An Efficient Algorithm for Tracking Multiple Maneuvering Targets

Songhwai Oh and Shankar Sastry

Abstract— **Tracking multiple maneuvering targets in a cluttered environment is a challenging problem. A combination of interacting multiple model (IMM) and joint probabilistic data association (JPDA) has been successfully applied to track multiple maneuvering targets. In IMM, the motion of a maneuvering target is approximated by a finite number of simple, distinct kinematic models. However, the exact computation of the combined approach has the time complexity which is exponential in the numbers of kinematic models and measurements. When applying JPDA and IMM, the numbers of targets and kinematic models are known, so we can design a tracking system suitable for the given numbers of targets and kinematic models. But the number of measurements is not known in advance, and it poses a serious problem in computing association probabilities in JPDA. Hence, for a large problem, we need to seek for an efficient approximation algorithm. In this paper, we present a randomized algorithm which finds approximations of association probabilities with good fidelity and prove that the time complexity of the algorithm is polynomial in the size of the problem.**

I. Introduction

The data association problem arises in many applications such as computer vision, surveillance, clustering, and mobile robots. In computer vision, the data association problem is known as the correspondence problem in which the objective is to determine which observation belongs to which feature [1], [2]. In target tracking, it is the problem of determining which observation is generated by which target or clutter [3].

Tracking multiple maneuvering targets in a cluttered environment is a challenging problem. A combination of interacting multiple model (IMM) [4] and joint probabilistic data association (JPDA) [3] has been successfully applied to track multiple maneuvering targets, *e.g.*, IMM-JPDA [5]. In IMM, the motion of a maneuvering target is approximated by a finite number of simple, distinct kinematic models. For survey of IMM, see [6]. However, the exact computation has the time complexity which is exponential in the numbers of kinematic models and measurements. When applying JPDA and IMM, the numbers of targets and kinematic models are known, so we can design a tracking system suitable for the given numbers of targets and kinematic models, *e.g.*, parallel computing. However, the number of measurements is not known in advance, and it poses a serious problem in computing association probabilities in JPDA.

Joint probabilistic data association (JPDA) is developed to solve the data association problem arises in multiple-target tracking [3]. JPDA is a suboptimal single-scan approximation to the optimal Bayesian filter, in which the associations between the "known" tracks and the *latest* observations are made sequentially. At each time step, instead of finding a single best association between latest observations and known tracks, JPDA enumerates all possible associations between observations and tracks and computes association probabilities $\{\beta_{jk}\}$, where β_{jk} is the probability that j-th observation is from k-th track. Given an association, the state of a target is estimated by a filtering algorithm and this conditional expectation of state is weighted by the association probability. Then the state of a target is estimated by summing over the weighted conditional expectations. It has proved very effective in a cluttered environment compared with the nearest neighbor approach which finds a single best association [3].

The exact calculation of association probabilities β_{jk} in JPDA is NP-hard [7] since the related problem of finding the permanent of a matrix is #P-complete [8]. To overcome the complexity of the problem, many approximation algorithms have been proposed. Some heuristic approaches to approximate JPDA include a "cheap" JPDA algorithm [9], "suboptimal" JPDA [10] and "near-optimal" JPDA [11]. In [12], a single-stage data association problem is considered and a leave-one-out heuristic is developed to avoid the enumeration of all possible associations. Sampling methods have been applied before [13], [2]. In [14], Markov chain Monte Carlo (MCMC) is applied to compute the association probabilities in JPDA and it is shown that MCMC outperforms Fitzgerald's cheap JPDA. Unfortunately, in all cases, the performance of an approximation algorithm for JPDA is measured in experiment only.

This paper presents a randomized algorithm, named Markov chain Monte Carlo data association (MCMCDA), for computing association probabilities required for IMM-JPDA and proves that the time complexity of the algorithm is polynomial in the number of kinematic models and the number of measurements. In [15], a general-purpose MCMCDA algorithm is developed to track an unknown number of targets. It has been shown that MCMCDA is computationally efficient compared to the multiple hypothesis tracker (MHT) [16] and outperforms MHT with heuristics, such as pruning, gating, clustering, N-scan-back logic and k-best hypotheses, under extreme conditions, such as a large number of targets in a dense environment, low detection probabilities, and high false alarm rates [15]. The MCMCDA algorithm has been extended to sensor networks in a hierarchical manner to be scalable and it has been shown that MCMCDA is robust against sensor localization error, transmission failures and

The authors are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, {sho,sastry}@eecs.berkeley.edu.

This material is based upon work supported by the National Science Foundation under Grant No. EIA-0122599.

communication delays, *i.e.*, out-of-sequence measurements [17]. In [18], it has been shown that a special case MCMCDA finds good estimates of association probabilities in polynomial time. This paper extends the results to the problem of tracking a known number of *maneuvering* targets.

The remainder of this paper is structured as follows. We describe a method to estimate states of multiple linear hybrid systems in Section II and describe the MCMC method in Section III. The MCMCDA algorithm is presented in Section IV and analysis about the algorithm is shown in Section V. We also present an experiment confirming our results in Section VI.

II. State Estimation of Multiple Linear Hybrid Systems

Let K be the number of targets moving around the surveillance region $\mathcal{R}$. The state dynamics of target k is modeled as

$$x_{t+1}^k = A_t^k(\nu_t^k)x_t^k + G_t^k(\nu_t^k)w_t^k(\nu_t^k), \quad \text{for } t = 1, 2, \ldots, \tag{1}$$

where $x_t^k \in \mathbb{R}^{n_x}$ is the state of target k at time t, ν_t^k denotes the kinematic model representing the motion of target k at time t; $A_t^k(\nu_t^k)$ and $G_t^k(\nu_t^k)$ are matrices with appropriate sizes; and $w_t^k(\nu_t^k)$ is a white Gaussian process with zero mean and covariance $Q_t^k(\nu_t^k)$. The evolution of ν_t^k is modeled by a finite state Markov chain taking values from $\{1, \ldots, M\}$ according to a transition probability matrix $P_{\mathrm{m}}^k = [p_{ij}^k]$. This linear hybrid system is also known as a jump linear system [19] and has been successfully applied to approximate a class of nonlinear systems [20].

The noisy observation of the state of a target is measured with a detection probability p_{d} which is less than unity. There are also false alarms (or outliers) and the number of false alarms has a Poisson distribution with a parameter $\lambda_{\mathrm{f}}V$ where V is the volume of $\mathcal{R}$ and λ_{f} is the false alarm rate per unit time, per unit volume. Let n_t be the number of observations at time t, including both noisy observations and false alarms. Let $y_t^j \in \mathbb{R}^{n_y}$ be the j-th observation at time t for $j = 1, \ldots, n_t$. Each target generates a unique observation at each sampling time if it is detected. The measurement model is

$$y_t^j = \begin{cases} C_t^j(\nu_t^k)x_t^k + w_t'^j(\nu_t^k) & \text{if } y_t^j \text{ is from } x_t^k \\ u_t^j & \text{otherwise,} \end{cases} \tag{2}$$

where $w_t'^j(\nu_t^k)$ is a white Gaussian process with zero mean and covariance $R_t^j(\nu_t^k)$, $C_t^j(\nu_t^k)$ is a matrix with appropriate size, and $u_t^j \sim \mathrm{Unif}(\mathcal{R})$ are random processes for false alarms. Notice that, with probability $1 - p_{\mathrm{d}}$, the target is not detected and we call this a missing observation. Let $Y_t = \{y_t^j : 1 \le j \le n_t\}$ and $Y_{1:t} = \{Y_1, \ldots, Y_t\}$.

