Reduced Spatio-Temporal Complexity MMPP and Image-Based Tracking Filters for Maneuvering Targets

VIKRAM KRISHNAMURTHY
University of British Columbia
Canada

SUBHRAKANTI DEY
University of Melbourne
Australia

We present reduced-complexity nonlinear filtering algorithms for image-based tracking of maneuvering targets. In image-based target tracking, the mode of the target is observed as a Markov modulated Poisson process (MMPP) and the aim is to compute optimal estimates of the target's state. We present a reduced complexity algorithm in two steps. First, a gauge transformation is used to reexpress the filtering equations in a form that is computationally more efficient for time discretization than naive discretization of the filtering equations. Second, a spatial aggregation algorithm with guaranteed performance bounds is presented for the time-discretized filters. A numerical example illustrating the performance of the resulting reduced-complexity filtering algorithms for a switching turn-rate model is presented.

Manuscript received May 17, 2002; revised June 11 and July 24, 2003; released for publication July 24, 2003.

IEEE Log No. T-AES/39/4/822057.

Refereeing of this contribution was handled by X. R. Li.

Parts of this work appear in the Fusion 2002 conference.

Authors' addresses: V. Krishnamurthy, Dept. of Electrical and Computer Engineering, University of British Columbia, Vancouver, V6T 1Z4, Canada, E-mail: (vikramk@ece.ubc.ca); S. Dey, Dept. of Electrical and Electronic Engineering, University of Melbourne, VIC 3010, Australia, E-mail: (sdey@ee.mu.oz.au).

I. INTRODUCTION

There is significant motivation to develop reduced-complexity filtering algorithms (with explicit performance bounds) for tracking maneuvering targets. Maneuvering target estimation is an important problem in target tracking due to the uncertainty in maneuvers of the target. In a hostile environment a target will try to avoid being tracked by maneuvering in such a way so that its motion is difficult to follow. The idea behind image-based and image-enhanced tracking [1–3] is to use two-dimensional imagery to obtain information about the mode of the target (e.g., oorientation information) apart from conventional measurements. Simulation studies in [1–3] demonstrate that this modal information can lead to marked improvements in the target tracking performance.

As is widely done [1–3], we assume the mode of the target with time is modeled as a finite state Markov chain and the target's trajectory is modeled as a jump Markov linear system. For further motivation on the use of finite state Markov chains and jump Markov linear systems for modelling a maneuvering target, see the classic books of [4, 5]. The image sensor processor response to the modal information is blurred due to the range of the target, weather conditions, etc. Finally, the blurred images are processed by an imager which generates a marked Poisson process according to the noisy state of the Markov chain. In summary the image-based target tracking model involves estimating the state of a jump Markov linear system when the underlying mode (finite state Markov chain) is observed via a multivariate Markov modulated Poisson process (MMPP) [6].

It is well known that if noisy measurements of the target's trajectory were available (rather than noisy modal information as considered here), then estimating the trajectory of the target (jump Markov linear system) is an NP hard problem (requires S^T computational cost for S modes and T data points). There are numerous suboptimal algorithms including particle filters [7] and the interacting multiple model (IMM) algorithm [4]. However, here we focus on the image-based tracking problem where optimal state estimates can be computed with linear complexity in T. In the image-based tracking problem considered here, estimating the target's mode and coordinates involves the following two filtering algorithms.

1) The optimal (minimum mean square error (MMSE)) estimate of the orientation is computed by an MMPP filter (which is essentially a continuous-time hidden Markov model (HMM) filter).

2) The trajectory of the target (modeled as a jump Markov linear system) given the noisy modal measurements is estimated using an image-based

filter. This is a finite dimensional filter (i.e. given by an ordinary differential equation driven by a Poisson observation process) which requires estimates from the MMPP filter [1].

The main contributions of this work are to present reduced complexity temporal and spatial approximations to the above MMPP and image-based filters. The computational cost of the image-based tracking filters is $O(S^2T)$. The time discretization and spatial aggregation proposed here reduce this $O(S^2T)$ complexity when S and T are large as follows.

1) *Temporal Discretization*: We present a computationally efficient time discretization of MMPP filter and the optimal image-based filter. In particular, we use a gauge transformation to convert the filtering equations to deterministic differential equations with random coefficients. Such "robust" transformations have been widely used in stochastic calculus and continuous-time HMM filtering [8]. A major advantage of our approach is that the discretization interval of the resulting filters can be much larger than that required by a standard Euler first order discretization of the filter. This leads to substantial computational savings in the numerical implementation of the MMPP and image-based filters. We present explicit upper bounds on the discretization interval for the two filters.

2) *Spatial Aggregation*: The computation complexity of the time discretized MMPP and image-based filters is a major issue when the state space is large. This happens for example when there are several modes or a large number of targets each with a finite number of maneuver types. We present a novel algebraic methodology (similarity transformation) for reducing the complexity of both filters. Under the natural assumption that the target's dynamics are similar when the target's modes are similar, the underlying Markov chain has a nearly completely decomposable structure [9, 10]. With $\epsilon > 0$ denoting a perturbation parameter signifying the weak interactions between dissimilar target modes, our novel similarity transformation decouples the components of the MMPP filter and image-based filters resulting in low-complexity filtering algorithms with provable $O(\epsilon)$ accuracy for the modal filter and $O(\epsilon^2)$ accuracy for the image filter. In order to prove the existence of the similarity transformation we demonstrate the existence of solutions to certain nonlinear matrix difference equations. The decoupling transformations are a generalization of the results in [11].

Typically, the evolution of the target's mode can be modeled as a nearly completely decomposable Markov chain (NCDMC) if there are several modes with each mode having a large number of finer states (either for higher resolution or for modelling parametric uncertainty), or when there are several independent targets with diagonally dominant transition probability matrices for the evolution of their modes.

The methodology proposed here is sufficiently general to apply to other problems requiring reduced-complexity HMM filtering and estimation of the state of a jump Markov linear system given noisy modal information (e.g., Poisson observations).

II. SIGNAL MODEL

All processes are defined on the probability space $(\Omega, \mathcal{F}, P)$.

Target Model: Let $\{X_t, t \geq 0\}$ be a continuous-time S state Markov chain defined on the state space $\{e_1, e_2, \ldots, e_S\}$ where $e_i \in \mathbb{R}^S$ is the unit vector with 1 in the ith position. (This choice of state space simplifies our subsequent notation). Let the infinitesimal generator or transition rate matrix be denoted by A where $\sum_{j=1}^{S} a_{ij} = 0$, $\forall i \in \{1,2,\ldots,S\}$. Define $P(X_t = i) = p_t^i$, $i \in \{1,2,\ldots,S\}$. The probability distribution $p_t = (p_t^1 \;\; p_t^2 \ldots p_t^S)'$ satisfies the forward equation $dp_t/dt = A'p_t$ where $'$ denotes the transpose operation. The process $\{X_t\}$ denotes the mode or regime of operation of the target and drives the jump linear system for the target dynamics as

$$ds_t = c(X_t)s_t dt + R dw_t. \tag{1}$$

Here $s_t \in \mathbb{R}^L$ denotes the coordinates of the target and the matrices $c(e_i)$, $i = 1,\ldots,S$ are each $L \times L$ matrices. R denotes an arbitrary known matrix. $\mathbf{E}\{s_0\}$ is assumed known and $\{w_t\}$ denotes a Wiener process which is independent of $\{X_t\}$. (If $\mathbf{E}\{s_0\}$ is not known, it can be estimated via an associated stochastic optimization problem (see [2])).

Image Sensor: The image sensor uses two-dimensional imagery to obtain orientation information of the target. (For example, [3] considers the profile of a T-62 tank with three different orientations.) The output of the image sensor is an S-variate MMPP as outlined below (see [3, 12] for excellent expositions). The image sensor is modeled in two steps. First, the appearance of an image frame depends on the mode dependent rates $\lambda^{(i)}$, $i = 1,\ldots,S$. Second, the image sensor processor generates an output statement $\rho_t \in \{e_1,\ldots,e_S\}$. Because of the blurring of the image due to range of the target, weather conditions, etc., the output statement ρ_t is not necessarily the same as the true target mode X_t. This error is modeled probabilistically in terms of the $S \times S$ discernibility matrix $D = (d_{ij})$, where $d_{ij} = P(\rho_t = e_i \mid X_t = e_j)$, $1 \leq i,j \leq S$.

As a result the output of the imager is an S-variate MMPP $N_t = (N_t^{(1)} \;\; N_t^{(2)} \ldots N_t^{(S)})'$ as follows:

$$dN_t^{(i)} = \langle X_t, g^{(i)} \rangle dt + dm_t^{(i)}, \qquad i = 1,2,\ldots,S. \tag{2}$$

Here $N_t^{(i)}$ denotes the number of events with mark i that occur during the interval $[0,t]$, and $g^{(i)} =$

$(g_1^{(i)}\ g_2^{(i)} \dots g_S^{(i)})' = \lambda^{(i)}[d_{i1} \dots d_{iS}]'$, $i = 1,\dots,S$ is the vector of intensities of the ith component of the process N_t. In (2), $\langle \cdot,\cdot \rangle$ denotes the scalar product in $\mathbb{R}^S$ and $m_t^{(i)}$ is an $\mathcal{F}_t$ Poisson martingale where $\mathcal{F}_t$ denotes the sigma algebra $\sigma(X_s, N_s;\ s \le t)$. Denote the observation history as $\mathcal{N}_t^{(i)} = \sigma(N_s^{(i)} : s \le t)$ and $\mathcal{N}_t = \bigvee_{i=1}^S \mathcal{N}_t^{(i)}$ (i.e., $\mathcal{N}_t$ is the sigma algebra generated by $\mathcal{N}_t^{(i)}$, $i = 1,\dots,S$).

The following somewhat technical assumption, which essentially states that the jumps of the Markov chain and Poisson process cannot occur at exactly the same time instant, is widely assumed [3] and is necessary for deriving the filters presented here. The assumption can be relaxed as in [1], however, the resulting filters then need careful accounting of the coquadratic variation of X_t and the individual processes $N_t^{(i)}$.

Assumption 1 We assume that $X_t, N_t^{(1)}, N_t^{(2)}, \dots, N_t^{(S)}$ do not have simultaneous jumps, i.e., $[X, N^{(i)}]_t = 0$, $\forall i \in \{1,2,\dots,S\}$ where $[Y,Z]_t$ denotes the optional coquadratic variation of the process Y,Z. Also, assume that $[N^{(i)}, N^{(j)}]_t = 0$, $\forall i,j \in \{1,2,\dots,S\}$, $i \ne j$.

