum_{R' \in \mathbb{A}_{M_t-1,N-1}} \exp \mathrm{tr}\, R'\Omega'_t} .$$

  Comparing this with the definition of $\alpha_t(P)$, it becomes a simple exercise to show that α_t and α'_t are related via

  $$\alpha'_t(D_j^T P D_i) \quad = \quad \frac{1}{\tau}\alpha_t(P)$$

  where $D_i^T A$ is the linear operator removing A's i-th row, so that $D_j^T P D_i$ is equal to P but with row j and column i deleted, and where τ is the marginal probability

  $$\tau \equiv \mathrm{pr}(x_t^i \to z_t^j) = \sum_{P \in \mathbb{A}_{M_t,N} : P_{j,i}=1} \alpha_t(P) .$$

  In other words, deleting row i and column j in Ω_t corresponds to normalizing α_t over all associations that make up the event $x_t^i \to z_t^j$. $\square$

- **Corollary:** when τ is close to 1, $\alpha'_t(D_j^T P D_i) \approx \alpha_t(P)$.

The conditioning proposition can be extended to the more general case where an event of the form E is conditioned upon. This validates the original claim that the CIMBal's PMF α_t is the result of conditioning upon non-potential identities of interest associating only with non-relevant measurements.

Apart from elucidating the connection between partial and full data association PMFs, the conditioning proposition also enables the CIMBal to prune associations $x_t^i \to z_t^j$ when their marginal probabilities become close to 1. This will be discussed further under the ”Pruning Step”.

5.5 Continuous state tracking

Like the IMKF, the CIMBal uses Kalman filters to track the continuous states of world objects corresponding to identities of interest $\mathcal{I}d_0$; we denote this set of Kalman filters by K. Since the measurements' association transitions between timesteps are assumed to be conditionally independent of the world objects' motions (given the tracker state T_t), each identity of interest/world object is dedicated its own Kalman filter, for a total of $|K| = N_0' \equiv |\mathcal{I}d_0|$ filters. At every time step, the most likely data association $\hat{P}_t$ is used to determine each identity's corresponding measurement z_t^j for the Kalman filter update.

6 The *CIMBal* Algorithm

In this section we cover the details of the CIMBal's operations, as listed in Algorithm 1.

6.1 *CIMBal* Initialization

As input, the CIMBal requires $\mathcal{I}d_0$ as well as the $M_0' = N_0'$ true corresponding initial measurements $\mathcal{M}_0$. There are many ways to initialize Ω_0, $\mathrm{R2I}_0$ and $\mathrm{C2M}_0$, but for concreteness we shall define them as such:

- $\mathrm{R2I}_0$ is a monotonically increasing function, i.e. increasing row numbers correspond to higher-numbered identities in id.

- $\mathrm{C2M}_0$ is such that identity $\mathrm{R2I}_0(i)$ truly corresponds to measurement $\mathrm{C2M}_0(i)$ for all $1 \leq i \leq N_0'$.

- The above definitions require the true data association at time 0 to be the $N_0' \times N_0'$ identity matrix $I_{N_0' \times N_0'}$. This in turn requires that we initialize Ω_0 to be:

$$\Omega_0 \quad = \quad c \cdot I_{N_0' \times N_0'} \tag{7}$$

  where c is a parameter controlling the degree of certainty that $I_{N_0' \times N_0'}$ is the true data association. Increasing c will increase the probability $\alpha_0(\hat{P}_0 = I_{N_0' \times N_0'})$, recalling that $\hat{P}_t$ is the most likely data association matrix at time t.

6.2 Diffusion Step

The purpose of the diffusion step is to model uncertainty introduced by the tracker's potential confusion of measurements. We have assumed that the tracker provides a distribution ψ_t, that encodes whatever ambiguities the tracker is aware of. The exact diffusion is given by the convolution of α_t with ψ_t; however this is computationally intractable for large Ω_t. Rather than perform the convolution, the IMKF approximates it via the following equation:

$$\Omega_{t+1}^{-} = \ln \left[\left(\exp \Omega_t^{+} \right) \sum_{Q \in \psi_t} \left(\psi_t(Q) Q^T \right) \right], \tag{8}$$

where $\exp$ and $\ln$ are computed element-wise, and where the notation $Q \in \psi_t$ denotes the set of association transition matrices[3] in the support of ψ_t. The notations Ω_{t+1}^{-} and Ω_t^{+} denote the prior and posterior values of Ω_{t+1} and Ω_t respectively; the distinction will be explained under the Update Step.

The CIMBal also uses (8) to perform diffusion. This poses some problems because the PMF α_t maintained by the CIMBal is already conditioned on a certain set of data associations; furthermore the column indices of Ω_t only indirectly correspond[4] to measurements through $\mathrm{C2M}_t$. We resolve these issues via the following guidelines:

1. After distributing $\exp \Omega_t^{+}$ over the summation, we see that the Q^T in each term of the summation acts to reorder or discard measurement identity information (represented by columns) in $(\exp \Omega_t^{+})Q^T$. This is because the $M'_t \times M'_{t+1}$ inverse transition association matrices Q^T have either one or zero 1's per row.

2. More specifically, $Q^T_{i,j} = 1$ means that Q^T moves column $\mathrm{C2M}_t^{-1}(i)$ of $\exp \Omega_t^{+}$ to column $\mathrm{C2M}_{t+1}^{-1}(j)$ of $(\exp \Omega_t^{+})Q^T$. On the other hand, row i of Q^T being zero means that column $\mathrm{C2M}_t^{-1}(i)$ does not appear in $(\exp \Omega_t^{+})Q^T$.

3. In cases where $Q^T_{i,j} = 1$ but $\mathrm{C2M}_t^{-1}(i)$ is undefined — i.e. measurement i is not represented by any column in Ω_t — we assume that measurement i is implicitly represented by a zero column, representing our lack of identity information. Because zero columns contribute nothing to the final summation regardless of how they are reordered, we can simply ignore references to such columns, which in turn speeds up computations.

4. If any column of the summation $\exp \Omega_{t+1}^{-}$ turns out to be zero, we simply delete it from $\exp \Omega_{t+1}^{-}$. The justification for this is twofold: First, the corresponding measurement is highly unlikely to be truly associated with any of the potential identities of interest, so its removal does not change the PMF α_t significantly. Second, its removal helps to keep Ω_{t+1}^{-} compact, which in turn improves runtime.