For notational convenience, we assume $A(\cdot) = A_t^k(\cdot)$, $G(\cdot) = G_t^k(\cdot)$, $Q(\cdot) = Q_t^k(\cdot)$, $C(\cdot) = C_t^j(\cdot)$, $R(\cdot) = R_t^j(\cdot)$, and $P_{\mathrm{m}} = P_{\mathrm{m}}^k$, for all k, t, and j. Since we are operating at the filtering step t, we further simplify our notations by dropping the subscript t. Let us denote the event $\{\nu_t^k = i\}$ by μ_i^k and let Ω be a set of all *feasible joint association events*. For each $\omega \in \Omega$, $\omega = \{(j, k)\}$, where (j, k) denotes an event that j-th observation is associated with target k. Then the state of target k can be estimated as

$$\mathbb{E}(x_t^k|Y_{1:t}) = \sum_{i=1}^{M} \mathbb{E}(x_t^k|\mu_i^k, Y_{1:t})P(\mu_i^k|Y_{1:t}), \tag{3}$$

where

$$\begin{aligned} \mathbb{E}(x_t^k|\mu_i^k, Y_{1:t}) &= \sum_{\omega} \mathbb{E}(x_t^k|\omega, \mu_i^k, Y_{1:t})P(\omega|\mu_i^k, Y_{1:t}) \quad (4) \\ &= \sum_{j=0}^{n_t} \mathbb{E}(x_t^k|\omega_{jk}, \mu_i^k, Y_{1:t})P(\omega_{jk}|\mu_i^k, Y_{1:t}), \end{aligned}$$

where ω_{jk} denotes the event $\{\omega \ni (j, k)\}$ and ω_{0k} denotes the event that no observation is associated with target k. Notice that other parameters, such as covariance matrices, can be computed similarly and interested readers are referred to [3], [5]. In (4), $\mathbb{E}(x_t^k|\omega_{jk}, \mu_i^k, Y_{1:t})$ can be computed easily by considering it as a single target estimation problem where the current observation is the j-th observation. Hence, the computation of $\mathbb{E}(x_t^k|\mu_i^k, Y_{1:t})$ reduces to the computation of association probability $\beta_{jk}(i)$, where

$$\beta_{jk}(i) := P(\omega_{jk}|\mu_i^k, Y_{1:t}) = \sum_{\omega:(j,k)\in\omega} P(\omega|\mu_i^k, Y_{1:t}). \tag{5}$$

The computation of $\beta_{jk}(i)$ requires a summation over the posteriors, hence the enumeration of all joint association events. JPDA is a method for computing expectations such as (4) using the association probabilities $\beta_{jk}(i)$ in the presence of the identity uncertainty while IMM computes (3) using model posteriors $P(\mu_i^k|Y_{1:t})$. As mentioned earlier, the exact calculation of $\beta_{jk}(i)$ in JPDA is NP-hard [7] and it is the major drawback of JPDA. In next sections, we introduce a randomized algorithm which approximates $\beta_{jk}(i)$ and $P(\mu_i^k|Y_{1:t})$ without the enumerations of all joint association events and all possible combinations of kinematic models.

III. Markov Chain Monte Carlo

Markov chain Monte Carlo (MCMC) plays a significant role in many fields such as physics, statistics, economics, and engineering [21]. In some cases, MCMC is the only known general algorithm that finds a good approximate solution to a complex problem in polynomial time [22]. MCMC techniques have been applied to complex probability distribution integration problems, counting problems such as #P-complete problems, and combinatorial optimization problems [22], [21].

MCMC is a general method to generate samples from a distribution π by constructing a Markov chain $\mathcal{M}$ with states ω and stationary distribution $\pi(\omega)$. If we are at state $\omega \in \Omega$, we propose $\omega' \in \Omega$ following the proposal distribution $q(\omega, \omega')$. The move is accepted with an acceptance probability $A(\omega, \omega')$ where

$$A(\omega, \omega') = \min\left(1, \frac{\pi(\omega')q(\omega', \omega)}{\pi(\omega)q(\omega, \omega')}\right), \tag{6}$$

otherwise the sampler stays at ω, so that the detailed balance condition is satisfied, *i.e.*,

$$Q(\omega, \omega') = \pi(\omega)P(\omega, \omega') = \pi(\omega')P(\omega, \omega'), \qquad (7)$$

for all $\omega, \omega' \in \Omega$, where $P(\omega, \omega') = q(\omega, \omega')A(\omega, \omega')$ is the transition probability from ω to ω' for $\omega' \neq \omega$. The described MCMC algorithm is known as the Metropolis-Hastings algorithm. If $\mathcal{M}$ is irreducible and aperiodic, then $\mathcal{M}$ converges to its stationary distribution by the ergodic theorem [23]. Hence, for a given bounded function f, the sample mean $\hat{f}$ of f over the sampled states converges to $\mathbb{E}_\pi f(\omega)$. Notice that (7) requires only the ability to compute the ratio $\pi(\omega')/\pi(\omega)$, avoiding the need to normalize π.

An ergodic chain $\mathcal{M}$ on state space Ω converges to its stationary distribution asymptotically. But a practical question is how fast $\mathcal{M}$ becomes close to stationarity. One way to measure the rate of convergence of $\mathcal{M}$ to stationarity is the "mixing time" of the Markov chain. Let P be the transition probabilities of $\mathcal{M}$ and let $P_x^t(\cdot)$ be the distribution of the state at time t given that $\mathcal{M}$ is started from the initial state $x \in \Omega$. If π is the stationary distribution of $\mathcal{M}$, then the *total variation distance* at time t with initial state x is defined as

$$\Delta_x(t) = \|P_x^t - \pi\| = \max_{S \subset \Omega} |P_x^t(S) - \pi(S)| \qquad (8)$$

The rate of convergence of $\mathcal{M}$ to stationarity can be measured by the *mixing time*:

$$\tau_x(\epsilon) = \min\{t : \Delta_x(s) \leq \epsilon \text{ for all } s \geq t\}. \qquad (9)$$

One approach to bound $\tau_x(\epsilon)$ of a Markov chain with a complex structure is the canonical path method [22]. In this paper, we consider a highly complex Markov chain, hence we use the canonical path method to bound $\tau_x(\epsilon)$ of the Markov chain simulated by the MCMCDA algorithm given in Section IV.

For a finite, reversible and ergodic Markov chain $\mathcal{M}$ with state space Ω, consider an undirected graph $G = (V, E)$ where $V = \Omega$ and $E = \{(x, y) : Q(x, y) > 0\}$. For each ordered pair $(x, y) \in \Omega^2$, the canonical path γ_{xy} is a simple path from x to y in G. In terms of $\mathcal{M}$ the canonical path γ_{xy} is a sequence of legal transitions from x to y in $\mathcal{M}$. Let $\Gamma = \{\gamma_{xy} : x, y \in \Omega\}$ be the set of all canonical paths. Now the mixing time of the chain is related to the *maximum edge loading*:

$$\bar{\rho} = \bar{\rho}(\Gamma) = \max_e \frac{1}{Q(e)} \sum_{\gamma_{xy} \ni e} \pi(x)\pi(y)|\gamma_{xy}|. \qquad (10)$$

If $\bar{\rho}$ is not so big, *i.e.*, no single edge is overloaded, then the Markov chain can move around fast and achieve the rapidly mixing property. The main result for the canonical path method is as follows [22]:

Theorem 1: Let $\mathcal{M}$ be a finite, reversible, ergodic Markov chain with loop probabilities $P(x, x) \geq \frac{1}{2}$ for all states x. Let Γ be a set of canonical paths with maximum edge loading $\bar{\rho}$. Then the mixing time of $\mathcal{M}$ satisfies $\tau_x(\epsilon) \leq \bar{\rho}(\log \pi(x)^{-1} + \log \epsilon^{-1})$, for any choice of initial state x.