A. Example. Switching Turn Rate Model

In addition to the image-based tracking problem outlined above, the switching turn rate model described in [13] serves as another motivation for the spatio-temporal reduced-complexity filters proposed here.

Consider tracking an agile maneuvering target in two dimensions. Denote the state vector $s = [x\ \dot{x}\ y\ \dot{y}]'$ with x and y denoting the x,y Cartesian coordinates of the horizontal plane.

A target moving with constant speed can be described by the continuous-time stochastic differential equation (1) with

$$c(X_t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\omega(X_t) \\ 0 & 0 & 0 & 1 \\ 0 & \omega(X_t) & 0 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3}$$

There are two accelerations modeled above, an omni-directional (white) acceleration described by the vector Wiener process w_t, and the structured maneuver acceleration represented by the turn rate processes $\omega(X_t)$ which switches between a finite number of possible turn rates.

The image sensors captures a sequence of images of the target from which orientation information can be obtained through image processing. However, we do not deal directly with orientation of the target but with turn rate. We rely on the strong relationship between orientation and turn rate to allow the transformation from the orientation estimates to the turn rate estimates. For example, the roll angle is intimately related to the magnitude of a turn and thus estimates of the target roll angle provides information for classification of the target turn rate into one of the prespecified bins.

III. IMAGE-BASED AND MODAL FILTERS AND TIME-DISCRETIZATION

Here we present the optimal filtering equations for the image-based and modal filters. Then a gauge transformation is used to transformed into a "robust form" and discretized for numerical implementation. An event-based discretization procedure and hybrid implementation is given in Section IIIB which is computationally efficient.

A. Preliminaries

Results regarding the estimation of the mode X_t given the observation history $\mathcal{N}_t$ were derived in [14] and later generalized in [1]. Here, we briefly state the Zakai form (unnormalized filtered density) for computing the optimal mode estimate $\mathbf{E}\{X_t \mid \mathcal{N}_t\}$ and optimal target trajectory $\mathbf{E}\{s_t \mid \mathcal{N}_t\}$, see [1] for details.

Define the measure $\bar{P}$ and the process $n_t^{(i)}$ such that

$$\frac{dP}{d\bar{P}}\Big|_{\mathcal{F}_t} = \Lambda_t = \prod_{i=1}^S \prod_{0 \le r \le t} \langle g^{(i)}, X_r \rangle (\triangle N_r^{(i)}) \times \exp\left(- \int_0^t [\langle g^{(i)}, X_r \rangle - 1] dr \right) \tag{4}$$

$$n_t^{(i)} = N_t^{(i)} - t. \tag{5}$$

Then the following results hold.

1) Λ_t is a $(\bar{P}, \mathcal{F}_t)$ martingale. and satisfies $\Lambda_t = 1 + \sum_{i=1}^S \int_0^t \Lambda_t (\langle g^{(i)}, X_r \rangle - 1) dn_r^{(i)}$ [14, p. 171].

2) A straightforward invocation of Girsanov's theorem [14, Theorem T2, p. 166] yields that under $\bar{P}$, each of the S components of N_t are independent unit-intensity Poisson processes.

3) Let $\bar{E}$ denote the expectation operator under $\bar{P}$. Let ϕ_t be an $\mathcal{F}_t$ adapted process. Define the unnormalized density $q_t = \bar{\mathbf{E}}\{\Lambda_t \langle \phi, X_t \rangle \mid \mathcal{N}_t\}$—note $q_t \in \mathbb{R}^N$ with nonnegative elements. Then an abstract version of Bayes' theorem states [15, p. 243] that

$$\mathbf{E}\{\langle \phi, X_t \rangle \mid \mathcal{N}_t\} = \frac{\bar{\mathbf{E}}\{\Lambda_t \langle \phi, X_t \rangle \mid \mathcal{N}_t\}}{\bar{\mathbf{E}}\{\Lambda_t \mid \mathcal{N}_t\}} = \frac{\langle \phi, q_t \rangle}{\mathbf{1}_S' q_t}.$$

Notation. Define the $S \times S$ diagonal matrix $B^{(i)}$, $i = 1,\dots,S$ and $(LS) \times (LS)$ matrix C, respectively, as

$$B^{(i)} = \mathrm{diag}[g^{(i)}], \qquad C = \begin{bmatrix} C_{11} & \dots & C_{1L} \\ \vdots & \ddots & \vdots \\ C_{L1} & \dots & C_{LL} \end{bmatrix}$$

where $C_{ij} = \mathrm{diag}[c_{ij}(e_1), \dots, c_{ij}(e_S)]$

and $c_{ij}(e_l)$ denotes the i,j element of $c(e_l)$ defined in (1).

Below we give the Zakai filtering equations for estimating the target mode X_t and the target state s_t. As shown in [1] to estimate the target state s_t one first needs to construct the filtered estimate of $s_t \otimes X_t$ where $\otimes$ denotes the tensor product. Accordingly, define the LS dimensional process

$$r_t = \bar{\mathbf{E}}\{\Lambda_t(s_t \otimes X_t) \mid \mathcal{N}_t\}.$$

THEOREM 1 (Zakai equations and estimates for mode and state) *The unnormalized filtered densities for the target mode and state are given by the following equations called the modal and image filter, respectively:*

Modal Filter:

$$q_t = q_0 + \int_0^t A' q_\tau d\tau + \sum_{i=1}^{S} (B^{(i)} - I) q_\tau dn_\tau^{(i)}. \tag{6}$$

Image Filter:

$$r_t = r_0 + \int_0^t C r_\tau d\tau + \int_0^t (I_L \otimes A') r_\tau d\tau + \int_0^t \sum_{i=1}^{S} (I_L \otimes (B^{(i)} - I)) r_\tau dn_\tau^{(i)} \tag{7}$$

$$\mathbf{E}\{s_t \mid \mathcal{N}_t\} = \frac{\bar{\mathbf{E}}\{\Lambda_t s_T \mid \mathcal{N}_t\}}{\bar{\mathbf{E}}\{\Lambda_t \mid \mathcal{N}_t\}} = \frac{1}{\mathbf{1}' q_t} \mathrm{diag}\underbrace{\{\mathbf{1}_S', \ldots, \mathbf{1}_S'\}}_{L \text{ times}} r_t. \tag{8}$$

Note that the above integrals involving $dn_\tau^{(i)}$ are Steiltjes integrals. Recall $dn_t^{(i)} = dN_t^{(i)} - dt$ and $dN_t^{(i)} = 1$ if an event with mark i occurs at time t, and is zero otherwise. The proof of the above theorem appears in [1, sec. 5.2]. The second equality in (8) follows because $\bar{\mathbf{E}}\{\Lambda_t s_t \mid \mathcal{N}_t\} = \bar{\mathbf{E}}\{\sum_{i=1}^{S} \Lambda_t s_t X_t(i)\} = \mathrm{diag}\{\mathbf{1}_S', \ldots, \mathbf{1}_S'\} r_t$ (where the diagonal matrix is of dimension $L \times LS$) and $\bar{\mathbf{E}}\{\Lambda_t\} = \bar{\mathbf{E}}\{\sum_{i=1}^{S} \Lambda_t X_t(i)\} = \mathbf{1}_S' q_t$.

Remark. The modal filter (6) and image filter (7) can be implemented exactly. Let $\tau_k, k = 1,2,\ldots$ denote the event times of $\{N_t\}$. Then (6) can be written as

$$q_t = q_0 + \int_0^t \left[A' - \left(\sum_{i=1}^{S} (B^{(i)} - I) \right) \right] q_s ds + \sum_{i=1}^{S} (B^{(i)} - I) \sum_{\tau_k < t} q_{\tau_k -}. \tag{9}$$

This leads to the following exact implementation: For $\tau_k \le t < \tau_{k+1}$,

$$q_t = \exp\left[\left(A' - \sum_{i=1}^{S} (B - I) \right) (t - \tau_k) \right] q_{\tau_k}. \tag{10}$$

That is, q_t evolves deterministically between Poisson events. At the event time $t = \tau_{k+1}$, update $q_{\tau_{k+1}}$ as

$$q_{\tau_{k+1}} = q_{\tau_{k+1}-} + \sum_{i=1}^{S} (B - I) q_{\tau_{k+1}-}.$$

However, computing the matrix exponential over the inter-arrival time $\tau_{k+1} - \tau_k$ requires $O(S^3)$ computations at each τ_k which can be expensive for large S. Similarly, an exact implementation of the image filter requires $O(S^3 L^3)$ computations at each τ_k.

An obvious approximation is to precompute and store the matrix exponentials for several possible time intervals (over a uniform grid of possible time intervals). Then one can approximate the matrix exponential of $\tau_{k+1} - \tau_k$ by the precomputed matrix exponential corresponding to the time interval in the grid that is closest to $\tau_{k+1} - \tau_k$. However, for large state spaces and a wide range of Poisson rates $g^{(i)}$, unless a very fine grid is used (which incurs significant memory overheads), the approximation can be quite inaccurate.

In contrast, the gauge transformed filters and their hybrid implementation presented in this section permit several possible implementations that can yield significant savings in memory and computation overheads. For example, the matrix exponentials can be precomputed over a coarse grid and the precomputed matrix exponential corresponding to the time interval in the grid that is closest to the inter-arrival time chosen. Then the gauge transformed filter can be run from the grid time point to the arrival time. The sampling time calculations in Theorem 3 give upper bounds on how coarse the time grid can be while still ensuring that the normalized filtered estimate densities are nonnegative.

B. Gauge Transformation and Time-Discretization

The aim of this section is to transform the Poisson driven stochastic differential equations (6), (7) into linear ordinary differential equations (ODEs) with random coefficients. This is exploited in two ways. 1) It permits the use of standard time discretization techniques for ODEs and computationally efficient numerical implementation. 2) The form of the resulting time-discretized filters are similar to a discrete-time HMM filter. This allows us in Section IV to use spatial aggregation methods developed in [11] which results in reduced-complexity filters.

For $i = 1,\ldots,S$ define

$$\Gamma_t^{(i)} = (B^{(i)})^{N_t^{(i)}} \exp(-(B^{(i)} - I)t), \qquad \Gamma_t = \prod_{i=1}^{S} \Gamma_t^{(i)}. \tag{11}$$

Define the gauge transformed processes $\bar{q}_t$ and $\bar{r}_t$ as

$$\bar{q}_t = \Gamma_t^{-1} q_t, \qquad \bar{r}_t = (I_L \otimes \Gamma_t)^{-1} r_t. \tag{12}$$

Note that since Γ_t is an exponential matrix, its inverse is guaranteed to exist. The following theorem gives ODEs for the dynamics of gauge transformed processes $\bar{q}_t$ and $\bar{r}_t$. The final discretized implementation of the resulting filters is given in (19) and (20). These recursions do not require computing matrix exponentials or inverses.