Together, these guidelines suggest that we compute (8) sparsely by summing over individual contributions to each column of $\exp \Omega_{t+1}^{-}$, as opposed to naively performing the summation over matrix products $(\exp \Omega_t^{+})Q^T$. The only remaining issues are how to determine the number of columns M'_{t+1} in $\exp \Omega_{t+1}^{+}$, as well as their mappings $\mathrm{C2M}_{t+1}$ to corresponding measurements. Notice that

$$M'_{t+1} \quad = \quad \left| \left\{ j \mid \exists i \exists Q \in \psi_t (Q^T_{i,j} = 1 \wedge i \in \mathrm{RangeOf}(\mathrm{C2M}_t)) \right\} \right| . \tag{9}$$

[3] The IMKF only allows association transition matrices $Q \in \mathbb{S}_n$, i.e. each measurement at time t must be associated with some measurement at time $t+1$. This is in contrast to the CIMBal, which allows its $M_{t+1} \times M_t$ matrices Q to lose track of old measurements or introduce new measurements.

[4] It should be noted that (8) does not change the number of rows in Ω_t, nor their correspondence $\mathrm{R2I}_t$ with identities.

In words, M'_{t+1} is equal to the number of measurement indices j at time $t+1$ that are associated with a measurement in Ω_t through some $Q \in \psi_t$. Once M'_{t+1} has been determined, the exact mapping from columns to measurements $\mathrm{C2M}_{t+1}$ is left up to implementation — for example, one might require $\mathrm{C2M}_{t+1}$ to be monotonically increasing, giving only one possible column ordering of Ω^-_{t+1}.

With these details in mind, (8) can be implemented with runtime $\mathrm{O}(pM'_tN'_t)$, where $p = |Q \in \psi_t|$. If we assume that $p = \mathrm{O}(z)$ where $z = \max(M'_t, N'_t)$, then the Diffusion Step runs in $\mathrm{O}(z^3)$. We believe this assumption to be reasonable because for most real-world tracking scenarios, the proportion of measurements that confuse with each other in a given time step is expected to be small.

6.3 Update Step

In the update step, the CIMBal (and IMKF) uses Bayes' rule to update its information-form matrix Ω_t with new identity readings, which are provided by the tracker as a matrix ω_t. Recall that identity readings are information on measurement to identity associations, as opposed to transition association information ψ_t relating measurements at different time steps.

More specifically, the goal of the Update Step is to use ω_t as evidence to update the prior Ω^-_t (from the Diffusion Step) to the posterior Ω^+_t. We require ω_t to be an information-form matrix, so that the identity readings are encoded in the distribution $\mathcal{A}(\omega_t)$. Let P_t denote the event that $P \in \mathbb{A}_{M'_t,N'_t}$ is the true data association at time t before considering the evidence in ω_t, i.e. $\mathrm{pr}(P_t) = \alpha(P;\Omega^-_t)$. By defining

$$\mathrm{pr}(\omega_t|P_t) \;\equiv\; \alpha(P;\omega_t)\;,$$

it follows that the posterior of Ω_t given evidence ω_t is

$$\Omega^+_t \;=\; \Omega^-_t + \omega_t \tag{10}$$

Proof of Update Step:

- We need to show that $\mathrm{pr}(P_t|\omega_t) = \alpha(P;\Omega^+_t)$. Using Bayes' Rule,

$$\begin{aligned}
\mathrm{pr}(P_t|\omega_t) &= \frac{\mathrm{pr}(\omega_t|P_t)\mathrm{pr}(P_t)}{\sum_{R\in\mathbb{A}_{M'_t,N'_t}} \mathrm{pr}(\omega_t|R_t)\mathrm{pr}(R_t)} \\
&= \frac{\alpha(P;\omega_t)\alpha(P;\Omega^-_t)}{\sum_{R\in\mathbb{A}_{M'_t,N'_t}} \alpha(R;\omega_t)\alpha(R;\Omega^-_t)} \\
&= \frac{(\exp\operatorname{tr} P\omega_t)(\exp\operatorname{tr} P\Omega^-_t)}{\sum_{R\in\mathbb{A}_{M'_t,N'_t}} (\exp\operatorname{tr} R\omega_t)(\exp\operatorname{tr} R\Omega^-_t)} \\
&= \frac{\exp\operatorname{tr} P(\omega_t + \Omega^-_t)}{\sum_{R\in\mathbb{A}_{M'_t,N'_t}} \exp\operatorname{tr} R(\omega_t + \Omega^-_t)} \\
&= \alpha(P;\Omega^-_t + \omega_t) \;=\; \alpha(P;\Omega^+_t) \quad \square
\end{aligned}$$

6.3.1 Constructing ω_t

In general, the tracker will not provide identity readings as information-form matrices ω_t. We therefore need to construct ω_t from whatever information is provided. Let us suppose that the tracker's sensing model produces marginal probabilities of the form $\mathrm{pr}(x_t^i \to z_t^j | T_t)$, where $i \in \mathcal{I}d_t$ and $j \in \mathcal{M}_t$. Supposing that for some identity i,

$$\mathrm{pr}(x_t^i \to z_t^j | T_t) = \begin{cases} \gamma & j = i \\ (1-\gamma)/(M_t' - 1) & j \neq i \end{cases} \tag{11}$$

then we can easily construct a satisfying ω_t by making row i equal $\ln \mathrm{pr}(x_t^i \to z_t^j | T_t)$ as defined above, and setting all other rows to zero. Furthermore, it is trivial to show that distributions $\mathcal{A}(\omega_t)$ are invariant to the addition of a scalar to any row of ω_t, provided $M_t' \geq N_t'$.[5] Since the CIMBal requires $M_t' \geq N_t'$ for all t, we may therefore subtract $\ln(1-\gamma)/(M_t' - 1)$ from the ith row to obtain:

$$\omega_t^{i,j} = \begin{cases} \ln \gamma (M_t' - 1)/(1-\gamma) & i = j \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

In the case where marginals $\mathrm{pr}(x_t^i \to z_t^j | T_t)$ are provided for multiple identities i, one may generate a unique ω_i for each identity as described above, then sum the ω_i's to get ω_t. This is equivalent to performing the Update Step on each ω_i in succession. Using this formulation, the Update Step has a runtime upper bound of $\mathrm{O}(N_t')$.