IV. MCMC Data Association Algorithm

In this section, we describe the MCMC data association (MCMCDA) algorithm for efficiently approximating the association probabilities $\beta_{jk}(i)$ and model posteriors $P(\mu_i^k|Y_{1:t})$.

In JPDA, *measurement validation* is used to reduce the number of measurement considered in computation of association probabilities. The same measurement validation is used for MCMCDA but we later find that it is a critical step to approximate association probabilities in polynomial time. Let $\hat{y}^{k,i}$ be the predicted observation for target k when μ_i^k, *i.e.*, $\hat{y}^{k,i} = C(i)\mathbb{E}(x_t^k|\mu_i^k, Y_{1:t-1})$. Suppose there are N observations and let $v_j^{k,i} = y^j - \hat{y}^{k,i}$ for $j = 1, \ldots, N$. The covariance of $v_j^{k,i}$ is $B^{k,i} = \mathbb{E}(v_j^{k,i} v_j^{k,i^T}|\mu_i^k, Y_{1:t-1})$. For each target k, let $B^k = \max_i |B^{k,i}|$. The measurement y^j is validated for target k, if and only if, for some i,

$$v_j^{k,i^T} (B^k)^{-1} v_j^{k,i} < \delta, \qquad (11)$$

where δ is an appropriate threshold. We assume that all measurements are validated with at least one target. If not, we can always consider the reduced problem, which consists of only validated measurements and targets with at least one validated measurement, and separately estimate states of targets with no validated measurement.

We encode the feasible joint association events in a graph. Let $G = (V_{\text{ex}}, E_{\text{ex}})$ be a graph with vertex set $V_{\text{ex}} = W \cup U \cup V$ and edge set $E_{\text{ex}} = F \cup E$, where $W = \{\mu_i^k : 1 \leq k \leq K, 1 \leq i \leq M\}$ is a set of kinematic models, $U = \{k : 1 \leq k \leq K\}$ is a set of target indices, $V = \{y^j : 1 \leq j \leq N\}$ is a set of observations, $F = \{(w, u) : \kappa(w) = u, w \in W, u \in U\}$ with $\kappa : W \to \{1, \ldots, K\}$ mapping kinematic model $w \in W$ to its target index, and $E = \{(u, v) : \exists i \text{ s.t. } (\hat{y}^{u,i} - v)^T (B^u)^{-1} (\hat{y}^{u,i} - v) < \delta, u \in U, v \in V\}$. An edge $(u, v) \in E$ represents that observation v is validated for target u according to (11). An edge $(w, u) \in F$ represents that the kinematic model of target u is w. Note that all kinematic models $\{\mu_i^u : 1 \leq i \leq M\}$ of target u are only connected to u. For future reference, let $G^{\text{a}} = (U, V, E)$ and $G^{\text{m}} = (W, U, F)$; and let $\iota : W \to \{1, \ldots, M\}$ be a mapping from kinematic model $w \in W$ to its model index.

We now define a set of feasible events over G. Let Ω be a set of all feasible events defined over G such that $\omega = (\omega^{\text{m}}, \omega^{\text{a}}) \in \Omega$ with $\omega^{\text{m}} \subset F$ and $\omega^{\text{a}} \subset E$ forms a subgraph in G. An event $\omega = (\omega^{\text{m}}, \omega^{\text{a}}) \in \Omega$ is feasible if ω^{a} is a matching in G^{a} and ω^{m} is a k-matching in G^{m}. Based on the parametric false alarm model, the posterior of $\omega \in \Omega$ can be computed as

$$\begin{aligned} P(\omega|Y_{1:t}) &= P(\omega^{\text{a}}, \omega^{\text{m}}|Y_{1:t}) \\ &= \tfrac{1}{Z} P(\omega^{\text{a}}|Y_{1:t-1}) P(Y_t|\omega^{\text{a}}, \omega^{\text{m}}, Y_{1:t-1}) P(\omega^{\text{m}}|Y_{1:t-1}) \\ &= \tfrac{1}{Z} \lambda_{\text{f}}^{N-|\omega^{\text{a}}|} p_{\text{d}}^{|\omega^{\text{a}}|} (1 - p_{\text{d}})^{K-|\omega^{\text{a}}|} \textstyle\prod_{(u,v)\in\omega^{\text{a}}} \mathcal{N}_u(v; \omega^{\text{m}}) \\ &\quad \times \textstyle\prod_{(w,u)\in\omega^{\text{m}}} P\left(\mu_{\iota(w)}^u|Y_{1:t-1}\right), \end{aligned} \qquad (12)$$

where Z is a normalizing constant and $\mathcal{N}_u(v; \omega^{\text{m}})$ is the Gaussian density function of variable v with mean $\hat{y}^{u,\iota(w)}$ and covariance $B^{u,\iota(w)}$ for $(w, u) \in \omega^{\text{m}}$. Notice that

Algorithm 1 (Multiple-model) MCMCDA (single step)

Sample U from Unif$[0,1]$
if $U < \frac{1}{2}$ **then**
  $\omega' = \omega$
else
  $(\omega^{\mathrm{m}}, \omega^{\mathrm{a}}) = \omega$
  Choose $e = (u,v) \in E_{\mathrm{ex}} \setminus \omega^{\mathrm{m}}$ uniformly at random
  if $e \in F$ **then**
    $\omega'^{\mathrm{m}} = \omega^{\mathrm{m}} + e - e'$, where $e' = (w,v) \in \omega^{\mathrm{m}}$
  else
    if $e \in \omega^{\mathrm{a}}$ **then**
      $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} - e$
    else if both u and v are unmatched in ω^{a} **then**
      $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} + e$
    else if exactly one of u and v is matched in ω^{a} and e' is the matching edge **then**
      $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} + e - e'$
    else
      $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}}$
    end if
  end if
  $\omega' = (\omega'^{\mathrm{m}}, \omega'^{\mathrm{a}})$
end if
$\omega = \omega'$ with probability $A(\omega, \omega')$

$P(\mu^u_{\iota(w)}|Y_{1:t-1})$ are computed from the interaction step of IMM [5].

The MCMC data association (MCMCDA) algorithm is an MCMC algorithm whose state space is the set of all feasible events Ω and whose stationary distribution is the posterior (12). Each step of the MCMCDA algorithm is described in Algorithm 1, where $A(\omega,\omega') = \min\left(1, \frac{\pi(\omega')}{\pi(\omega)}\right)$ and $\pi(\omega) = P(\omega|Y_{1:t})$ from (12). There are four MCMC moves and we name them for future reference: (i) a model switch move proposes $\omega'^{\mathrm{m}} = \omega^{\mathrm{m}} + e - e'$; (ii) an addition move proposes $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} + e$; (iii) a deletion move proposes $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} - e$; and (iv) a switch move proposes $\omega'^{\mathrm{a}} = \omega^{\mathrm{a}} + e - e'$.