THEOREM 2 *$\bar{q}_t$ and $\bar{r}_t$ satisfy the linear ODEs*

$$\frac{d\bar{q}_t}{dt} = \Gamma_t^{-1} A' \Gamma_t \bar{q}_t, \qquad \bar{q}_0 = q_0 \tag{13}$$

$$\frac{d\bar{r}_t}{dt} = (I_L \otimes \Gamma_t)^{-1}(I_L \otimes A' + C)(I_L \otimes \Gamma_t)\bar{r}_t, \qquad \bar{r}_0 = r_0. \tag{14}$$

Remark. The term "robust" has widely been used to describe the filtering equations (13) in the case when the observation process was the Markov chain corrupted by Brownian motion [16, 17]. In that case, [18] has proved that because $\bar{q}_t$ satisfies an ODE, it has the robustness property of being Lipschitz continuous (in terms of the sup norm) with respect to the observation trajectory. We use the term robust in the context that the equations are ODEs.

PROOF We prove (14). The proof of (13) is identical and hence omitted. Define $\bar{\Gamma}_t = I_L \otimes \Gamma_t$ and $\bar{\Gamma}_t^{(i)} = I_L \otimes \Gamma_t^{(i)}$. Suppose $\tilde{r}_t = \bar{\Gamma}_t \bar{r}_t$. Then

$$d\tilde{r}_t = (d\bar{\Gamma}_t)\bar{r}_t + \bar{\Gamma}_t \frac{d\bar{r}_t}{dt} dt. \tag{15}$$

Let us compute $d\bar{\Gamma}_t$. Because $\Gamma_t = \prod_{i=1}^S \Gamma_t^{(i)}$, it follows from the product rule of differentiation that

$$d\bar{\Gamma}_t = \sum_{i=1}^S d\bar{\Gamma}_t^{(i)} \prod_{j=1 j\neq i}^S \bar{\Gamma}_t^{(j)}. \tag{16}$$

But because $\bar{\Gamma}_t^{(i)} = I_L \otimes \Gamma_t^{(i)}$ it follows that

$$\begin{aligned} d\bar{\Gamma}_t^{(i)} &= I_L \otimes [(B^{(i)} - I)\Gamma_t^{(i)}(dN_t^{(i)} - dt)] \\ &= (I_L \otimes (B^{(i)} - I))(I_L \otimes \Gamma_t^{(i)})(dN_t^{(i)} - dt). \end{aligned}$$

Substituting into (16) yields

$$d\bar{\Gamma}_t = \sum_{i=1}^S \bar{\Gamma}_t (I_L \otimes (B^{(i)} - I))(dN_t^{(i)} - dt). \tag{17}$$

Finally substituting (17) into (15) yields

$$d\tilde{r}_t = C\tilde{r}_t dt + (I_L \otimes A')\tilde{r}_t dt + \sum_{i=1}^S (I_L \otimes (B^{(i)} - I))\tilde{r}_t dn_t^{(i)}.$$

Therefore $\tilde{r}_t$ is a solution of (7). Solutions of (7) which are continuous on the right with limits on the left are unique [14, chap. VI], hence the result follows.

Discretization Interval for Gauge Transformed Filters: We now consider time discretization of the above gauge transformed filters (13) and (14). Consider a regular time partition $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n < \cdots$ with constant time step $\Delta = t_n - t_{n-1}$. Define the discrete-time observation probability diagonal matrix

$$B_{t_{n+1}} = \mathrm{diag}[b_{t_{n+1}}(1), \ldots, b_{t_{n+1}}(S)] \stackrel{\Delta}{=} \Gamma_{t_{n+1}} \Gamma_{t_n}^{-1}. \tag{18}$$

Note that $B_{t_{n+1}}$ is precomputable since

$$B_{t_{n+1}} = \prod_{i=1}^S [(B^{(i)})^{(N_{t_{k+1}}^{(i)} - N_{t_k}^{(i)})} \exp(-(B^{(i)} - I)\Delta)]$$

and for sufficiently fine discretization $N_{t_{k+1}}^{(i)} - N_{t_k}^{(i)}$ is either 0 or 1.

A first order (Euler) explicit discretization of (13) yields

$$\bar{q}_{t_{n+1}} = \bar{q}_{t_n} + \Delta \Gamma_{t_n}^{-1} A' \Gamma_{t_n} \bar{q}_{t_n}.$$

Multiplying both sides by $\Gamma_{t_{n+1}}$ yields the discrete-time gauge transformed mode filter

$$q_{t_{n+1}} = B_{t_{n+1}}(I + A'\Delta) q_{t_n}. \tag{19}$$

Similarly, a first order explicit discretization of (14) followed by multiplying both sides by $(I_L \otimes \Gamma_t)$ yields the discrete-time gauge transformed image filter

$$r_{t_{n+1}} = (I_L \otimes B_{t_{n+1}})[I_L \otimes (I_S + A'\Delta) + C\Delta] r_{t_n}. \tag{20}$$

Note that for all n, $q_{t_{n+1}}$ should be an S dimensional vector with nonnegative elements. Similarly, consider r_t in (14). If $\bar{r}_0 > 0$ (element wise) then from (14), $\bar{r}_t \geq 0$ for all t. This is easily seen since the right-hand side (RHS) of (14) is 0 when $r_t = 0$, meaning that 0 is an absorbing point and hence the components of r_t can never go negative. Our aim is to determine an upper bound for the size of the time-discretization step Δ in (19) and (20) to ensure nonnegativity of the discretized processes q_{t_n} in (19) and r_{t_n} in (20).

THEOREM 3 *To ensure nonnegativity of the elements of the discretized gauge transformed modal filter q_{t_n} in* (19) *requires the sampling period Δ to satisfy*

$$\Delta \leq \Delta_{\text{mode}}^{\text{gauge}} \stackrel{\Delta}{=} \frac{1}{\max_{j\in\{1,\ldots,S\}} |a_{jj}|}. \tag{21}$$

To ensure nonnegativity of the elements of the discretized gauge transformed image filter r_{t_n} in (20) *requires Δ to satisfy*

$$\Delta \leq \Delta_{\text{image}}^{\text{gauge}} \stackrel{\Delta}{=} \frac{1}{\max_m |a_{jj} + c_{mm}|} \qquad \textit{where} \quad j = m \bmod S. \tag{22}$$

PROOF Using (19) we thus require for any fixed $j \in \{1,\dots,S\}$ that

$$q_{t_{n+1}}(j) = b_{t_{n+1}}(j)\sum_{i=1}^{S}(\delta_{ij} + \Delta a_{ij})q_{t_n}(i) \geq 0. \quad (23)$$

Define the normalized filtered estimate $\hat{q}_{t_n} = q_{t_n}/(\mathbf{1}_S' q_{t_n})$. We can then normalize the above inequality by dividing both sides by $\mathbf{1}_S' q_{t_n} = \sum_{i=1}^{S} q_{t_n}(i)$ which then reads

$$\Delta\left(\sum_{i=1}^{S}\hat{q}_{t_n}(i)a_{ij}b_{t_{n+1}}(j)\right) + \hat{q}_{t_n}(j)b_{t_{n+1}}(j) \geq 0.$$

Since $a_{jj} < 0$ (because A is a generator matrix), the left-hand size of the above inequality is minimized when $\hat{q}_{t_n} = e_j$. With this choice of $\hat{q}_{t_n}$, the condition $\hat{q}_{t_{n+1}}(j) \geq 0$ requires $\Delta \leq 1/|a_{jj}|$. Thus the maximum possible discretization interval to ensure nonnegativity of $\hat{q}_{t_{n+1}}(j)$ and hence $q_{t_{n+1}}(j)$, for all $j \in \{1,\dots,S\}$ satisfies (21).

Note that (20) is of the form $r_{t_{n+1}} = \bar{B}_{t_{n+1}}(I_{LS} + (\bar{A} + C)\Delta)r_{t_n}$ where

$$\bar{B}_{t_{n+1}} = \mathrm{diag}(\bar{b}_{t_{n+1}}(1),\dots,\bar{b}_{t_{n+1}}(LS)) \stackrel{\Delta}{=} I_L \otimes B_{t_{n+1}}$$

and $\bar{A} = I_L \otimes A$ with elements $\bar{a}_{lm}$, $l,m = 1,\dots,LS$. Then similar to the above argument for q_{t_n}, define the normalized estimate $\hat{r}_{t_n} = r_{t_n}/\sum_{l=1}^{LS} r_{t_n}(l)$. We have for each $m = 1,\dots,LS$ that $\hat{r}_{t_{n+1}}(m) \geq 0$ implies that

$$\Delta\sum_{l=1}^{LS}(\bar{a}_{lm} + c_{lm})\hat{r}_{t_n}(l) + \hat{r}_{t_n}(m) \geq 0. \quad (24)$$

The left-hand side of the above inequality is minimized when $\hat{r}_{t_n}(m) = 1$. With this choice we require

$$\Delta(\bar{a}_{mm} + c_{mm}) \geq -1$$

There are two possible cases as follows. 1) If $(\bar{a}_{mm} + c_{mm}) > 0$ then any choice of $\Delta > 0$ suffices. 2) If $(\bar{a}_{mm} + c_{mm}) < 0$ or equivalently $c_{mm} < |\bar{a}_{mm}|$ then $\Delta \leq 1/|\bar{a}_{mm} + c_{mm}|$. Given that $\bar{A} = I_L \otimes A$, it follows that $\bar{a}_{mm} = a_{jj}$ where $j = m \bmod S$. Taking the minimum over all $m \in \{1,\dots,LS\}$ yields the upper bound (22) for Δ.