6.3.2 Caveats

The aforementioned construction does not cover situations where marginal probabilities $\mathrm{pr}(x_t^i \to z_t^j | T_t)$ such that

1. $i \in \mathcal{I}d_t$ and $j \notin \mathcal{M}_t$, or
2. $i \notin \mathcal{I}d_t$ and $j \in \mathcal{M}_t$, or
3. $i \notin \mathcal{I}d_t$ and $j \notin \mathcal{M}_t$

are given. In all 3 cases, the given marginal references entries outside of Ω_t^-. One solution is to expand Ω_t^- to include a new zero row and/or column corresponding to the identity and measurement in the marginal, similar to what was done in the Diffusion Step for references to measurements outside $\mathcal{M}_t$. This solution is premised on the notion that zero entries of Ω_t loosely indicate a lack of information regarding the corresponding measurement-to-identity associations. The overall, qualitative effect is to diffuse the probability mass in α_t over the newly-introduced measurements and targets. The precise effects, however, vary depending on the current values of Ω_t^-.

[5] If $N_t' \geq M_t'$, the distributions are instead invariant to the addition of a scalar to any column of ω_t.

6.4 Pruning Step

The CIMBal uses the *conditioning proposition* to prune highly likely data associations from Ω_t, reducing its size and hence the runtime of all other CIMBal operations. This is the key operation that allows the CIMBal substantial runtime and space savings over the IMKF.

More specifically, the CIMBal checks for potential identities of interest $i \in \mathcal{I}d_t$, $i \notin \mathcal{I}d_0$ and relevant measurements $j \in \mathcal{M}_t$ such that $\mathrm{pr}(x_t^i \rightarrow z_t^j) \approx 1$, deleting row $\mathrm{R2I}_t^{-1}(i)$ and column $\mathrm{C2M}_t^{-1}(j)$ when this condition has been met. Recall that such a deletion conditions α_t on the event $x_t^i \rightarrow z_t^j$; this produces a good approximation to the unconditioned α_t when $\mathrm{pr}(x_t^i \rightarrow z_t^j) \approx 1$. When there are multiple pruning candidates i and j, the CIMBal prunes one pair (i, j) at a time until there are no longer any candidate pairs.

6.4.1 Identifying Pruning Candidates

Identifying $(i, j) : \mathrm{pr}(x_t^i \rightarrow z_t^j) \approx 1$ is normally prohibitive because of the factorial complexity of computing such marginals. In order to overcome this, the CIMBal selects (i, j) pairs in two steps. The First Step uses a fast $\mathrm{O}(N_t' M_t')$ approximation to compute all $N_t' M_t'$ data associations marginals in Ω_t. In the Second Step, the highest marginals (up to a maximum of H) from the first step that

1. Do not involve $i \in \mathcal{I}d_0$
2. Are not significantly lower than some pruning threshold κ

are re-estimated using the Metropolis-Hastings algorithm (Robert and Casella 2005) in $\mathrm{O}(z^2)$, where $z = \max(N_t', M_t')$. For each marginal $\mathrm{pr}(x_t^i \rightarrow z_t^j)$ re-estimated to be greater than κ, the CIMBal deletes row $\mathrm{R2I}_t^{-1}(i)$ and column $\mathrm{C2M}_t^{-1}(j)$ from Ω_t, while updating $\mathrm{R2I}_t$ and $\mathrm{C2M}_t$ appropriately.

6.4.2 First Step: Fast Approximation

The first approximation produces an $N_t' \times M_t'$ matrix A, where element $A_{i,j}$ approximates $\mathrm{pr}(x_t^i \rightarrow z_t^j)$. A is generated in $\mathrm{O}(N_t' M_t')$ time via the following equations:

$$
\begin{aligned}
&A = \tfrac{R'+C'}{2} \\
&R'_{x,y} = \tfrac{R_{x,y}}{\sum_a R_{a,y}} \quad R_{x,y} = \tfrac{\exp \Omega_t^{x,y}}{\sum_b \exp \Omega_t^{x,b}} \\
&C'_{x,y} = \tfrac{C_{x,y}}{\sum_b C_{x,b}} \quad C_{x,y} = \tfrac{\exp \Omega_t^{x,y}}{\sum_a \exp \Omega_t^{a,y}}
\end{aligned}
\tag{13}
$$

where R', C', R and C are also $N_t' \times M_t'$ matrices, and exp denotes element-wise matrix exponentiation. In short, $\exp \Omega_t$ is normalized across columns and then across rows to get C'. This normalization is also performed across rows and then across columns to get R'. Finally, R' and C' are averaged to get A.

In general, the matrix A will not be doubly stochastic. Experimentation on 10×10 random matrices Ω shows that for high cutoffs $\kappa \geq 0.9$, while the Fast Approximation generates a sizeable number of false negatives, it generates very few false positives. This behavior is illustrated in Figure 1.

6.4.3 Second Step: Metropolis-Hastings Approximation

Observe that relative probabilities $\alpha_t(P_1)/\alpha_t(P_2)$ can be computed in $\mathrm{O}(\min(N_t', M_t'))$:

$$\frac{\alpha_t(P_1)}{\alpha_t(P_2)} = \frac{\exp \operatorname{tr} P_1 \Omega_t}{\exp \operatorname{tr} P_2 \Omega_t} \tag{14}$$

This allows us to use the Metropolis-Hastings algorithm to accurately re-estimate candidate marginals $\mathrm{pr}(x_t^i \to z_t^j)$ provided by the First Step. Moreover, when $\mathrm{pr}(x_t^i \to z_t^j)$ is being re-estimated, a proposal density favoring data associations $P : P_{j,i} = 1$ should be used. This is because we expect $\mathrm{pr}(x_t^i \to z_t^j)$ to be close to 1 with high probability, assuming that the pruning threshold κ is close to 1 while candidate marginals coming from the First Step were not estimated to be significantly lower than κ. When Ω_t is square, one such proposal density can be implemented by generating a random data association[6] P uniformly, and then:

1. If $P(i) \neq j$ (i.e. $P_{j,i} \neq 1$), with probability p swap $P(i)$ with $P(k) = j$.

2. If $P(i) = j$, with probability $1 - p$ swap $P(i)$ with a randomly chosen element $P(k) : k \neq i$.

As these swaps can be performed in constant time, the use of this proposal density does not affect the asymptotic complexity of the Second Step. Other proposal densities also exist, and their use is encouraged should they better fit the expected distribution of marginals from the First Step.