Now suppose that Algorithm 1 is repeated for S steps and let ω_s be the state of the chain at step s. Let $\hat{\mu}_i^k = \frac{1}{S}\sum_{s=1}^{S} \mathbb{I}((\mu_i^k, k) \in \omega_s^{\mathrm{m}})$, $\hat{\beta}_{jk}(i) = \frac{1}{S\hat{\mu}_i^k}\sum_{s=1}^{S} \mathbb{I}((k, y^j) \in \omega_s^{\mathrm{a}}, (\mu_i^k, k) \in \omega_s^{\mathrm{m}})$ for $\hat{\mu}_i^k > 0$ and $\hat{\beta}_{jk}(i) = 0$ for $\hat{\mu}_i^k = 0$, where $\mathbb{I}$ is an indicator function. If the Markov chain is ergodic, by the ergodic theorem [23], $\hat{\beta}_{jk}(i) \to \beta_{jk}(i)$ and $\hat{\mu}_i^k \to P(\mu_i^k|Y_{1:t})$ almost surely as $S \to \infty$. In the next section, we show that the Markov chain simulated by Algorithm 1 is ergodic and analyze the rate of this convergence. If the chain converges to its stationary distribution fast, we can find good approximations with small S.

V. Analysis

Let $\mathcal{M}$ be the Markov chain simulated by Algorithm 1. Since the self-loop probability is nonzero, $\mathcal{M}$ is aperiodic. It can be easily seen that $\mathcal{M}$ is irreducible, *i.e.*, all states communicate, for example via the empty matching. In addition, the transitions described in Algorithm 1 satisfy the detailed balance condition (7) so $\mathcal{M}$ is reversible. Hence, by the ergodic theorem, the chain converges to its stationary distribution [23].

We first establish few facts. In (12), the normalizing constant is

$$Z = \sum_{\omega\in\Omega} P(\omega^{\mathrm{a}}|Y_{1:t-1})P(Y_t|\omega^{\mathrm{a}},\omega^{\mathrm{m}},Y_{1:t-1})P(\omega^{\mathrm{m}}|Y_{1:t-1}). \tag{13}$$

We can bound each likelihood term as $\underline{L} \le \mathcal{N}_u(v;\omega^{\mathrm{m}}) \le \bar{L}$, for all $(u,v) \in E_{\mathrm{a}}$ and ω^{m}, where

$$\begin{aligned}\bar{L} &= \max_{1\le k\le K, 1\le i\le M}\left\{\left((2\pi)^{n_y}|B^{k,i}|\right)^{-\frac{1}{2}}\right\}\\ \underline{L} &= \min_{1\le k\le K, 1\le i\le M}\left\{\left((2\pi)^{n_y}|B^{k,i}|e^{c(k,i)\delta}\right)^{-\frac{1}{2}}\right\}.\end{aligned}$$

Here, $B^{k,i}$ are positive definite matrices and $c(k,i) = \lambda_{\max}((B^{k,i})^{-1})/\lambda_{\min}((B^k)^{-1})$, where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the maximum and minimum eigenvalues of matrix A, respectively. We assume that targets are maintained by covariance control and $c(k,i) < \bar{c}$ for some constant $\bar{c}$. Notice that the lower bound $\underline{L}$ is due to the measurement validation.

To prove Theorem 2, we need the following lemmas. Note that the omitted proofs are given in Appendix.

Lemma 1: For any $\omega_0, \omega_1, \omega_2 \in \Omega$, if $\omega_0^{\mathrm{m}} = \omega_1^{\mathrm{m}} = \omega_2^{\mathrm{m}}$ and $\omega_1 = \omega_0 - e_0$, for some edge $e_0 \in \omega_0$, and $\omega_2 = \omega_1 - e_1$, for some edge $e_1 \in \omega_1$, then $\pi(\omega_0)/\pi(\omega_1) \le C$, $\pi(\omega_0)/\pi(\omega_2) \le C^2$, $\pi(\omega_1)/\pi(\omega_0) \le D$, and $\pi(\omega_2)/\pi(\omega_0) \le D^2$, where $C = \frac{p_{\mathrm{d}}L}{\lambda_{\mathrm{f}}(1-p_{\mathrm{d}})}$ and $D = \frac{\lambda_{\mathrm{f}}(1-p_{\mathrm{d}})}{\underline{L}p_{\mathrm{d}}}$.

Lemma 2: Suppose that $P\left(\mu_i^k|Y_{1:t-1}\right) \ge \underline{\mu}$ for all $1 \le i \le M$ and $1 \le k \le K$. For any $\omega_1, \omega_2 \in \Omega$, if $\omega_1^{\mathrm{a}} = \omega_2^{\mathrm{a}}$ and $\omega_1 = \omega_2 - e_2 + e_1$ for edges $e_1 \in \omega_1$ and $e_2 \in \omega_2$ with e_1 and e_2 sharing a common vertex, then $\frac{\pi(\omega_1)}{\pi(\omega_2)} \le H$, where $H = \bar{L}/(\underline{\mu}\underline{L})$.

Lemma 3: Suppose that $P\left(\mu_i^k|Y_{1:t-1}\right) \ge \underline{\mu}$ for all $1 \le i \le M$ and $1 \le k \le K$. Let $R = \max\{1, C, D, H\}$, where C and D are defined in Lemma 1 and H is defined in Lemma 2. Then the maximum edge loading of the Markov chain $\mathcal{M}$ is bounded as $\bar{\rho} \le 6R^4K^2(N+M)$.

Remark 1: In Lemma 3, we have assumed that $P\left(\mu_i^k|Y_{1:t-1}\right) \ge \underline{\mu}$ for all i and k. If $P\left(\mu_i^k|Y_{1:t-1}\right) < \underline{\mu}$, then the contribution from model i for target k is small in (3) and its contribution can be safely ignored. Thus, one strategy is to ignore models whose priors are less than the threshold $\underline{\mu}$. It can be seen as "model validation" similar to the measurement validation in JPDA. Hence, from now on, we assume that $P\left(\mu_i^k|Y_{1:t-1}\right) \ge \underline{\mu}$ for all $1 \le i \le M$ and $1 \le k \le K$.

For Theorem 2 below, define $m_1 = \max\{1, \bar{L}\}$, $m_2 =$

$\min\{1, \underline{L}\}$,

$$
\begin{aligned}
m_3(K,N) &= \max_{0\le k\le K}\{\lambda_{\mathrm{f}}^{N-k} p_{\mathrm{d}}^k (1-p_{\mathrm{d}})^{K-k}\},\\
m_4(K,N) &= \min_{0\le k\le K}\{\lambda_{\mathrm{f}}^{N-k} p_{\mathrm{d}}^k (1-p_{\mathrm{d}})^{K-k}\}, \quad \text{and}\\
m_5(K,N) &= K\log\frac{m_1}{m_2}\frac{M}{\underline{\mu}} + \log\frac{m_3(K,N)}{m_4(K,N)}\\
&+ \sum_{k=1}^{K+1}\log k + \sum_{n=1}^{N}\log n.
\end{aligned}
$$

Theorem 2: Suppose that $\lambda_{\mathrm{f}} > 0$ and $0 < p_{\mathrm{d}} < 1$. Then the mixing time of the Markov chain $\mathcal{M}$ is bounded by $\tau_x(\epsilon) \le 6R^4K^2(N+M)(m_5(K,N)+\log\epsilon^{-1})$ for all $x \in \Omega$.

Remark 2: Let $\bar{\tau}(\epsilon)$ be the upper bound found in Theorem 2. Since $m_5(K,N)$ is polynomial in K and N, $\bar{\tau}(\epsilon)$ is polynomial in the number of kinematic models K and the number of measurements N. If we assume that both K and M are fixed, $\bar{\tau}(\epsilon) = O(N(N\log N + \log\epsilon^{-1}))$.