Discretization Interval for Zakai Filters: Here we consider the Euler discretization of the Zakai (nonrobust) equations (6) and (7). In the interval $t_n < t \leq t_{n+1}$,

$$q_{t_{n+1}} = q_{t_n} + \int_{t_n}^{t_{n+1}} A' q_s ds + \int_{t_n}^{t_{n+1}}\sum_{i=1}^{S}(B^{(i)} - I)q_s dn_s^{(i)}. \quad (25)$$

First order discretization and normalization with $\hat{q}_{t_n} = q_{t_n}/(\mathbf{1}' q_{t_n})$ yields

$$\hat{q}_{t_{n+1}}(j) = \hat{q}_{t_n}(j) + \sum_{j=1}^{S}\Delta a_{ij}\hat{q}_{t_n}(i) + \sum_{i=1}^{S}(g_j^{(i)} - 1)\hat{q}_{t_n}(j)(\Delta N_t^{(i)} - \Delta), \qquad t_n < t \leq t_{n+1}$$

where $\Delta N_t^{(i)} = 1$ if a Poisson event occurs at time t and 0 otherwise. Again considering the case $\hat{q}_{t_n} = e_j$, it follows that $q_{t_{n+1}}(j) \geq 0$ for $j = 1,\dots,S$ requires

$$\Delta \leq \Delta_{\mathrm{mode}}^{\mathrm{Zakai}} \stackrel{\Delta}{=} \min_{j\in\{1,\dots,S\}} \frac{1 + \sum_{i=1}^{S}(g_j^{(i)} - 1)\Delta N_t^{(i)}}{\sum_{i=1}^{S}(g_j^{(i)} - 1) + |a_{jj}|} \quad (26)$$

providing $\sum_{i=1}^{S}(g_j^{(i)} - 1) > a_{jj}$.

Consider now the Zakai equation of the image filter (7). Using similar steps to the above it follows that $r_{t_{n+1}}(m) \geq 0$ for all $m = 1,\dots,LS$ requires

$$\Delta \leq \Delta_{\mathrm{image}}^{\mathrm{Zakai}} \stackrel{\Delta}{=} \min_{m\in\{1,\dots,LS\}} \frac{1 + \sum_{i=1}^{S}(g_j^{(i)} - 1)\Delta N_t^{(i)}}{\sum_{i=1}^{S}(g_j^{(i)} - 1) - (a_{jj} + c_{mm})} \quad j = m \bmod S \quad (27)$$

providing that $\sum_{i=1}^{S}(g_j^{(i)} - 1) > (a_{jj} + c_{mm})$ for $j \in \{1,\dots,S\}$ and $m \in \{1,\dots,LS\}$.

Comparison and Event Driven Discretization: Comparing (21) with (26) or (22) with (27) shows that Δ for the gauge transformed filters are independent of the Poisson rates.

Consider the case when $\Delta N_t^{(i)} = 0$, i.e., no event has occurred at time $t_i \leq t < t_{i+1}$. Assuming $a_{jj} + c_{mm} < 0$ and $g_j^{(i)} > 1$, then comparing the sampling intervals (21) with (26), and (22) with (27), yields

$$\Delta_{\mathrm{mode}}^{\mathrm{gauge}} = \frac{1}{\max_j |a_{jj}|} > \Delta_{\mathrm{mode}}^{\mathrm{Zakai}} = \frac{1}{\max_j \sum_{i=1}^{S}(g_j^{(i)} - 1) + |a_{jj}|}$$

$$\Delta_{\mathrm{image}}^{\mathrm{gauge}} = \frac{1}{\max_m |a_{jj} + c_{mm}|} > \Delta_{\mathrm{image}}^{\mathrm{Zakai}} = \frac{1}{\max_m \sum_{i=1}^{S}(g_j^{(i)} - 1) + |a_{jj} + c_{mm}|}.$$

These strict inequalities show that the gauge transformed mode and image filters will tolerate coarser time discretization than the Euler discretization of the standard Zakai filters. The coarser partition for the gauge transformed filters means that the number of computations can be reduced while still maintaining the inequalities (23) and (24). This feature can be exploited to devise an event driven discretization for the gauge transformed filters driven by N_t as follows. Discretize the filters over the irregular time partition $0 = t_0 < t_1 \cdots < t_k < \cdots$ defined as follows: $t_{k+1} = t_k + \Delta_{t_k}$ where $\Delta_{t_k} = \min[\Delta, t^* - t_k]$ and t^* denotes the first Poisson event after time t_k. Here Δ is typically chosen close to (but less than) $\Delta_{\mathrm{mode}}^{\mathrm{gauge}}$ or $\Delta_{\mathrm{image}}^{\mathrm{gauge}}$. The

resulting gauge transformed filters are computationally cheaper than the discretized Zakai equations. We can compare the complexity of the filters as follows.

Computational Complexity of Event Driven Discretization of Filters:

Exact implementation: First, consider an event-driven discretization of the exact implementation of the modal filter (9). Let N_Δ denote the number of Poisson events that occur in a time interval Δ. Let $\bar{g}$ denote the expected number of events per unit time. Then clearly $\bar{g} = \sum_{i=1}^{S} g^{(i)\prime} p_\infty$ where p_∞ denotes the steady state distribution of X_t, i.e., p_∞ satisfies $p'_\infty A = 0$, $p'_\infty \mathbf{1}_S = 1$. If no event occurs in a period Δ, then this filter requires S^2 computations since (10) is a deterministic update and the matrix exponential in (10) over the interval Δ can be precomputed. If $N_\Delta \geq 1$ events occur in the period Δ and we update the filter N_Δ times, this costs $N_\Delta S^3$ computations because the matrix exponential (see discussion below (9)) needs to be evaluated over a random time interval. Hence the expected complexity in the interval Δ of the exact optimal filter is

$$S^2 P(N_\Delta = 0) + S^3 \sum_{k=1}^{\infty} k P(N_\Delta = k) \leq S^2 + \bar{g}\Delta S^3.$$

For a sample size of length T, there are T/Δ segments, hence the average complexity is

$$S^2 T/\Delta + S^3 \bar{g} T.$$

Gauge transformed filters: Using a similar argument, if an event driven discretization is used for the gauge transformed modal filter, the complexity is upper bounded by $S^2 T/\Delta + S^2 \bar{g} T$. The complexity of the image-based filters can be computed similarly—simply replace S by LS in the above complexity expressions.

Hybrid implementation: The above complexity calculations show that the following hybrid discretization implementation is possible: 1) If no event occurs in the interval Δ, i.e., $N_\Delta = 0$, then use the exact implementation. 2) If $N_\Delta \geq 1$, then use the gauge transformed filters. The average complexity of this hybrid scheme is identical to that of the gauge transformed filter—however, the discretization error is considerably reduced since no approximation is carried out when $N_\Delta = 0$.

IV. REDUCED SPATIAL COMPLEXITY IMAGE AND MODAL FILTERS

Having demonstrated the computational efficiency of the hybrid time-discretization scheme for the image and modal filters in Section III, in this section we describe spatial aggregation methods to further decrease the computational complexity of these filters. During intervals when $N_\Delta = 0$, and the exact implementation of the filters are used, it is straightforward to reduce the dimension of the matrix exponential and the vector q_t by aggregation since the evolution is deterministic. As shown in [19] spatial aggregation methods can be used to obtain reduced-complexity implementation of (10) with arbitrarily accurate approximations. Hence the key issue addressed in this section is the spatial aggregation of the gauge transformed filters.

Throughout, we work with the time-discretized gauge transformed filters (19) and (20). We impose the following nearly decomposable structure on the transition probability matrix of the underlying Markov chain X_t: $A = \bar{A} + \epsilon Q$ where $\bar{A}$ has a block diagonal structure

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & 0 & \cdot & \cdot & 0 \\ 0 & \bar{A}_{22} & 0 & \cdot & \cdot \\ 0 & 0 & \cdot\cdot & \cdot & \cdot \\ \cdot\cdot & \cdot\cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \cdot & \bar{A}_{MM} \end{bmatrix} \tag{28}$$

where $\bar{A}_{ii} \in \mathbb{R}^{s_i \times s_i}$, $\forall i$, $\sum_i s_i = S$, $\epsilon > 0$ is a small perturbation parameter, and $Q \in \mathbb{R}^{S \times S}$. $\bar{A}_{ii}$, $\forall i$ are also infinitesimal generators and $\bar{B}$ has zero row-sums. Denote the state partitions as $\mathcal{S}_1 = (e_1, e_2, \ldots, e_{s_1})$, $\mathcal{S}_2 = (e_{s_1+1}, e_{s_1+2}, \ldots, e_{s_1+s_2})$, and so on. We call the "superstates" $\mathcal{S}_1$, $\mathcal{S}_2$, etc, as macro-states. We assume that A and the individual infinitesimal generators A_{ii}, $\forall i$ represent irreducible Markov chains. The above structure of A implies that the Markov chain X_t has a two time-scale structure—it spends most of its time jumping between states in a particular macro-state, and only rarely jumps between macro-states.

We exploit a transformation method proposed in [19, 11] followed by the decoupling transformation used in [11] to obtain approximate reduced-complexity computations for the target trajectory (state) and modal (state of the underlying Markov chain) estimates. It is shown that for the case where the underlying Markov chain is nearly completely decomposable (with a sufficiently small weak coupling parameter ϵ), one can obtain $O(\epsilon)$ approximations to the MMPP and $O(\epsilon^2)$ approximations to the image-based filters when the Poisson arrival rates and the target dynamics are the same for all the constituent states within a given macro-state but are different from one macro-state to another. This leads to the following assumption (as in [11]).

Assumption 2 The intensities of the Poisson arrival process and the target dynamics are only different from one macro-state to another but are identical for all constituent states within a given macro-state. In other words, $g_l = \bar{g}_m$, $\forall X_t = e_l \in \mathcal{S}_m$. Similarly, $c_{ij}(e_l) = \bar{c}_{ij}(e_m)$ $\forall X_t = e_l \in \mathcal{S}_m$, $i, j \in \{1, 2, \ldots, L\}$, $m \in \{1, 2, \ldots, M\}$.

Notice that Assumption 2 also guarantees that the elements of $B_{t_{n+1}}$ are identical within each macro-state.

REMARK 1 Note that Assumption 2 helps in obtaining substantial reductions in computations for the MMPP modal as well as the target trajectory estimates. It also follows immediately that small perturbations (polynomial in ϵ^2) in the arrival rates and target dynamics do not change the results of this paper, i.e., the reduced-complexity MMPP and image-based filtering algorithms can still be applied to this mildly generalized observation dynamics. However, Assumption 2 can be lifted altogether and one can obtain reduced-complexity algorithms for very general observation processes (see [20] for details), at the expense of reduced-computational savings.