If $\mathrm{O}(z)$ iterations per marginal (where $z = \max(N_t', M_t')$) are used to estimate only a constant number of marginals, then the Metropolis-Hastings step can performed in $\mathrm{O}(z^2)$, which is still faster than the limiting $\mathrm{O}(z^3)$ runtime of the Hungarian algorithm required by the Association Step. Moreover, up to $\mathrm{O}(z^2)$ iterations per marginal may be used without exceeding $\mathrm{O}(z^3)$ runtime; alternatively one could choose to re-estimate up to $\mathrm{O}(z)$ marginals.

In our testing, we used $200z$ iterations per marginal to re-estimate up to $H = 5$ marginals. However, it should be noted that we inadvertently used an inefficient $\mathrm{O}(z \log z)$ means of generating data associations, causing our implementation of the Second Step to have $\mathrm{O}(z^2 \log z)$ runtime. Coupled with our observation that the Second Step accounts for a large fraction of the total CIMBal runtime, we therefore believe the CIMBal can perform better than our results would indicate.

[6] Note that a random data association $P \in \mathbb{A}_{M_t', N_t'}$ can be uniformly generated in $\mathrm{O}(z)$ time, by generating a permutation on $\mathbb{S}_z$ and then reading only the first $\min(N_t', M_t')$ elements.

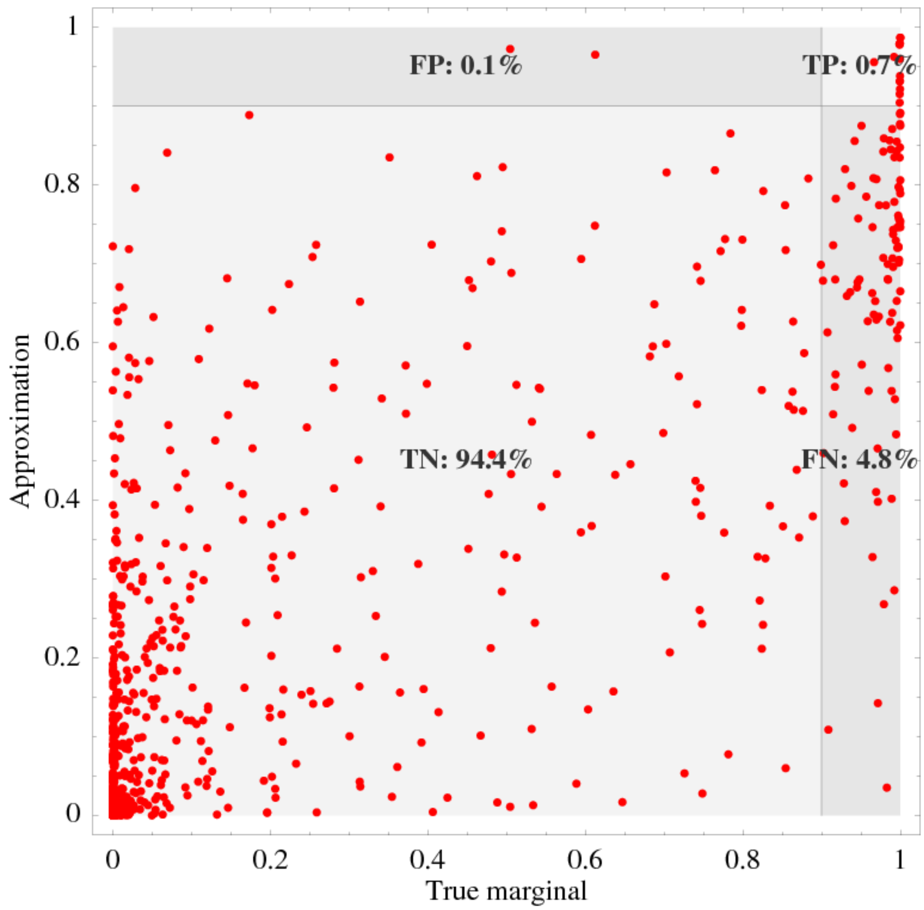

Figure 1: True marginals vs. fast approximation marginals for 20 randomly generated 10×10 Ω's, with each element distributed according to $\mathrm{U}(0, 25)$. Given a cutoff of $\kappa = 0.9$, the percentages of true/false postives/negatives are given in each case. The ratio of FP's to TP's to FN's is approximately $1 : 7 : 49$, indicating that the approximation is more often too conservative, but only 1 in 7 are false positives.

6.5 Association Step

The most likely data association $\hat{P}_t$ is required to update the $|\mathcal{I}d_0|$ Kalman filters K. This was given by (5), which can be solved in $\mathrm{O}(\max(N_t', M_t')^3)$ using the Hungarian max-matching algorithm on Ω_t (Papadimitriou and Steiglitz 1998), according to

$$\left[\hat{P}_t^T\right]_{i,j} = \begin{cases} 1 & \text{Hungarian chooses } [\Omega_t]_{i,j} \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

where $\hat{P}_t^T$ is the transpose of $\hat{P}_t$. Then, for each identity of interest $i \in \mathcal{I}d_0$, its corresponding measurement under the most likely data association is z_t^j where j : $\hat{P}_t^{\mathrm{C2M}_t^{-1}(j),\mathrm{R2I}_t^{-1}(i)} = 1$. Once determined, the $|\mathcal{I}d_0|$ corresponding measurements are used to update the Kalman filters. If few changes in Ω_t are expected between timesteps, then the dynamic Hungarian algorithm (Mills-Tettey, Stentz, and Dias 2007) should be used instead, as the former may be able to solve (5) in less than $\mathrm{O}(\max(N_t, M_t)^3)$.