Let $p(\omega)$ be the distribution of the states of $\mathcal{M}$ after simulating Algorithm 1 for at least $\bar{\tau}(\epsilon)$ steps. Then the total variation distance satisfies $\|p-\pi\| \le \epsilon$. Hence, for a given bounded function $f : \Omega \to \mathbb{R}$, we can estimate $\mathbb{E}_\pi f$ by the sample mean $\hat{f} = \frac{1}{S}\sum_{s=1}^{S} f(\omega_s)$, where $\{\omega_s\}$ are sampled from p. However, there is a small bias in our estimates since we are not sampling from π. The following theorem from [18] gives an upper bound on the number of samples needed for finding good estimates.

Theorem 3: Let $0 < \epsilon_1, \epsilon_2 \le 1$ and $0 < \eta < .5$. Suppose that $\|p-\pi\| \le \epsilon$ for $\epsilon \le \epsilon_1\epsilon_2/8$. If $f : \Omega \to [0,1]$, then, with a total of $504\epsilon_1^{-2}\epsilon_2^{-1}\lceil\log\eta^{-1}\rceil$ samples from p, we can find estimates $\hat{f}$ for $\mathbb{E}_\pi f$ with probability at least $1-\eta$, such that, for $\mathbb{E}_\pi f \ge \epsilon_2$, $\hat{f}$ estimates $\mathbb{E}_\pi f$ within ratio $1+\epsilon_1$, *i.e.*, $(1-\epsilon_1)\mathbb{E}_\pi f \le \hat{f} \le (1+\epsilon_1)\mathbb{E}_\pi f$, and, for $\mathbb{E}_\pi f < \epsilon_2$, $\hat{f} \le (1+\epsilon_1)\epsilon_2$.

Following Remark 2, for fixed K and M, $\bar{\tau}(\epsilon) = O(N(N\log N + \log\epsilon^{-1}))$. Combining this fact with Theorem 3, the time complexity of the overall procedure is $S = O(\epsilon_1^{-2}\epsilon_2^{-1}\log\eta^{-1}N(N\log N + \log(\epsilon_1^{-1}\epsilon_2^{-1})))$. Hence, with a total of S samples, Algorithm 1 finds estimates $\hat{f}$ for $\mathbb{E}_\pi f$ with probability at least $1-\eta$, such that, for $\mathbb{E}_\pi f \ge \epsilon_2$, $\hat{f}$ estimates $\mathbb{E}_\pi f$ within ratio $1+\epsilon_1$, and, for $\mathbb{E}_\pi f < \epsilon_2$, $\hat{f} \le (1+\epsilon_1)\epsilon_2$. We can simplify further by letting $\epsilon_0 = \epsilon_1\epsilon_2$. Then the time complexity is $O(\epsilon_0^{-2}\log\eta^{-1}N(N\log N + \log(\epsilon_0^{-1})))$.

Hence, we can estimate $P(\mu_i^k|Y_{1:t})$ by letting $f(\omega) = \mathbb{I}((\mu_i^k, k) \in \omega^{\mathrm{m}})$. The association probability $\beta_{jk}(i)$ can be estimated by $\hat{f}_1/\hat{f}_2$ where $f_1(\omega) = \mathbb{I}((k, y^j) \in \omega^{\mathrm{a}}, (\mu_i^k, k) \in \omega^{\mathrm{m}})$ and $f_2(\omega) = \mathbb{I}((\mu_i^k, k) \in \omega^{\mathrm{m}})$. As long as $\hat{f}_2 \ge \epsilon'$ for some $\epsilon' > 0$, we can find a good approximation of $\beta_{jk}(i)$ with a small number of samples. Again, when $P(\mu_i^k|Y_{1:t})$ is small, the contribution from model i for target k is small in (3), hence its contribution can be safely ignored. Thus, it is beneficial to apply the model validation method discussed in Remark 1 when computing $\beta_{jk}(i)$ and eventually (3).

VI. Simulation Results

In this section, we compare the performance of MCMCDA-based IMM algorithm against JPDA-based IMM algorithm for tracking multiple maneuvering targets. There are 8 targets and each target has three kinematic models based on the discrete-time linear dynamics (1):

(Model 1) Second-order kinematic model:

$$
A(\nu_t^k = 1) = \begin{bmatrix} 1 & \delta & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & \delta\\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
G(\nu_t^k = 1) = \begin{bmatrix} \delta^2/2 & 0\\ \delta & 0\\ 0 & \delta^2/2\\ 0 & \delta \end{bmatrix}
$$

and $Q(\nu_t^k = 1) = \mathrm{diag}(0.83, 0.83)$, where δ is the sampling period. The state vector is $x = (x_1, \dot{x}_1, x_2, \dot{x}_2)^T$. This model assumes that the variation in a velocity component is a discrete time white noise acceleration [24].

(Model 2) Third-order kinematic model:

$$
A(\nu_t^k = 2) = \begin{bmatrix} 1 & \delta & \delta^2/2 & 0 & 0 & 0\\ 0 & 1 & \delta & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & \delta & \delta^2/2\\ 0 & 0 & 0 & 0 & 1 & \delta\\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$
G(\nu_t^k = 2) = \begin{bmatrix} \delta^2/2 & 0\\ \delta & 0\\ 1 & 0\\ 0 & \delta^2/2\\ 0 & \delta\\ 0 & 1 \end{bmatrix}
$$

and $Q(\nu_t^k = 2) = \mathrm{diag}(13.22, 13.22)$. The state vector is $x = (x_1, \dot{x}_1, \ddot{x}_1, x_2, \dot{x}_2, \ddot{x}_2)^T$. This is a third-order kinematic model with accelerations modeled as a discrete time Wiener process [24].

(Model 3) Third-order kinematic model: The same third-order kinematic model used in (model 2) but $Q(\nu_t^k = 3) = \mathrm{diag}(0.83, 0.83)$.

In all cases, the measurement model (2) is used with

$$
C = \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 \end{bmatrix}
$$

and $R = \mathrm{diag}(0.83, 0.83)$.

The trajectories of 8 targets are shown in Figure 1. There were about 200 measurements per each scan including false alarms. In this example, $\lambda_{\mathrm{f}} = .0002$, $p_{\mathrm{d}} = .998$, and the model transition matrix is

$$
P_{\mathrm{m}} = \begin{bmatrix} .6 & .2 & .2\\ .2 & .6 & .2\\ .2 & .2 & .6 \end{bmatrix}.
$$

The initial distributions of kinemetic models are $P(\nu_0^k = 1) = .4, P(\nu_0^k = 2) = .3$ and $P(\nu_0^k = 3) = .3$. The

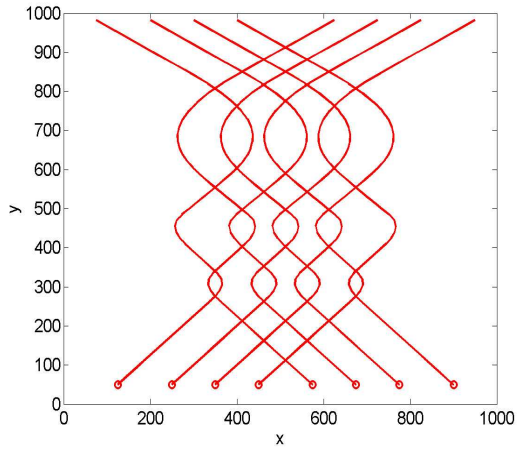

Fig. 1. A scenario used in simulation - trajectories of 8 targets (initial positions are circled)