Now, rewrite (19), (20) by using $\alpha_{t_{n+1}} = q'_{t_{n+1}}$, $\beta_{t_{n+1}} = r'_{t_{n+1}}$, as

$$\alpha_{t_{n+1}} = \alpha_{t_n}(I_S + A\Delta)B_{t_{n+1}} \tag{29}$$

$$\beta_{t_{n+1}} = \beta_{t_n}[I_L \otimes (I_S + A\Delta) + C'\Delta](I_L \otimes B_{t_{n+1}}). \tag{30}$$

In the following discussion, we denote the corresponding normalized measures for $\alpha_{t_{n+1}}, \beta_{t_{n+1}}$ as $\hat{\alpha}_{t_{n+1}}, \hat{\beta}_{t_{n+1}}$. In the next two subsections, we describe how one can obtain reduced-complexity computations for $\hat{\alpha}_{t_{n+1}}$ and $\hat{\beta}_{t_{n+1}}$. These computations are of $O(M^2)$ instead of $O(S^2)$ as required by the exact computations. Since typically, $M \ll S$, this implies a large savings in computations. These computations result in an $O(\epsilon)$ approximation to $\hat{\alpha}_{t_n}$ and $O(\epsilon^2)$ approximation to $\hat{\beta}_{t_n}$.

A. Examples

In this section, we consider two examples where the Poisson arrival process and the target dynamics are modulated by underlying nearly completely decomposable irreducible Markov chains. These chains have several macro-states with each macro-state having numerous states (for finer resolution and/or parametric uncertainty), whereas the Poisson arrival rates and the target dynamics are identical for all states within the same macro-state but can be different from one macro-state to another. Of course, one can consider small perturbations of such rates and target dynamics in our current framework. The first example considers only one target whereas the second example examines how our method can be applied to the case of multiple independent targets.

Single Target: Assume that the target dynamics are given by (1) where $c(X_t)$ assumes the switching turn rate model (3). Assume that the turn rate has $S = 361$ possible values: $\omega(X_t) \in \{-\pi + n\pi/180, n = 0,1,\ldots,360\}$. Depending on the resolution of the imager and subsequent processing, the turn rate could be quantized to a much smaller set of possible values, say $M = 21$ possible values $\{-\pi + n\pi/10, n = 0,1,\ldots,20\}$. Due to the underlying physics of the target model, the target will switch between these aggregated 20 states much less frequently than than between a small subset of the 360 possible values. Hence a nearly completely decomposable Markov chain with transition probability matrix $A = \bar{A} + \epsilon Q$ (see (28)) is a natural model for this example.

The imager output is an S-variate MMPP with the underlying irreducible NCDMC having S states with M macro-states, where the ith macro-state has s_i number of states such that $\sum_i s_i = S$. The imager output is described by (2). Typically the response of the imager would be similar for all the states within a macro-state. In our example, the Poisson arrival rates $g^{(i)}$, $i = 1,2,\ldots,S$ satisfy Assumption 2. However, our algorithm can be readily applied to the case where the Poisson arrival rates are within $O(\epsilon^2)$ for all states within a given macro-state and are substantially different from one macro-state to another (see Remark 1). Thus one can use the reduced-complexity image filtering and target trajectory algorithms (to be developed in the following section). We present simulation studies involving a single target example (with a univariate Poisson process for simplicity) in Section V.

Multiple Independent Targets: Consider the case of having M_T independent targets where the jth target, $(j = 1,2,\ldots,M_T)$, evolves according to a jump linear system modulated by a nearly completely decomposable irreducible Markov chain X_t^j. Assume these M_T Markov chains are statistically independent. Assume that all the targets have coordinate state spaces of dimension L, but note that the underlying Markov chains may have different states and hence possibly different number of macro-states, with the jth chain evolving according to the transition matrix A^j. In this case, one can write the combined target dynamics as

$$\begin{bmatrix} ds_t^1 \\ ds_t^2 \\ \vdots \\ ds_t^{M_T} \end{bmatrix} = \begin{bmatrix} c(X_t^1) & 0 & \cdots & 0 \\ 0 & c(X_t^2) & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & c(X_t^{M_T}) \end{bmatrix} \begin{bmatrix} s_t^1 \\ s_t^2 \\ \vdots \\ s_t^{M_T} \end{bmatrix} dt + R_T \begin{bmatrix} dw_t^1 \\ dw_t^2 \\ \vdots \\ dw_t^{M_T} \end{bmatrix} \tag{31}$$

where the superscripts $1,2,\ldots,M_T$ refer to the corresponding targets and R_T is an $LM_T \times LM_T$ dimensional matrix. Note that the vector process $(X_t^1, X_t^2, \ldots, X_t^{M_T})$ is merely a Markov chain with transition probability matrix $A^1 \oplus A^2 \oplus \cdots \oplus A^{M_T}$ (see

[21] for a proof) where $\oplus$ denotes the Kronecker sum operator, defined for two matrices $A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{m\times m}$ as

$$A \oplus B = (I_m \otimes A) + (B \otimes I_n) \tag{32}$$

where $\otimes$ denotes Kronecker product.

The imager output is the sum of the Markov modulated Poisson arrival processes (for each dimension of the multivariate process). That is, the imager output is given (for each $l = 1,2,\ldots,S$)

$$dN_t^{(l)} = \sum_{j=1}^{M_T} N_t^{(l)}(j) = \sum_{j=1}^{M_T} \langle X_t^{(j)}, g^{(l,j)} \rangle dt + dm_t^{(l,j)}$$

where $N_t^{(l)}(j)$ denotes the arrival process for the jth target and lth dimension of the multivariate Poisson process and $g^{(l,j)}$ is the vector of intensities of the lth component of the process $N_t^{(l)}(j)$ (driven by the Markov chain $\{X_t^{(j)}\}$). The combined superimposed arrival process $\sum_{j=1}^{M_T} N_t^{(l)}(j)$ is also an MMPP with an intensity vector $g^{(i)}(1) \oplus g^{(i)}(2) \oplus \cdots \oplus g^{(i)}(M_T)$ (see [21] for a proof), where $g^{(i)}(j)$ is the corresponding intensity vector for the jth target.

The corresponding gauge transformed discretized MMPP and image filter equations can be obtained by appropriately modifying (19), (20) with substitution of $B_{t_{n+1}}, A, C$ by their corresponding large dimensional counterparts. It can be shown that the large dimensional state transition probability matrix $A^1 \oplus A^2 \oplus \cdots \oplus A^{M_T}$ is also nearly completely decomposable if every A^j, $j \in \{1,2,\ldots,M_T\}$ is nearly completely decomposable. In order to show this, one needs to prove the property that $A^i \oplus A^j$ is nearly completely decomposable if both A^i, A^j are nearly completely decomposable. This follows fairly easily from the definition of Kronecker sum (32). The number of macro-states in the Kronecker sum can be seen to be the minimum of the number of macro-states in the individual matrices. One can then repeatedly apply this result to prove the nearly completely decomposable property for the sum $A^1 \oplus A^2 \oplus \ldots \oplus A^{M_T}$. The fact that the individual blocks in the block diagonal part of this sum are irreducible generator matrices follows directly from the assumption that the individual matrices A^j have this property.

With the assumption that the individual target dynamics and Poisson arrival rates satisfies Assumption 2, the combined target dynamics and the superimposed Poisson arrival process also satisfies this assumption. Therefore, the reduced-complexity MMPP filtering and image filtering results of the following section can be applied to the case of multiple independent targets. Since the combined multitarget has a large dimensional underlying Markov chain, the substantial savings to be obtained by using our reduced-complexity estimation algorithms make the significance of our algorithm even more compelling.

REMARK 2 Note that the Assumption 2 only simplifies the computations of the reduced-complexity modal and trajectory estimates by resulting in substantial reduction in computation. The multiple independent target tracking problem (with or without this assumption) is not a completely decomposable problem because of the fact that the observation equation involves the superposition of the various individual MMPPs. While Assumption 2 helps in reducing computational complexity, it certainly does not imply that by making this assumption, we are only interested in obtaining target level estimation. In the case of a single target, we are able to obtain state estimates at the macro-state and micro-state level. In the case of multiple independent targets also, we are able to obtain estimates of the macro-states (superstates) and micro-states for each target. Note also that as pointed out in Remark 1, this assumption can be lifted and reduced-complexity algorithms can be obtained for general observation equations, but with less computational savings (see [20]).

B. Reduced-Complexity Modal Estimation for MMPP Filter

Denote $[\zeta_n \eta_n] = \alpha_{t_n}[W_1 W_2]$ where $W_1 \in \mathbb{R}^{S\times M}$ is given by

$$W_1 = \begin{bmatrix} 1_{s_1} & 0 & 0 & 0 \\ 0 & 1_{s_2} & \cdot & \cdot \\ \cdot & 0 & \cdot & 0 \\ 0 & \cdot & \cdot & 1_{s_M} \end{bmatrix}.$$

Similarly, $W_2 \in \mathbb{R}^{S\times(S-M)}$ is a block diagonal matrix with the ith diagonal block $W_2^{(i)} \in \mathbb{R}^{s_i \times s_i - 1}$ is given by

$$W_2^{(i)} = \begin{bmatrix} 0 & \ldots & 0 \\ & I_{s_i-1} & \end{bmatrix}.$$

Note that for these choices $[W_1\ \ W_2]$ is nonsingular and has an inverse $\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$ where V_1, V_2 are also simple matrices independent of the system parameters. Notice an obvious fact that $\zeta_n \in \mathbb{R}^M$ is the aggregate version of the unnormalized measure α_{t_n}. Also, similar relationships can be established between the normalized measures $\hat{\zeta}_n, \hat{\eta}_n$ with $\hat{\alpha}_{t_n}$. It is easy to show that $\hat{\alpha}_{t_n} = \hat{\zeta}_n V_1 + \hat{\eta}_n V_2$.

Also, define

$$\tilde{A}_1^n = V_1(I + \bar{A}\Delta)B_{t_{n+1}}W_1, \qquad \tilde{Q}_1^n = V_1(Q\Delta)B_{t_{n+1}}W_1$$

$$\tilde{A}_2^n = V_1(I + \bar{A}\Delta)B_{t_{n+1}}W_2, \qquad \tilde{Q}_2^n = V_1(Q\Delta)B_{t_{n+1}}W_2$$

$$\tilde{C}_1^n = V_2(I + \bar{A}\Delta)B_{t_{n+1}}W_1, \qquad \tilde{D}_1^n = V_2(Q\Delta)B_{t_{n+1}}W_1$$

$$\tilde{C}_2^n = V_2(I + \bar{A}\Delta)B_{t_{n+1}}W_2, \qquad \tilde{D}_2^n = V_2(Q\Delta)B_{t_{n+1}}W_2.$$

In fact, under Assumption 2, one can show that $C_1^n = 0$, $\forall n$. Using these definitions and notations, one

can then write the recursion (29) in terms of $[\zeta_n \ \ \eta_n]$ and then decouple η_n from ζ_n by using the following decoupling transformation:

$$[\bar{\zeta}_n \quad \bar{\eta}_n] = [\zeta_n \quad \eta_n]\begin{bmatrix} I_M & J_n \\ 0 & I_{S-M} \end{bmatrix} \tag{33}$$

where (see [11] for more details)

$$[\tilde{A}_1^n + \epsilon(\tilde{Q}_1^n - J_n\tilde{D}_1^n)]J_{n+1} = J_n(\tilde{C}_2^n + \epsilon\tilde{D}_2^n) - (\tilde{A}_2^n + \epsilon\tilde{Q}_2^n) \tag{34}$$

where for a sufficiently small ϵ, one can show that $\{J_n\}$ is a sequence of uniformly bounded matrices (see [11] for a proof). Thus one can expand J_n in terms of a power series of ϵ, i.e., $J_n = J_n(0) + \epsilon J_n(1) + \cdots$. One can also show that $J_n(0) \to J(0)$ as $n \to \infty$ where $J(0) = V_1\bar{A}W_2(V_2\bar{A}W_2)^{-1}$. Finally, one can write $O(\epsilon^2)$ approximations for $\hat{\zeta}_n, \hat{\eta}_n$ (hence $\hat{\alpha}_n$) as a result of this decoupling.