7 Experiments

Assuming that the support of ψ_t has size $\mathrm{O}(z)$ (where $z = \max(M_t', N_t')$), then the CIMBal requires $\mathrm{O}(z^3)$ to run through time step t. When this is compared to the IMKF's per-time step runtime of $\mathrm{O}(N^3)$, we see that CIMBal has the potential to run much faster, especially when the Pruning Step is successful in keeping $z \ll N$ for most time steps. We support this claim by testing the CIMBal against the IMKF on the same data set used by Schumitsch *et al.* (2006). We will also see that the CIMBal suffers no significant loss of tracking accuracy, despite running an entire order of magnitude faster compared to the IMKF.

7.1 IMKF Data Set

7.1.1 Test Setup

We tested the CIMBal on the data used in Schumitsch *et al.* (2006) (provided courtesy of Schumitsch *et al.*), which was generated via a DARPA-designed program for use in the DARPA ACIP. The generated data contains over 3,000 objects moving through an urban environment in a realistic fashion. $N = 2412$ of these objects moving over 678 time steps were selected to make up the world; every one of these objects is labeled with a unique identity in $[N]$, and Figure 2 gives their tracks. The number of measurements generated was fixed at $M_t = N$. We selected 30 objects at random to be the *identities of interest* $\mathcal{I}d_0$.

We ran the CIMBal and IMKF in a series of trials varying over two parameters:

1. The Frequency of Identity inforMation (FIM) per measurement per timestep at $FIM = 0.100$, 0.033, 0.010. This represents the rate at which identity readings are provided by a tracker. More specifically, FIM is the probability for each measurement that the tracker will provide the true corresponding identity, at any timestep. For example, $FIM = 0.100$ means that at every time step, there is a 0.1 probability per measurement to get an identity reading for that measurement.

2. The pruning threshold κ at $\kappa = 0.99,\ 0.9$.

Moreover, in each trial the IMKF was run once, while the CIMBal was run in two setups:

1. 30T setup: The CIMBal was run once with $\mathcal{I}d_0$ as described earlier.
2. 1T setup: The CIMBal was run 30 times, each time with a different $i \in \mathcal{I}d_0$ as the single *identity of interest*.

The details of our Metropolis-Hastings implementation are given in the Pruning Step, Second Step.

Accuracies for the IMKF and 30T CIMBal setups are given as the mean fraction of timesteps where identity i's true corresponding measurement was correctly identified, over all 30 $i \in \mathcal{I}d_0$. The accuracies for 1T CIMBal setups are given as the mean fraction of timesteps where the single identity's true corresopnding measurement was correctly identified, over all 30 runs. **Runtimes shown are for the IMKF/CIMBal only** — they do not include low-level tracker runtime, which includes the time required to generate ψ_t and ω_t. As the same low-level tracker is used for both the IMKF and CIMBal, its contribution to overall runtime is the same for both filters. Also, note that the runtime for 1T CIMBal setups is given as the *mean* over all 30 runs, not the sum.

7.1.2 Results and Discussion

Our results show that the CIMBal is as accurate as the IMKF at tracking the identities of interest $\mathcal{I}d_0$ (Figure 3). Compared to the IMKF, the 30T CIMBal setup runs significantly faster; in particular Figure 4 shows that even under sparse identity readings $FIM = 0.01$, the 30T CIMBal is *an order of magnitude faster* than the IMKF, when low-level tracker runtime has been excluded from both. Increasing FIM to 0.1 makes the 30T CIMBal another 5 times faster.

We have noticed one drawback to the CIMBal: under low $FIM = 0.01$, runtime per time step (iteration time) and $|\Omega_t|$ increase over time for both the 1T and 30T setups. However, with high $FIM = 0.1$ these quantities remain more or less constant over time (Figures 5, 6). Figure 7 shows that CIMBal runtime scales according to $\mathrm{O}(z^3)$: when iteration time is plotted against z on a log-log scale, the slope of any of the trends — where each trend represents a specific number of Metropolis-Hastings re-estimation candidates — does not exceed 3.

Additionally, we have observed that the Metropolis-Hastings algorithm accounts for a significant fraction of the CIMBal's runtime. This is evidenced by the large gaps between the trends in Figure 7; notice how going from 1 to 3 re-estimation candidates doubles the runtime. This, coupled with the fact that H is only a maximum on re-estimation candidates, suggests that CIMBal runtime can fluctuate significantly from one time step to the next — a trend that is clearly observed in Figure 5.

8 Conclusion

We have devised an IMKF variant we call the CIMBal, with several major advantages over its predecessor:

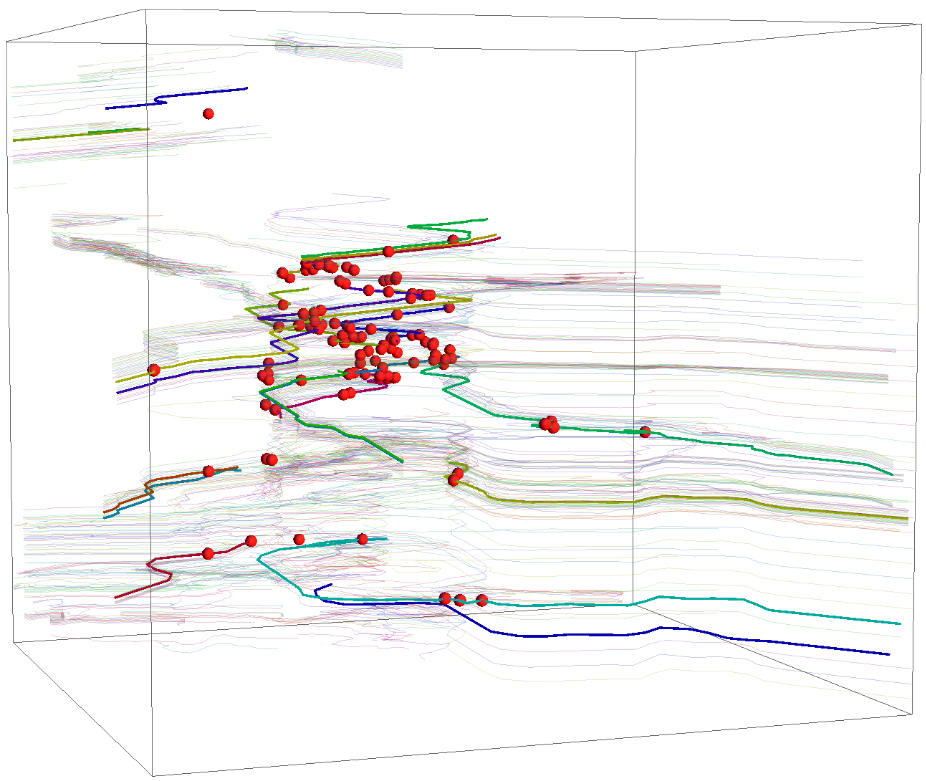

Figure 2: Object tracks in the test data. The vertical axis represents time. Thick opaque lines are identities of interest. Transparent lines are other targets. Spheres represent events when two nearby tracks become confused.