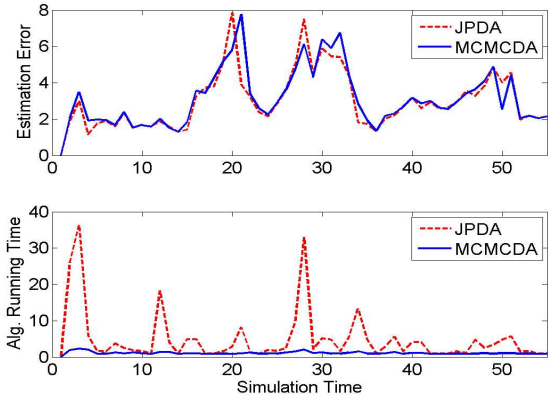

Fig. 2. Comparison between MCMCDA against JPDA

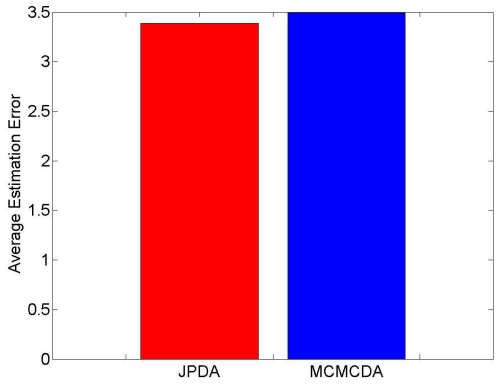

Fig. 3. Average estimation error

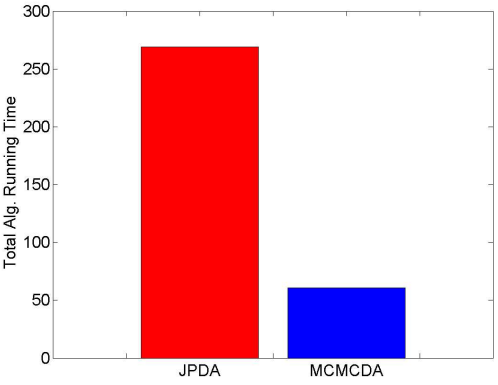

Fig. 4. Total running time

average velocity is about 20 unit length per sampling time. This example shows high levels of maneuvers and difficulties in data association when targets move close by or cross over each other. Without good initial position estimates, tracking is not possible in this cluttered environment. Both JPDA and MCMCDA have maintained tracks and this shows the robustness of Bayesian formulation used in JPDA and MCMCDA against clutter. In Figure 2, the performance of MCMCDA-based IMM is compared against JPDA-based IMM. The overall root-mean-square error of MCMCDA was 3.49 and it was slightly higher than JPDA's 3.39 (see Figure 3). But MCMCDA outperformed in terms of the algorithm running time. The overall running time of MCMCDA was 60.92s while it was 269.11s for JPDA (see Figure 4). Hence, MCMCDA saved more than 75% of computation time compared to JPDA while achieving about the same level of performance.

VII. Conclusion

Tracking multiple maneuvering targets in a cluttered environment is a challenging problem. A combination of IMM and JPDA has proved very effective for solving this problem. However, the exact computation of the combined approach has the time complexity which is exponential in the numbers of kinematic models and measurements. In this paper, we have presented an efficient approximation algorithm for tracking multiple maneuvering targets based on Markov chain Monte Carlo data association (MCMCDA) and proved that the time complexity of the algorithm is polynomial in the size of the problem.

Appendix

Proof of Lemma 1

ω_0 and ω_1 are identical except that ω_1 is missing the edge $e_0 \in E$. So $|\omega_0| = |\omega_1| + 1$. If $e_0 = (u, v)$ and $k = |\omega_0^{\mathrm{a}}|$,

$$\begin{aligned}\pi(\omega_0)/\pi(\omega_1) &= \frac{\lambda_{\mathrm{f}}^{N-k} p_{\mathrm{d}}^{k} (1-p_{\mathrm{d}})^{K-k}}{\lambda_{\mathrm{f}}^{N-(k-1)} p_{\mathrm{d}}^{k-1} (1-p_{\mathrm{d}})^{K-(k-1)}} \mathcal{N}_u(v; \omega_0^{\mathrm{m}}) \\ &= \frac{p_{\mathrm{d}}}{\lambda_{\mathrm{f}}(1-p_{\mathrm{d}})} \mathcal{N}_u(v; \omega_0^{\mathrm{m}}) \qquad \leq C.\end{aligned}$$

On the other hand,

$$\begin{aligned}\pi(\omega_1)/\pi(\omega_0) &= \frac{\lambda_{\mathrm{f}}^{N-(k-1)} p_{\mathrm{d}}^{k-1} (1-p_{\mathrm{d}})^{K-(k-1)}}{\lambda_{\mathrm{f}}^{N-k} p_{\mathrm{d}}^{k} (1-p_{\mathrm{d}})^{K-k}} \frac{1}{\mathcal{N}_u(v; \omega_0^{\mathrm{m}})} \\ &= \frac{\lambda_{\mathrm{f}}(1-p_{\mathrm{d}})}{p_{\mathrm{d}}} \frac{1}{\mathcal{N}_u(v; \omega_0^{\mathrm{m}})} \qquad \leq D.\end{aligned}$$

Since $\pi(\omega_0)/\pi(\omega_2) = \pi(\omega_0)/\pi(\omega_1) \times \pi(\omega_1)/\pi(\omega_2)$, by repeating the above argument twice, we get $\pi(\omega_0)/\pi(\omega_2) \leq C^2$. Similarly, we have $\pi(\omega_2)/\pi(\omega_0) \leq D^2$.

Proof of Lemma 2

Suppose that $e_1 = (w_1, k)$ and $e_2 = (w_2, k)$ and let $i_1 = \iota(w_1)$ and $i_2 = \iota(w_2)$. Since $\omega_1^{\mathrm{a}} = \omega_2^{\mathrm{a}}$,

$$\frac{\pi(\omega_1)}{\pi(\omega_2)} = \frac{P\left(\mu_{i_1}^k | Y_{1:t-1}\right)}{P\left(\mu_{i_2}^k | Y_{1:t-1}\right)} \frac{P(Y_t|\omega_1, Y_{1:t-1})}{P(Y_t|\omega_2, Y_{1:t-1})}. \quad (14)$$

Now at most one observation is connected to the vertex k. If there is an observation connected to k, likelihoods $P(Y_t|\omega_1, Y_{1:t-1})$ and $P(Y_t|\omega_2, Y_{1:t-1})$ differ only for this observation. Hence, the likelihood ratio in (14) is bounded above by $\bar{L}/\underline{L}$. Notice that $\bar{L}/\underline{L} \geq 1$. Since $P\left(\mu_{i_2}^k | Y_{1:t-1}\right) \geq \underline{\mu}$, $\frac{\pi(\omega_1)}{\pi(\omega_2)} \leq H$.