Note that the approximations are written with a ˜ sign, that is, an approximation for normalized measures like $\hat{\zeta}_n$ is written as $\tilde{\zeta}_n$, whereas the corresponding approximation for an unnormalized measure like ζ_n is written as $\tilde{\zeta}_n^u$ where the superscript u stands for "unnormalized." We now present our main result regarding the reduced-complexity $O(\epsilon)$ computations for the time-discretized gauge transformed filter $\alpha_{t_{n+1}}$.

THEOREM 4 *Consider the MMPP described by* (2) *where the modulating Markov chain has a nearly completely decomposable structure* $A = \bar{A} + \epsilon Q$ *and the transition probability matrices* $A, \bar{A}_{ii}$ *are irreducible* $\forall i$. *Suppose Assumption* 2 *holds. Then, for a sufficiently small* ϵ, *there exists a large enough but finite* n_0 *such that for* $n \geq n_0$, *an* $O(\epsilon)$ *approximation for the discretized approximate filter* (29), *denoted by* $\tilde{\alpha}_{t_n}$ *can be described by the following recursions:*

$$\tilde{\zeta}_{n+1}^u = \tilde{\zeta}_n[\tilde{A}_1^n + \epsilon(\tilde{Q}_1^n - J(0)\tilde{D}_1^n)], \qquad \tilde{\zeta}_{n_0-1} = \hat{\zeta}_{n_0-1}$$
$$\tilde{\zeta}_{n+1} = \frac{1}{[\tilde{\zeta}_{n+1}^u 1_M]}\tilde{\zeta}_{n+1}^u \tag{35}$$

$$\tilde{\eta}_n = -\tilde{\zeta}_n J(0), \qquad \tilde{\eta}_{n_0-1} = \hat{\eta}_{n_0-1} \tag{36}$$

$$\tilde{\alpha}_{t_n} = \tilde{\zeta}_n V_1 + \tilde{\eta}_n V_2, \qquad \tilde{\alpha}_{t_{n_0-1}} = \hat{\alpha}_{t_{n_0-1}}. \tag{37}$$

PROOF For a similar proof, and various other inequalities that determine the range of a sufficiently small ϵ, see [11].

REMARK 3 Notice that in the above approximations, one can actually obtain an $O(\epsilon^2)$ approximation to ζ_n (or $\hat{\zeta}_n$) but an $O(\epsilon)$ approximation to $\hat{\eta}_n$ and hence an $O(\epsilon)$ approximation to $\hat{\alpha}_{t_n}$. One can obtain an $O(\epsilon^2)$ approximation to $\hat{\eta}_n$ and hence $\hat{\alpha}_{t_n}$ at a slightly higher computational cost. For the above recursions, notice that the computational cost is $O(M^2)$ per time instant (as opposed to $O(S^2)$ per time instant for exact computations).

C. Reduced-Complexity Estimation for Target Trajectory

For the purpose of this section define

$$[\xi_n^{(1)} \quad \chi_n^{(1)} \ldots \xi_n^{(L)} \quad \chi_n^{(L)}] = \beta_{t_n}(I_L \otimes [W_1 \quad W_2]). \tag{38}$$

One can then use a generalized decoupling transformation to define

$$[\bar{\xi}_n^{(1)} \quad \bar{\chi}_n^{(1)} \ldots \bar{\xi}_n^{(L)} \quad \bar{\chi}_n^{(L)}] = [\xi_n^{(1)} \quad \chi_n^{(1)} \ldots \xi_n^{(L)} \quad \chi_n^{(L)}] \times \begin{bmatrix} \begin{pmatrix} I_S & \bar{J}_n^1 \\ 0 & I_{S-M} \end{pmatrix} & 0 & \cdot & 0 \\ 0 & \begin{pmatrix} I_S & \bar{J}_n^2 \\ 0 & I_{S-M} \end{pmatrix} & 0 & \cdot \\ 0 & 0 \cdot & \cdot & \cdot \\ \cdot\cdot & \cdot\cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \begin{pmatrix} I_S & \bar{J}_n^L \\ 0 & I_{S-M} \end{pmatrix} \end{bmatrix}. \tag{39}$$

Rewriting the recursion (30) in terms of the decoupled variables, one gets

$$[\bar{\xi}_{n+1}^{(1)} \quad \bar{\chi}_{n+1}^{(1)} \ldots \bar{\xi}_{n+1}^{(L)} \quad \bar{\chi}_{n+1}^{(L)}] = [\bar{\xi}_n^{(1)} \quad \bar{\chi}_n^{(1)} \ldots \bar{\xi}_n^{(L)} \quad \bar{\chi}_n^{(L)}] \times \begin{bmatrix} T_{n+1}^{11} & T_{n+1}^{12} & \cdot & \cdot & T_{n+1}^{1L} \\ T_{n+1}^{21} & T_{n+1}^{22} & \cdot & \cdot & T_{n+1}^{2L} \\ \cdot\cdot & \cdot\cdot & \cdot & \cdot & \cdot \\ T_{n+1}^{L1} & T_{n+1}^{L2} & \cdot & \cdot & T_{n+1}^{LL} \end{bmatrix} \tag{40}$$

where

$$\begin{aligned} T_{n+1}^{ii} &= \begin{pmatrix} I_S & -\bar{J}_n^i \\ 0 & I_{S-M} \end{pmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}(I_S + A\Delta + \Delta C_{ii}) \\ &\quad \times B_{t_{n+1}}[W_1 \quad W_2]\begin{pmatrix} I_S & \bar{J}_{n+1}^i \\ 0 & I_{S-M} \end{pmatrix} \\ T_{n+1}^{ij} &= \begin{pmatrix} I_S & -\bar{J}_n^i \\ 0 & I_{S-M} \end{pmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}\Delta C_{ji} \\ &\quad \times B_{t_{n+1}}[W_1 \quad W_2]\begin{pmatrix} I_S & \bar{J}_{n+1}^j \\ 0 & I_{S-M} \end{pmatrix}, \qquad i \neq j \end{aligned} \tag{41}$$

and after denoting $G_{n+1}^{ii} = \Delta C_{ii}B_{t_{n+1}}$, $G_{n+1}^{ij} = \Delta C_{ji}B_{t_{n+1}}$, $\bar{J}_n^i$, $i = 1,2,\ldots,L$ satisfy the following recursive equations (under Assumption 2):

$$\begin{aligned} &[V_1(I_S + A\Delta)B_{t_{n+1}}W_1 + V_1G_{n+1}^{ii}W_1]\bar{J}_{n+1}^i \\ &\quad = \bar{J}_n^i[V_2(I_S + A\Delta)B_{t_{n+1}}W_2 \\ &\qquad + V_2G_{n+1}^{ii}W_2 + V_2(I_S + A\Delta)B_{t_{n+1}}W_1\bar{J}_{n+1}^i] \\ &\qquad - V_1(I_S + A\Delta)B_{t_{n+1}}W_2 \\ &(V_1G_{n+1}^{ij}W_1)\bar{J}_{n+1}^j = \bar{J}_n^i(V_2G_{n+1}^{ij}W_2). \end{aligned} \tag{42}$$

Under Assumption 2, one can show that the first order approximations to $\bar{J}_n^i$, $i = 1,2,\ldots,L$ (i.e., when $\epsilon = 0$) denoted by $\bar{J}_n^i(0) \to J(0)$, $\forall i$ as $n \to \infty$ (this result is similar to that in the previous subsection and the case presented in [11]). Hence, for a sufficiently small ϵ, using analyticity arguments, $\{J_n^i\}$, $\forall i$ can be shown to be sequences of uniformly bounded sequences of matrices which can be approximated by $J(0)$ for a sufficiently large n. Note that in [11], we provide explicit inequalities to obtain ranges for ϵ to prove a similar uniform boundedness result. In this paper, we refrain from obtaining such results for two reasons: 1) these inequalities are quite complicated, and 2) solving them to obtain the desired range of ϵ can be quite cumbersome as well. Some simulation results are presented to illustrate that one can indeed use this approximation in practice to obtain reduced-complexity computations for the image-based filter recursions (30). In fact, one can obtain $O(\epsilon^2)$ recursions to the normalized measures $\hat{\xi}_n^{(1)}, \hat{\xi}_n^{(2)}, \ldots, \hat{\xi}_n^{(L)}$ and hence the normalized measure $\hat{\beta}_{t_n}$ for a sufficiently large n. This result is summarized in the following theorem. The proof is analogous to that of Theorem 4 and is omitted.

THEOREM 5 *Consider the target model given by* (1) *and the MMPP given by* (2) *with an underlying Markov chain that is nearly completely decomposable with* $A = \bar{A} + \epsilon Q$ *where A,* $\bar{A}_{ii}$, $\forall i = 1,2,\ldots,M$ *satisfy the irreducibility conditions. Suppose also that Assumption* 2 *holds. Then there exists a sufficiently large integer* $n_1 > 0$ *such that for* $n \geq n_1$, *one can obtain* $O(\epsilon^2)$ *approximations to the normalized measures* $\hat{\xi}_n^{(i)}$ *and* $\hat{\beta}_{t_n}$ *(and hence to the gauge transformed discretized target trajectory estimator* $\hat{r}_{t_n}$*) using the following approximate computations denoted by* $\tilde{\xi}_n^{(i)}, \tilde{\beta}_n$ *which are given by (for* $n \geq n_1$*) (with the unnormalized estimates denoted with a* $\check{}$*)* $(i,j = 1,2,\ldots,L)$

$$\begin{aligned}
\check{\xi}_{n+1}^{(i)} &= \tilde{\xi}_n^{(i)}(V_1[(I_S + \bar{A}\Delta)B_{t_{n+1}} + G_{n+1}^{ii}]W_1 \\
&\quad + \epsilon(Q_n^1 - J(0)D_n^1)) + \sum_{j \neq i} \tilde{\xi}_n^j (V_1 G_{n+1}^{ji} W_1) \\
\tilde{\xi}_{n+1}^{(i)} &= \check{\xi}_{n+1}^{(i)} / (\tilde{\zeta}_{n+1}^u 1_M) \\
\tilde{\beta}_{n+1} &= [\tilde{\xi}_{n+1}^{(1)} \ldots \tilde{\xi}_{n+1}^{(L)}](I_L \otimes 1_M).
\end{aligned} \tag{43}$$

REMARK 4 Note that the order of approximation for $\tilde{\beta}_n$ is $O(\epsilon^2)$ as opposed to an $O(\epsilon)$ approximation to the MMPP filter as described in Theorem 4.