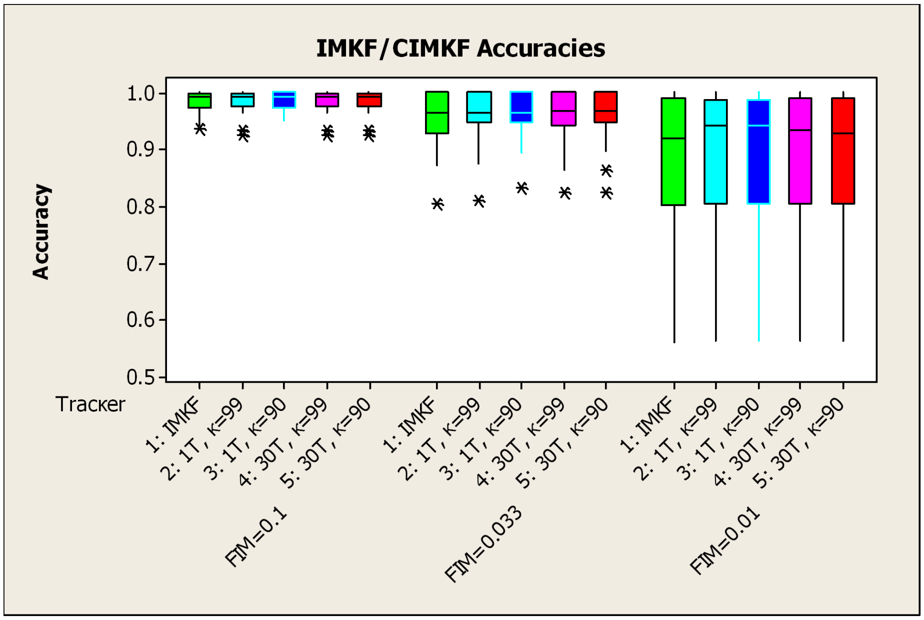

Figure 3: IMKF and CIMBal accuracies. 1T and 30T denote the single and 30 *identities of interest* CIMBal setups respectively.

1. The CIMBal runs much faster and uses less space than the IMKF, in situations where only a small subset of objects/identities need to be tracked.

2. The CIMBal can handle changing numbers M_t of measurements with time.

3. The CIMBal does not require $M_t = N$ (where N is the number of identities).

4. The CIMBal does not even require N to be known, in the sense that no CIMBal operation requires the value of N. This has the consequence that the low-level tracker may increase N at any time, provided that existing identity-to-object mappings are left unchanged. Decreasing N is not permitted, however.

Key to the CIMBal's performance is the Pruning Step, which removes high-probability identity-to-measurement associations, thus decreasing the size of Ω_t with minimal loss in accuracy. We expect the CIMBal to be of utility in applications requiring its advantages, such as aerial surveillance and camera networks.

9 Further Research

There are a few avenues for further improvement. First, the Fast Approximation might be improved to be more statistically sound, allowing the CIMBal to propose better

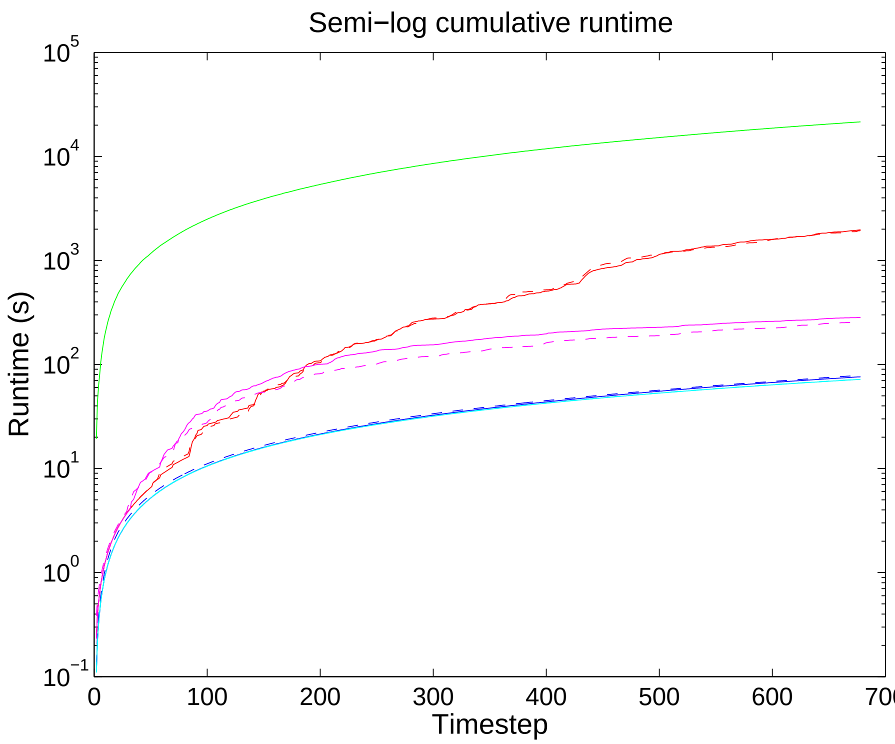

Figure 4: Semi-log plot of cumulative runtime. Color codes — Green: IMKF with $FIM = 0.01$, Magenta/Red: 30T CIMBal with $FIM = 0.1$ (magenta) and 0.01 (red), Cyan/Blue: 1T CIMBal average with $FIM = 0.1$ (cyan) and 0.01 (blue). Line styles: Solid: $\kappa = 0.99$, Dashed: $\kappa = 0.90$.