Proof of Lemma 3

For $X, Y \in \Omega$, the canonical path γ_{XY} is defined as follows. Consider the symmetric differences $X^{\mathrm{m}} \oplus Y^{\mathrm{m}}$ and $X^{\mathrm{a}} \oplus Y^{\mathrm{a}}$, where $X \oplus Y = (X - Y) \cup (Y - X)$. $X^{\mathrm{m}} \oplus Y^{\mathrm{m}}$ is a disjoint collection of paths in G^{m}, each of which has edges that belong to X^{m} and Y^{m} alternately. $X^{\mathrm{a}} \oplus Y^{\mathrm{a}}$ is a disjoint collection of paths in G^{a} including closed cycles, each of which has edges that belong to X^{a} and Y^{a} alternately. We first fix ordering on simple paths in G^{m} followed by simple paths in G^{a}. For paths in G^{a}, designate a "start vertex" to each of the paths, which is arbitrary if the path is a closed cycle but must be an endpoint otherwise. This gives a unique ordering $P_1, P_2, \ldots, P_m$ on the paths in $X^{\mathrm{m}} \oplus Y^{\mathrm{m}}$ followed by $P_{m+1}, P_{m+1}, \ldots, P_n$ on the paths appearing in $X^{\mathrm{a}} \oplus Y^{\mathrm{a}}$. The canonical path from X to Y involves "unwinding" each of the P_i in turn as follows. We need to consider three cases:

(i) P_i is a path in $X^m \oplus Y^m$. P_i contains three vertices since each target is connected to a single model. Let P_i consist of the sequence (w_1, k, w_2). Perform the switch move replacing (w_1, k) with (w_2, k).

(ii) P_i is a path in $X^a \oplus Y^a$ and P_i is not a cycle. Let P_i consist of the sequence $(v_0, v_1, \ldots, v_l)$ of vertices with the start vertex v_0. If $(v_0, v_1) \in Y$, perform a sequence of switching moves replacing (v_{2j+1}, v_{2j+2}) by (v_{2j}, v_{2j+1}) for $j = 0, 1, \ldots$, and finish with an addition move if l is odd. If $(v_0, v_1) \in X$, remove (v_0, v_1) and proceed as before for the reduced path $(v_1, \ldots, v_l)$.

(iii) P_i is a path in $X^a \oplus Y^a$ and P_i is a cycle. Let P_i consist of the sequence $(v_0, v_1, \ldots, v_{2l+1})$ of vertices, for $l \geq 1$, where v_0 is the start vertex, and $(v_{2j}, v_{2j+1}) \in X$ for $j = 0, \ldots, l$, with remaining edges belonging to Y. We first remove the edge (v_0, v_1). Now we are left with an open path O with endpoints v_0, v_1, with the start vertex v_k of O, for $k \in \{0, 1\}$. Then we unwind O as in (i) above but treating v_{1-k} as the start vertex to identify that it was a cycle.

Let q be an arbitrary edge in the Markov chain $\mathcal{M}$, *i.e.*, a transition from ω to $\omega' \neq \omega$. Let $cp(q) = \{(X, Y) : \gamma_{XY} \ni q\}$ be the set of canonical paths that use q. We define a function $\eta_q : cp(q) \to \Omega$ as follows:

$$\eta_q(X,Y) = \begin{cases} X \oplus Y \oplus \omega, \\ \quad \text{if } q \text{ is a model switch move;} \\ X \oplus Y \oplus (\omega \cup \omega') - e_{XY_q}, \\ \quad \text{if } q \text{ is a switch move and} \\ \quad \text{the current path is a cycle;} \\ X \oplus Y \oplus (\omega \cup \omega'), \quad \text{otherwise,} \end{cases}$$

where e_{XY_q} is the edge in X adjacent to the start vertex that was removed first in (iii) above. Notice that $\eta_q(X, Y) \in \Omega$ and an injective function. Since $|F| = MK$,

$$\begin{aligned} Q(q) &= Q(\omega, \omega') = \pi(\omega) P(\omega, \omega') \quad (15) \\ &= \frac{1}{2(|E| + (M-1)K)} \min\{\pi(\omega), \pi(\omega')\}. \end{aligned}$$

Next, we bound $\pi(X)\pi(Y)$ and we need to consider five cases:

(i) q is a model switch move. We have $\omega' = \omega + e - e'$ and $\omega \cup \eta_q(X, Y)$ and $X \cup Y$ are identical when viewed as multisets. Hence,

$$\begin{aligned} \pi(X)\pi(Y) &= \pi(\omega)\pi(\eta_q(X,Y)) \\ &= \tfrac{2(|E|+(M-1)K)Q(q)}{\min\{\pi(\omega),\pi(\omega')\}} \pi(\omega)\pi(\eta_q(X,Y)) \\ &= 2(|E| + (M-1)K)Q(q) \max\left\{1, \tfrac{\pi(\omega)}{\pi(\omega')}\right\} \pi(\eta_q(X,Y)) \\ &\leq 2H(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)) \\ &\leq 2R(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)), \end{aligned}$$

where we used (15) in the second equality and Lemma 2 in the first inequality.

(ii) q is a deletion move. We have $\omega' = \omega - e$ and $\eta_q(X, Y) = X \oplus Y \oplus (\omega \cup \omega')$. Since $\omega \cup \eta_q(X, Y)$ and $X \cup Y$ are identical when viewed as multisets,

$$\begin{aligned} \pi(X)\pi(Y) &= \pi(\omega)\pi(\eta_q(X,Y)) \\ &= 2(|E| + (M-1)K)Q(q) \max\left\{1, \tfrac{\pi(\omega)}{\pi(\omega')}\right\} \pi(\eta_q(X,Y)) \\ &\leq 2R(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)), \end{aligned}$$

where we used 15) in the second equality and Lemma 1 for the last inequality.

(iii) q is an addition move. We have $\omega' = \omega + e$ and $\eta_q(X, Y) = X \oplus Y \oplus (\omega \cup \omega')$. Since $\omega \cup \eta_q(X, Y)$ and $X \cup Y$ are identical when viewed as multisets, using the arguments from (i),

$$\pi(X)\pi(Y) \leq 2R(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)).$$

(iv) q is a switch move and the current path is a cycle. Suppose $\omega' = \omega + e - e'$. Let $\omega_1 = \omega + e$. Then $\omega' = \omega_1 - e'$. Since $\frac{\pi(\omega)}{\pi(\omega')} = \frac{\pi(\omega_1)}{\pi(\omega')} \frac{\pi(\omega)}{\pi(\omega_1)}$, by Lemma 1, $\frac{\pi(\omega)}{\pi(\omega')} \leq CD \leq R^2$. Since $\eta_q(X, Y) = X \oplus Y \oplus (\omega \cup \omega') - e_{XY_q}$, the multisets $\omega \cup \eta_q(X, Y)$ differs from $X \cup Y$ only in that e and e_{XY_q} are missing from it. Hence, by Lemma 1,

$$\begin{aligned} \pi(X)\pi(Y) &\leq C^2\pi(\omega)\pi(\eta_q(X,Y)) \\ &= 2C^2(|E| + (M-1)K)Q(q) \max\left\{1, \tfrac{\pi(\omega)}{\pi(\omega')}\right\} \pi(\eta_q(X,Y)) \\ &\leq 2R^4(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)). \end{aligned}$$

(v) q is a switch move and the current path is not a cycle. This case is similar to (iii) but the multisets $\omega \cup \eta_q(X, Y)$ differs from $X \cup Y$ only in that e is missing from it. Hence, by Lemma 1,

$$\begin{aligned} \pi(X)\pi(Y) &\leq C\pi(\omega)\pi(\eta_q(X,Y)) \\ &= 2C(|E| + (M-1)K)Q(q) \max\left\{1, \tfrac{\pi(\omega)}{\pi(\omega')}\right\} \pi(\eta_q(X,Y)) \\ &\leq 2R^3(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)). \end{aligned}$$

In summary, we have

$$\pi(X)\pi(Y) \leq 2R^4(|E| + (M-1)K)Q(q)\pi(\eta_q(X,Y)).$$

Thus, for any transition q,

$$\begin{aligned} &\tfrac{1}{Q(q)} \textstyle\sum_{\gamma_{XY} \ni q} \pi(X)\pi(Y)|\gamma_{XY}| \\ &\leq 2R^4(|E| + (M-1)K) \textstyle\sum_{\gamma_{XY} \ni q} \pi(\eta_q(X,Y))|\gamma_{XY}| \\ &\leq 6R^4K(|E| + (M-1)K) \textstyle\sum_{\gamma_{XY} \ni q} \pi(\eta_q(X,Y)) \\ &\leq 6R^4K(|E| + (M-1)K)) \\ &\leq 6R^4K^2(N+M) \end{aligned}$$

where the second inequality follows from the fact that the length of any canonical path is bounded by $3K$, the third equality is due to the fact that η_q is injective and π is a probability distribution, and the last inequality follows from $|E| \leq KN$. Hence, $\bar{\rho} \leq 6R^4K^2(N+M)$.