V. NUMERICAL EXAMPLES

In this section, we present numerical examples to illustrate the performance of our reduced-complexity mode estimation and image enhanced target trajectory estimation schemes.

These simulations are carried out with the following choice of parameters for the signal model described in Section II. Throughout we consider a uni-variate MMPP observation process. The MMPP process was simulated using the thinning algorithm, see [22, p. 77]. The filters implemented are the time-discretized gauge transformed modal filter (19) and gauge transformed image filter (20).

We simulated an underlying NCDMC with a generator matrix $A = \bar{A} + \epsilon Q$ where

$$\bar{A} = \begin{bmatrix}
-0.7 & 0.45 & 0.25 & 0 & 0 & 0 & 0 & 0 \\
0.2 & -0.45 & 0.25 & 0 & 0 & 0 & 0 & 0 \\
0.24 & 0.35 & -0.59 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -0.8 & 0.8 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.6 & -0.6 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -0.9 & 0.75 & 0.15 \\
0 & 0 & 0 & 0 & 0 & 0.2 & -0.45 & 0.25 \\
0 & 0 & 0 & 0 & 0 & 0.55 & 0.15 & -0.70
\end{bmatrix} \tag{44}$$

$$\bar{Q} = \begin{bmatrix}
0.1 & 0.1 & -1.0 & 0.6 & 0.05 & 0.05 & 0.05 & 0.05 \\
0.05 & 0.05 & -0.9 & 0.5 & 0.05 & 0.05 & 0.1 & 0.1 \\
0.01 & 0.01 & -0.4 & 0.2 & 0.05 & 0.05 & 0.04 & 0.04 \\
0.02 & 0.42 & 0.01 & 0.01 & -0.61 & 0.025 & 0.1 & 0.025 \\
0.45 & 0.01 & 0.4 & -1.0 & 0.01 & 0.1 & 0.01 & 0.02 \\
0.01 & 0.05 & 0.01 & 0.01 & 0.05 & 0.01 & -0.15 & 0.01 \\
0.03 & 0.01 & 0.03 & 0.04 & 0.01 & 0.01 & 0.01 & -0.14 \\
0.01 & 0.05 & 0.01 & 0.01 & 0.05 & -0.16 & 0.01 & 0.02
\end{bmatrix}.$$

Clearly, for this example, $S = 8$, $M = 3$. The Markov modulated Poisson arrival rates are given by

$$g = (10\ \ 10\ \ 10\ \ 50\ \ 50\ \ 110\ \ 110\ \ 110)'$$

where we have chosen $g(1) = g(2) = g(3)$, $g(4) = g(5)$ and $g(6) = g(7) = g(8)$. Similarly, the target trajectory is approximated by an Euler discretization of (1) by

$$s(k) = s(k-1) + \Delta C(X_k)s(k-1) + Rw_k \tag{45}$$

where $s(k) \stackrel{\Delta}{=} (x\ \ \dot{x}\ \ y\ \ \dot{y}) \in \mathbb{R}^4$ which is also known as the switching turn rate model of a maneuvering target in two dimensions x, y. The sampling time interval Δ is taken to be 0.001 s, w_k is a zero mean white Gaussian sequence with covariance $Q = \sigma^2 I_2$. The matrix R is given by

$$R = \begin{bmatrix}
0.5\Delta^2 & 0 \\
\Delta & 0 \\
0 & 0.5\Delta^2 \\
0 & \Delta
\end{bmatrix}.$$

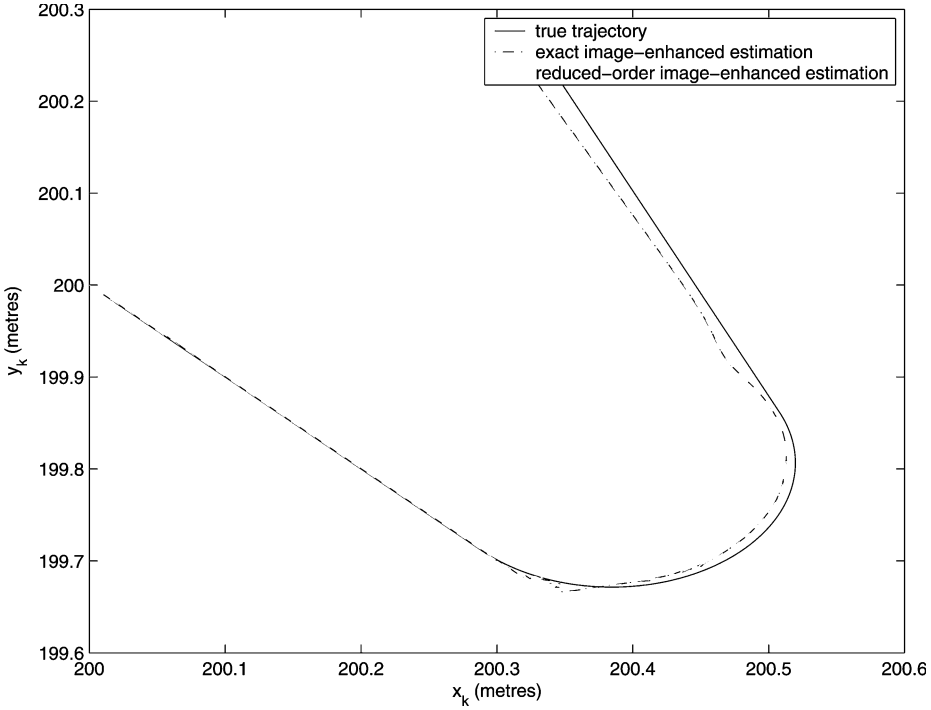

Fig. 1. Low-complexity image-enhanced target trajectory estimation, $\sigma^2 = 0.01$.

Here w_k models an omni-directional acceleration process whereas

$$C(1) = C(2) = C(3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

describe a constant velocity process for the modes 1,2,3. Similarly $C(4) = C(5)$ and $C(6) = C(7) = C(8)$ describe two different turn rate processes with

$$C(4) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\omega_1 \\ 0 & 0 & 0 & 1 \\ 0 & \omega_1 & 0 & 0 \end{bmatrix}$$

and

$$C(6) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -\omega_2 \\ 0 & 0 & 0 & 1 \\ 0 & \omega_2 & 0 & 0 \end{bmatrix}$$

with $\omega_1 = \pi/3$ rad/s and $\omega_2 = -\pi/3$ rad/s. The simulations are carried out with a 10000 point Markov chain (i.e., over a period of 10 s). The results are averaged over 10 simulation trials.

Table I shows the average approximation error between the exact aggregate modal filter ξ_k and the reduced-order approximate aggregate modal filter $\tilde{\xi}_k$ measured as $1/T \sum_{k=1}^{T} \sqrt{\|\xi_k - \tilde{\xi}_k\|^2}$ against various values of ϵ between 0.001 and 0.15. Similarly, Table II shows the average relative approximation error in image-enhanced trajectory estimation between the exact filter and the reduced-order filter.

TABLE I
Average Approximation Error in Low-Complexity Aggregate Modal Filtering

ϵ	Average approximation error in aggregate modal filtering
0.001	1.0594×10^{-6}
0.005	1.042×10^{-5}
0.01	4.204×10^{-5}
0.05	1.554×10^{-4}
0.1	6.6468×10^{-4}
0.15	1.3×10^{-3}

TABLE II
Average Relative Error in Low-Complexity Image-Enhanced Filtering

ϵ	Average relative error in image enhanced filtering
0.001	2.3929×10^{-10}
0.005	4.361×10^{-9}
0.01	1.097×10^{-7}
0.05	4.387×10^{-7}
0.1	2.8×10^{-6}
0.15	3.2872×10^{-5}

Figures 1 and 2 show snapshots of the true trajectory (solid line), the estimated trajectory (the "dash-dotted" line) according to the exact image-enhanced filter and the estimated trajectory (the "dashed" line) according to the reduced-order image enhanced filter, for $\sigma^2 = 0.01$ and $\sigma^2 = 0.1$,

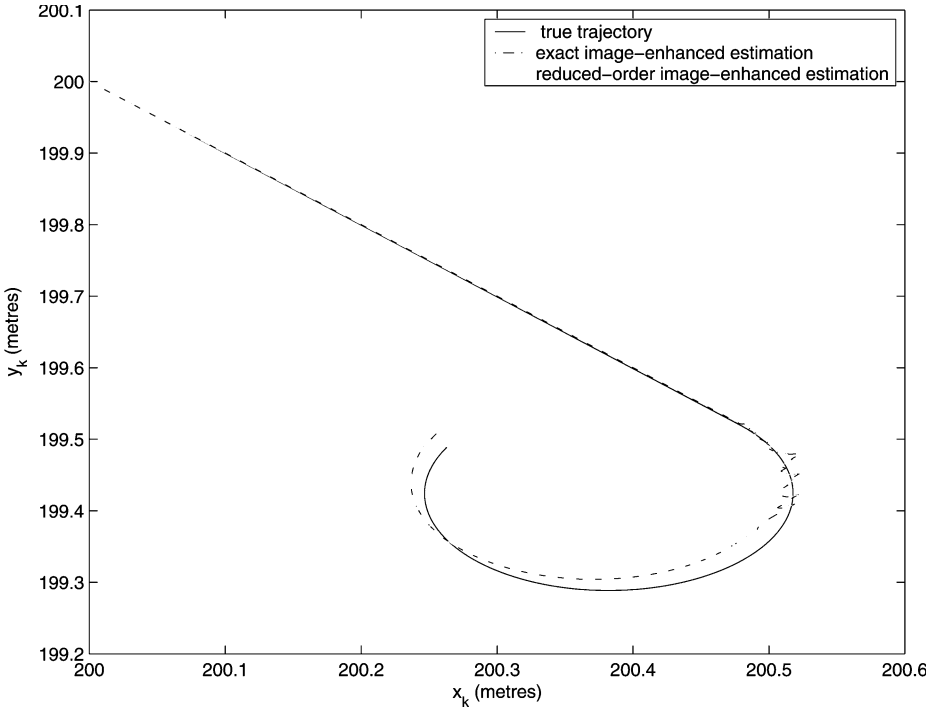

Fig. 2. Low-complexity image-enhanced target trajectory estimation, $\sigma^2 = 0.1$.