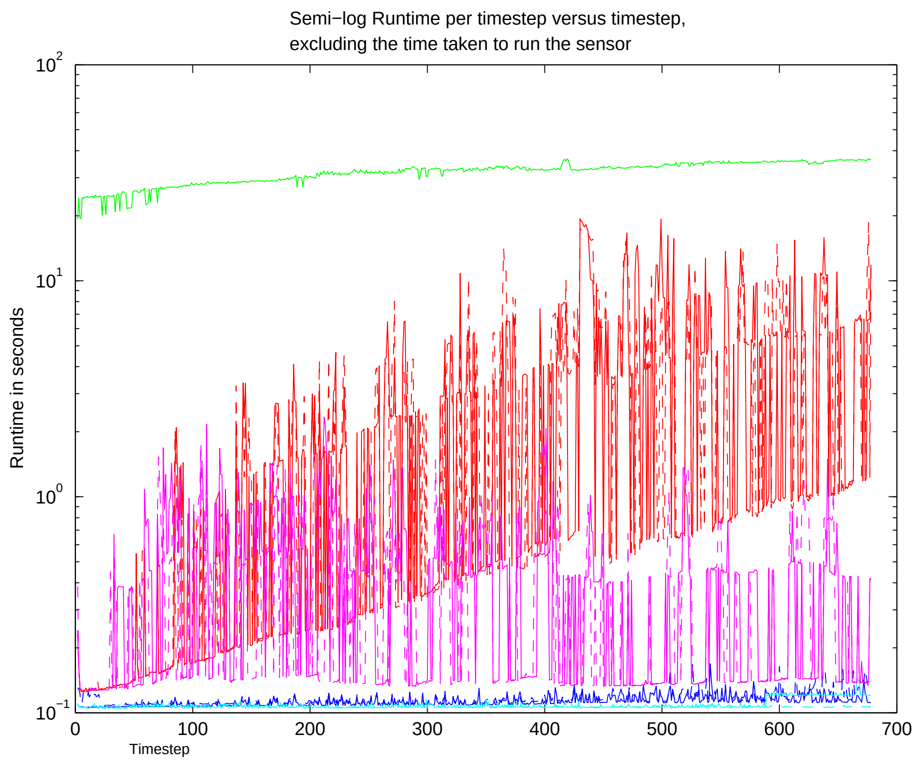

Figure 5: Semi-log plot of runtime for each time step. Color codes and line styles are the same as in Figure 4.

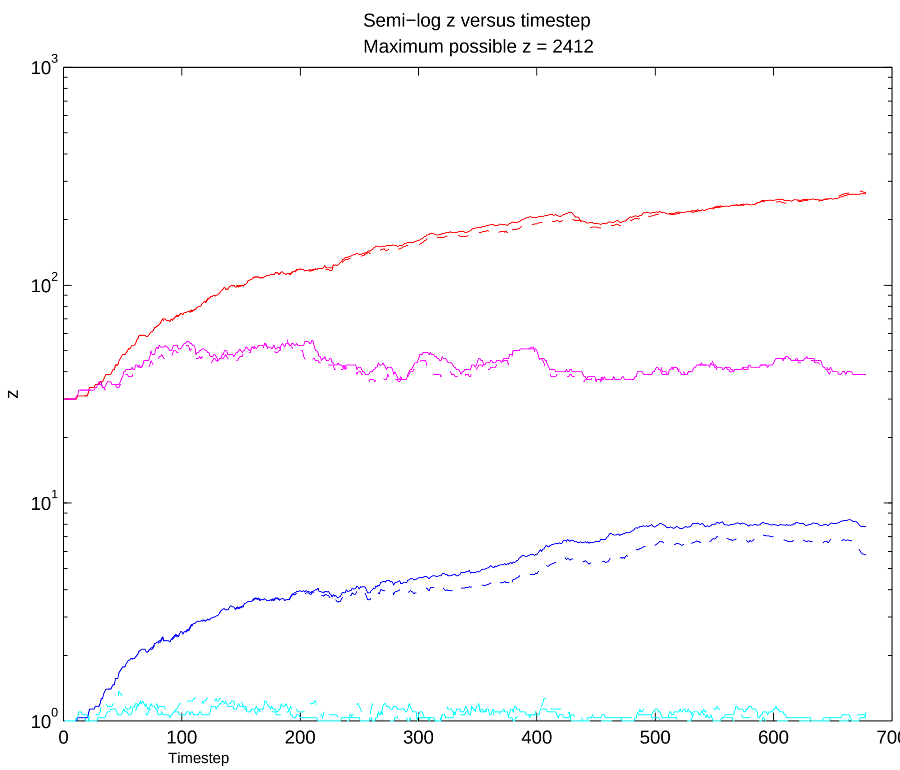

Figure 6: Semi-log plot of z over time, where $z = \max(N_t', M_t')$. Color codes and line styles are the same as in Figure 4.