Proof of Theorem 2

$\mathcal{M}$ is a finite, reversible, ergodic Markov chain with loop probabilities $P(x,x) \geq \frac{1}{2}$ for all states x (see Section IV). Hence, by Theorem 1, we have

$$\tau_x(\epsilon) \leq \bar{\rho}(\log \pi(x)^{-1} + \log \epsilon^{-1}). \tag{16}$$

The upper bound for $\bar{\rho}$ is computed from Lemma 3. Now we just need to find the upper bound for $\pi(x)^{-1}$. From (13),

$$\begin{aligned} Z &\leq \sum_{\omega \in \Omega} m_1^K m_3(K,N) \\ &= m_1^K m_3(K,N)|\Omega| \end{aligned}$$

From [18], for fixed ω^{m},

$$|\{\omega \in \Omega : \omega^{\mathrm{m}} \subset \omega\}| \leq \sum_{k=0}^{K} \binom{K}{k} \frac{N!}{(N-k)!}.$$

Since the number of different ω^{m} can be at most M^K,

$$|\Omega| \leq M^K \sum_{k=0}^{K} \binom{K}{k} \frac{N!}{(N-k)!}.$$

Hence,

$$\begin{aligned} Z &\leq m_1^K m_3(K,N) M^K \sum_{k=0}^{K} \binom{K}{k} \frac{N!}{(N-k)!} \\ &\leq m_1^K m_3(K,N) M^K (K+1)! N!, \end{aligned}$$

Although this bound on Z is not tight, it will serve our purpose. For any $\omega \in \Omega$, $\pi(\omega) \geq \frac{1}{Z} m_2^K m_4(K,N) \underline{\mu}^K$ so

$$\begin{aligned} \frac{1}{\pi(\omega)} &\leq \frac{Z}{m_2^K m_4(K,N) \underline{\mu}^K} \\ &\leq \left(\frac{m_1}{m_2}\frac{M}{\underline{\mu}}\right)^K \frac{m_3(K,N)}{m_4(K,N)} (K+1)! N!. \end{aligned}$$

Hence,

$$\begin{aligned} \log \tfrac{1}{\pi(\omega)} &\leq \log\left(\left(\tfrac{m_1}{m_2}\tfrac{M}{\underline{\mu}}\right)^K \tfrac{m_3(K,N)}{m_4(K,N)} (K+1)! N!\right) \\ &= m_5(K,N). \end{aligned}$$

Putting all together, we have $\tau_x(\epsilon) \leq 6R^4K^2(N+M)(m_5(K,N) + \log \epsilon^{-1})$ for all initial state $x \in \Omega$.

REFERENCES

[1] I. Cox, "A review of statistical data association techniques for motion correspondence," *International Journal of Computer Vision*, vol. 10, no. 1, pp. 53–66, 1993.

[2] F. Dellaert, S. Seitz, C. Thorpe, and S. Thrun, "EM, MCMC, and chain flipping for structure from motion with unknown correspondence," *Machine Learning*, vol. 50, pp. 45–71, 2003.

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

[4] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, pp. 780–783, 1988.

[5] B. Chen and J. Tugnait, "Tracking of multiple maneuvering targets in clutter using imm/jpda filtering and fixed-lag smoothing," *Automatica*, vol. 37, pp. 239–249, 2001.

[6] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, pp. 103–123, 1998.

[7] J. Collins and J. Uhlmann, "Efficient gating in data association with multivariate distributed states," *IEEE Trans. Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 909–916, July 1992.

[8] L. Valiant, "The complexity of computing the permanent," *Theoretical Computer Science*, vol. 8, pp. 189–201, 1979.

[9] R. Fitzgerald, "Development of practical PDA logic for multipltarget tracking by microprocessor," in *Multitarget-Multisensor Tracking: Advanced Applications*, Y. Bar-Shalom, Ed. Artech House: Norwood, MA, 1990.

[10] J. Roecker and G. Phillis, "Suboptimal joint probabilistic data association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-29, 2, pp. 510–517, April 1993.

[11] J. Roecker, "A class of near optimal JPDA algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-30, 2, pp. 504–510, April 1994.

[12] T. Huang and S. J. Russell, "Object identification in a Bayesian context," in *Proc. of the International Joint Conference on Artificial Intelligence*, Nagoya, Japan, Aug. 1997.

[13] H. Pasula, S. J. Russell, M. Ostland, and Y. Ritov, "Tracking many objects with many sensors," in *Proc. of the International Joint Conference on Artificial Intelligence*, Stockholm, 1999.

[14] S. Cong, L. Hong, and D. Wicker, "Markov-chain Monte-Carlo approach for association probability evaluation," *IEE Proceedings of Control, Theory and Applications*, vol. 151, no. 2, pp. 185–193, March 2004.

[15] S. Oh, S. Russell, and S. Sastry, "Markov chain Monte Carlo data association for general multiple-target tracking problems," in *Proc. of the 43rd IEEE Conference on Decision and Control*, Paradise Island, Bahamas, Dec. 2004.

[16] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transaction on Automatic Control*, vol. 24, no. 6, pp. 843–854, December 1979.

[17] S. Oh, L. Schenato, and S. Sastry, "A hierarchical multiple-target tracking algorithm for sensor networks," in *Proc. of the International Conference on Robotics and Automation*, Barcelona, Spain, April 2005.

[18] S. Oh and S. Sastry, "A polynomial-time approximation algorithm for joint probabilistic data association," in *Proc. of the American Control Conference*, Portland, OR, June 2005.

[19] R. Vidal, A. Chiuso, and S. Soatto, "Observability and identifiability of jump linear systems," in *Proc. of IEEE Conference on Decision and Control*, Las Vegas, NV, Dec. 2002.

[20] M. Johansson, *Piecewise Linear Control Systems - A Computational Approach*, ser. Lecture Notes in Control and Information Sciences no. 284. Heidelberg, Germany: Springer-Verlag, 2002.

[21] I. Beichl and F. Sullivan, "The Metropolis algorithm," *Computing in Science and Engineering*, vol. 2, no. 1, pp. 65–69, 2000.

[22] M. Jerrum and A. Sinclair, "The Markov chain Monte Carlo method: An approach to approximate counting and integration," in *Approximations for NP-hard Problems*, D. Hochbaum, Ed. PWS Publishing, Boston, MA, 1996.

[23] G. Roberts, "Markov chain concepts related to sampling algorithms," in *Markov Chain Monte Carlo in Practice*, ser. Interdisciplinary Statistics Series, W. Gilks, S. Richardson, and D. Spiegelhalter, Eds. Chapman and Hall, 1996.

[24] D. Lerro and Y. Bar-Shalom, "Interacting multiple model tracking with target amplitude feature," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 494–509, 1993.