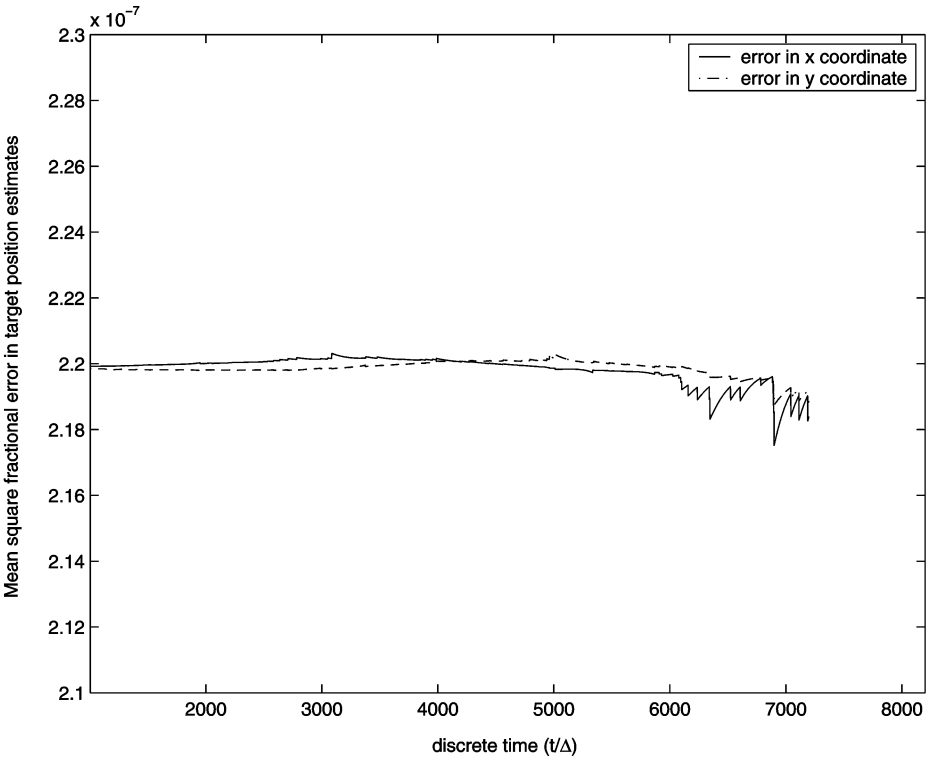

Fig. 3. Mean square fractional error in target position estimates for x and y coordinates, $\sigma^2 = 0.01$.

respectively (with $\epsilon = 0.1$). The estimated paths plotted in Fig. 1 and 2 are computed by averaging over 20 independent trials with the same fixed true trajectory. It is clear that the filters perform quite well with the exact filter and the reduced-order filter being almost indistinguishable. Note that although here we assume the "block"-structured measurement dynamics (g and C matrices being dependent only on the macro-state of the Markov chain), small perturbations of such dynamics will result in similar performances of the reduced-order filters (as commented on in Remark 1). As might be expected, in numerical studies (not presented here) we found that as the variance σ^2 of w_k in (45) is increased, the performance of the filters

are degraded, resulting in larger mean square tracking errors.

Unlike the Kalman filtering case, it is not possible to explicitly compute a "variance ellipse" for the MMPP (modal) filter and image-based filter. Thus in Fig. 3, we plot the estimated mean square fractional error of the target position estimates in x and y coordinates over 8200 discrete-time samples (8.2 s in real time) averaged over 2000 trials, for $\epsilon = 0.1$. The mean square fractional error is defined as

$$\mathbf{E}\left\{1 - \frac{\tilde{\beta}_k^i}{s^i(k)}\right\}^2, \qquad s^i(k) \neq 0$$

where $s^i(k)$ and $\tilde{\beta}_k^i$ denote the ith component of the vectors $s(k)$ and $\tilde{\beta}_k$ defined in (45) (43), respectively, $i = 1,3$ for the x and y coordinates, respectively.

As can be seen from Fig. 3, the estimated mean square fractional error is quite small, implying that the reduced-complexity filters perform satisfactorily.

VI. CONCLUSIONS AND EXTENSIONS

We have presented numerically efficient temporal discretization and spatial aggregation methods for the filters arising in image-based target tracking. The temporal discretization was achieved by using a gauge transformation. The spatial aggregation was achieved by using an algebraic transformation that decouples the components of the filters, leading to provable performance bounds. In future work it would be worthwhile to extend the spatial aggregation procedure to more general observation processes.

REFERENCES

[1] Kirshnamurthy, V., and Elliott, R. (1997)
Filters for estimating Markov modulated Poisson processes and image based tracking.
Automatica, **33** (May 1997), 821–833.

[2] Dufour, F., and Bertrand, P. (1996)
An image based filter for discrete-time Markovian jump linear systems.
Automatica, **32** (1996), 241–247.

[3] Sworder, D. D., and Hutchins, R. G. (1989)
Image-enhanced tracking.
IEEE Transactions on Aerospace and Electronic Systems, **25** (Sept. 1989), 701–709.

[4] Bar-Shalom, Y., and Li, X. (1995)
Multitarget Multisensor Tracking: Principles and Techniques.
Storrs, CT: YBS Publishing, 1995.

[5] Blackman, S., and Popoli, R. (1999)
Design and Analysis of Modern Tracking Systems.
Boston: Artech House, 1999.

[6] Allam, S., Dufour, F., and Bertrand, P. (2001)
Discrete-time estimation of a Markov chain with marked point process observations, application to Markovian jump filtering.
IEEE Transactions on Automatic Control, **46** (June 2001), 903–908.

[7] Doucet, A., Gordon, N., and Krishnamurthy, V. (2001)
Particle filters for state estimation of jump Markov linear systems.
IEEE Transactions on Signal Processing, **49** (2001), 613–624; also technical report CUED-F-INFENG TR.359, University of Cambridge, 1999.

[8] Bensoussan, A. (1992)
Stochastic Control of Partially Observable Systems.
Cambridge, UK: Cambridge University Press, 1992.

[9] Meyer, C. (1989)
Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems.
SIAM Review, **31** (1989), 240–272.

[10] Krishnamurthy, V. (1994)
Adaptive estimation of hidden nearly completely decomposable Markov chains with applications in blind equalization.
International Journal of Adaptive Control and Signal Processing, **8** (1994), 237–260.

[11] Dey, S. (2000)
Reduced-complexity filtering for partially observed nearly completely decomposable Markov chains.
IEEE Transactions on Signal Processing, **48** (Dec. 2000), 3334–3344.

[12] Yang, C., Bar-Shalom, Y., and Lin, C-F. (1992)
Discrete-time point process filter for mode estimation.
IEEE Transactions on Automatic Control, **37** (Nov. 1992), 1812–1816.

[13] Evans, J., and Evans, R. (1999)
Image-enhanced multiple model tracking.
Automatica, **35**, 11 (1999), 1769–1786.

[14] Bremaud, P. (1981)
Point Processes and Queues.
New York: Springer-Verlag, 1981.

[15] Wong, E., and Hajek, B. (1985)
Stochastic Processes in Engineering Systems (2nd ed.).
Berlin: Springer-Verlag, 1985.

[16] James, M., Krishnamurthy, V., and LeGland, F. (1996)
Time discretization of continuous-time filters and smoothers for HMM parameter estimation.
IEEE Transactions on Information Theory, **42** (Mar. 1996), 593–605.

[17] Krishnamurthy, V., and Elliott, R. (2002)
Robust continuous-time smoothers—without two-sided stochastic integrals.
IEEE Transactions on Automatic Control, **47** (Nov. 2002), 1824–1841.

[18] Clark, J. (1978)
The design of robust approximations to the stochastic differential equations of nonlinear filtering.
In J. Skwirzynski (Ed.), *Communication Systems and Random Processes Theory, Darlington 1977*, Alphen aan den Riju: Sijthoff and Noordhoff, 1978.

[19] Aldhaheri, R., and Khalil, H. (1991)
Aggregation of the policy iteration method for nearly completely decomposable Markov chains.
IEEE Transactions on Automatic Control, **36**, 2 (1991), 178–187.

[20] Dey, S., and Mareels, I.
Reduced-complexity estimation for large scale hidden Markov models.
IEEE Transactions on Signal Processing, to be published.

[21] Fischer, W., and Meier-Hellstern, K. (1993)
The Markov-modulated Poisson process (MMPP) cookbook.
Performance Evaluation, **18**, 2 (1993), 149–171.

[22] Ross, S. (1997)
Simulation.
New York: Academic Press, 1997.

Vikram Krishnamurthy was born in 1966. He received his bachelor's degree in electrical engineering from the University of Auckland, New Zealand, in 1988, and doctoral degree from the Australian National University, Canberra, in 1992.

Since July 2002 he has been a professor and Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, Canada. Prior to this he was a professor at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia for several years. His research interests span several areas including stochastic scheduling and wireless network optimization, time-series analysis, Bayesian estimation and statistical signal processing.

Dr. Krishnamurthy is currently an associate editor for *IEEE Transactions on Signal Processing* and *Systems and Control Letters*. He has served on the technical program committee of several international conferences.

Subhrakanti Dey was born in Calcutta, India in 1968. He obtained his Bachelor of Technology and Master of Technology degree from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively. He obtained his Doctor of Philosophy degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering in the Australian national University, Canberra, in 1996.

He is currently appointed as a senior lecturer with the Department of Electrical and Electronic Engineering at the University of Melbourne where he has been since February 2000. During September 1995–September 1997 and September 1998–February 2000, he was appointed as a postdoctoral research fellow in the Department of Systems Engineering, Australian National University. During September 1997–September 1998, he was a postdoctoral research associate at the Institute for Systems Research, University of Maryland, College Park.

His current research interests include signal processing for telecommunications, wireless communications and networks, performance analysis of communication networks, stochastic and adaptive estimation and control and statistical and adaptive signal processing.