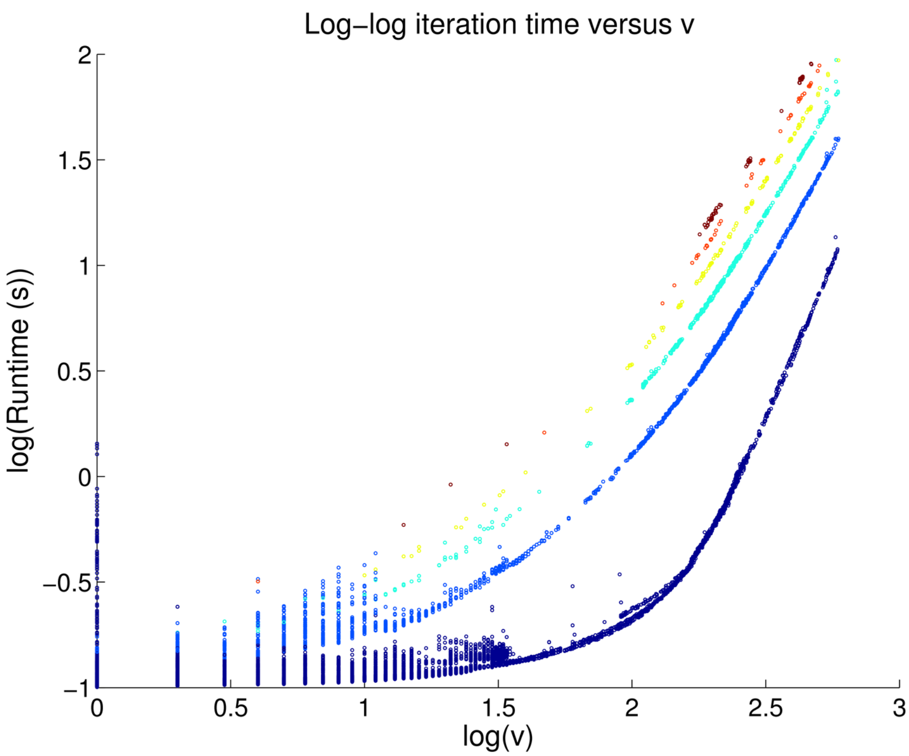

Figure 7: Log-log plot of iteration time versus v, where $v = \max(N'_t, M'_t)$. Each point represents a particular time step in either a 30T or 1T CIMBal setup, for the parameters $FIM = 0.01$ and $\kappa = 0.99$ or 0.9. Point colors indicate the number of candidates re-estimated using Metropolis-Hastings during that iteration, from 0 (dark blue) to 5 (dark red). Observe that the slope of any colored trend does not exceed 3, implying that the CIMBal achieves $\mathrm{O}(v^3)$ runtime in practice.

re-estimation candidates $\mathrm{pr}(x_t^i \rightarrow z_t^j)$ to the more accurate but slower Metropolis-Hastings step. In particular, the Fast Approximation's high rate of false negatives negatively affects the overall performance of the CIMBal. This is because certain pruning candidates might be consistently missed, thus keeping $|\Omega_t|$ larger than it needs to be. Ideally, we desire a Monte Carlo method that quickly estimates high probability marginal associations $\mathrm{pr}(x_t^i \rightarrow z_t^j)$, where by "quickly" we mean faster than the total time required by the Metropolis-Hastings step.

Second, the Metropolis-Hastings algorithm need not be run as often as it is now. It is likely that any marginals $\mathrm{pr}(x^i \rightarrow z^j)$ that are estimated by the Metropolis-Hastings algorithm during some timestep t_k, will be estimated again in future timesteps t_{k+r} for $r \leq$ some horizon h; this is especially true when identity information is scarce, since the marginals do not change as dramatically by the diffusion step as they do by the update step. Hence for a given marginal, one can apply the importance sampling technique (Liu 2001) on the sample sequence generated at t_k to estimate the new value of the marginal at t_{k+r}. With respect to the CIMBal, we can implement importance sampling during Metropolis-Hastings by taking a previous sample sequence, reweighing each sampled permutation P_i by $\alpha(P_i; \Omega_{t_{k+r}})/\alpha(P_i; \Omega_{t_k})$, and then using this reweighed sequence to re-estimate $\mathrm{pr}(x^i \rightarrow z^j)$. This requires a constant fraction fewer computations than generating proposal permutations and calculating acceptance probabilities normally, though the asymptotic complexity still remains $\mathrm{O}(z)$ per sample. The primary drawback to importance sampling is that additional space is required to store Ω_t's from previous timesteps, though we expect the improvement in runtime to be worth the extra space.

Finally, while the effects of deleting rows and columns from Ω are explained by the conditioning proposition, adding rows and columns is only understood to diffuse the mass of the distribution α over the new measurement-to-identity associations made possible; furthermore the diffusion is highly dependent on the existing entries of Ω. A formula or approximation for the new rows/columns that results in a more precise redistribution of the probability mass would be an important step forward.

10 Acknowledgements

We would like to thank Brad Schumitsch for providing us with his dataset, and we would also like to thank Jonathan Huang for his helpful discussions.

References

Betke, M., D. E. Hirsh, A. Bagchi, N. I. Hristov, N. C. Makris, and T. H. Kunz (2007). Tracking large variable numbers of objects in clutter. In *CVPR*. IEEE Computer Society.

Dellaert, F., S. Seitz, C. Thorpe, and S. Thrun (2007). Em, mcmc, and chain flipping for structure from motion with unknown correspondence. *Machine Learning 50*(1-2), 45–71.

Fortmann, T., Y. Bar-Shalom, and M. Scheffe (1980, December). Multi-target tracking using joint probabilistic data association. In *Proc. IEEE Conference on Decision and Control*, pp. 807 – 812.

Huang, J., C. Guestrin, and L. Guibas (2007). Efficient inference for distributions on permutations. In *Proc. NIPS 2007*.

Kondor, R., A. Howard, and T. Jebara (2007, March). Multi-object tracking with representations of the symmetric group. In *Artificial Intelligence and Statistics, AISTATS*, San Juan, Puerto Rico.

Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag.

Mills-Tettey, G. A., A. T. Stentz, and M. B. Dias (2007, July). The dynamic hungarian algorithm for the assignment problem with changing costs. Technical Report CMU-RI-TR-07-27, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Oh, S. and S. Sastry (2005, June). A polynomial-time approximation algorithm for joint probabilistic data association. In *Proc. American Control Conference (ACC)*, Portland, OR, USA.

Page, L. (2007, October). Boeing robo-chopper for darpa's super-spyeye. *The Register*.

Papadimitriou, C. H. and K. Steiglitz (1998, July). *Combinatorial Optimization : Algorithms and Complexity*. Dover Publications.

Pasula, H., S. J. Russell, M. Ostland, and Y. Ritov (1999). Tracking many objects with many sensors. In *IJCAI*, Stockholm, Sweden, pp. 1160–1171.

Reid, D. (1979). An algorithm for tracking multiple targets. *IEEE Transactions on Aerospace and Electronic Systems AC-24*, 843–854.

Robert, C. P. and G. Casella (2005, July). *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer.

Schumitsch, B., S. Thrun, L. Guibas, and K. Olukotun (2006, August). The identity management kalman filter (imkf). In *Proceedings of Robotics: Science and Systems*, Philadelphia, PA, USA.

Shin, J., L. J. Guibas, and F. Zhao (2003). A distributed algorithm for managing multi-target identities in wireless ad-hoc sensor networks. In F. Zhao and L. J. Guibas (Eds.), *IPSN*, Volume 2634 of *Lecture Notes in Computer Science*, pp. 223–238. Springer.