12

Probabilistic Multi-Hypothesis Tracking

Roy L. Streit
Combat Systems Department

Tod E. Luginbuhl
Submarine Sonar Department

DTIC
SELECTED
SEP 0 8 1995
G
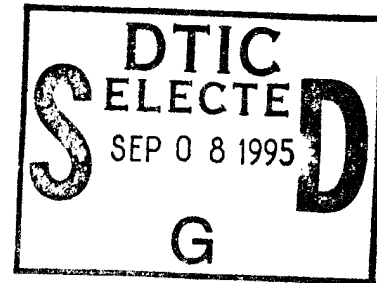
19950906 034

Naval Undersea Warfare Center Division
Newport, Rhode Island

DTIC QUALITY INSPECTED 5

PREFACE

The research described in this report was performed under Task 3 of the Platform Acoustic Warfare Data Fusion Project as part of the Submarine/Surface Ship USW Surveillance Technology Program sponsored by the Technology Directorate of the Office of Naval Research, Program Element 0602314N, ONR Technology Program UN3B, Project Number RJ14Q63, NUWC Job Order No. E67801, NUWC principal investigator M.E. Simard (Code 33A), program director G. C. Connolly (Code 2192). The ONR program manager is D. H. Johnson (ONR 321).

The technical reviewer for this report was M. L. Graham (Code 2214).

Reviewed and Approved: 15 February 1995

J. R. Short
Director for Submarine Combat Systems

Accesion For		
NTIS CRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

REPORT DOCUMENTATION PAGE

Public reporting for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 15 February 1995	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE

Probabilistic Multi-Hypothesis Tracking

5. FUNDING NUMBERS

6. AUTHOR(S)

Roy L. Streit
Tod E. Luginbuhl

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Undersea Warfare Center Division
1176 Howell Street
Newport, Rhode Island 02841-1708

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 10,428

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 No. Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

In a multitarget, multimeasurement environment, knowledge of the measurement-to-track assignments is typically unavailable to the tracking algorithm. This study is a probabilistic approach to the measurement-to-track assignment problem. Measurements are not assigned to tracks as in traditional multi-hypothesis tracking (MHT) algorithms; instead, the probability that each measurement belongs to each track is estimated using a maximum *a posteriori* (MAP) method. These measurement-to-track probability estimates are intrinsic to the multitarget tracker called the probabilistic multi-hypothesis tracking (PMHT) algorithm. The PMHT algorithm is computationally practical because it requires neither enumeration of measurement-to-track assignments nor pruning. The PMHT algorithm is an optimal MAP multitarget tracking algorithm.

14. SUBJECT TERMS		15. NUMBER OF PAGES 52
Targets	Tracking	
Bayesian Networks	Probabilistic Data	**16. PRICE CODE**

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT SAR

NSN 7540-01-280-5500

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS AND SYMBOLS

BIN	Bayesian inference network
CRLB	Cramer-Rao lower bound
EM	Expectation-maximization
FIM	Fisher information matrix
F_{tm}	Transition matrix for target m at time t for the linear-Gaussian case
H_{tm}	Measurement matrix for target m at time t for the linear-Gaussian case
i	Index of PMHT algorithm
J	FIM for X and Π, derived from the marginal PDF $P(Z,X|\Pi)$
JPDA	Joint probabilistic data association
$K \equiv K_T = (K_0, K_1, ..., K_T)$	Discrete component of the batch observer state
$K_t = \left(k_{t1}, ..., k_{t,n_t}\right)$	Vector of measurement-to-track assignments for Z_t, $t \geq 1$
k_{tr}	Target assigned to measurement z_{tr}, where $1 \leq k_{tr} \leq M$
$\kappa(K|Z,X)$	Conditional PDF of the discrete component used by the EM method
$\kappa(K|Z,X;\Pi)$	Same as $\kappa(K|Z,X)$ with parametric dependence on Π explicitly indicated
M	Number of targets (target motion models)

MAP	Maximum *a posteriori*
MHT	Multi-hypothesis tracking
ML	Maximum likelihood
$N = n_1 + \cdots n_T$	Total number of measurements in the batch
n_t	Number of measurements in the scan at time t, $n_t \geq 1$ for $t \geq 1$
N_{Xm}	State dimension of target m
N_z	Dimension of all measurements z_{tr}
$O \equiv O_T = (O_1,...,O_T)$ $\equiv (X_T, K_T)$	State of the batch observer
$O_t = (X_t, K_t)$	State of the scan observer at time t, $t \geq 1$
PDF	Probability density function
PMHT	Probabilistic multi-hypothesis tracking
$P(Z,O) \equiv P(Z,X,K)$	Joint PDF of the batch observer and batch measurement
$P(Z_t\|O_t)$	Scan measurement PDF, conditioned on scan observer state O_t, $t \geq 1$
$p(z_{tr}\|O_t)$	Single measurement PDF, conditioned on scan observer state O_t, $t \geq 1$
$P(Z,X,K;\Pi)$	Same as $P(Z,X,K)$ with parametric dependence on Π explicitly indicated
Q_{tm}	Process covariance matrix for target m at time t for the linear-Gaussian case
$Q(\cdot \| \cdot)$	Cross-entropy, or auxiliary, function defined by the EM method
r	Index used for measurements within a scan
R_{tm}	Noise covariance matrix for target m at time t for the linear-Gaussian case
s, m, v	Indices used for target states
T	Batch length (number of scans processed together)
t	Index used for discrete time
$X \equiv X_T = (X_0, X_1,..., X_T)$	Continuous component of the batch observer state
$X^m = (x_{0m}, x_{1m},..., x_{Tm})$	Vector of all target state random variables for target m
$X_t = (x_{t1},..., x_{tM})$	Vector of all target state random variables at time t
x_{tm}	State of target m at time t, where $0 \leq t \leq T$ and $1 \leq m \leq M$

$\left(\bar{x}_{0m}, \bar{\Sigma}_{0m}\right)$	Parameters of the prior PDF of target m for the linear-Gaussian case
$\mathbf{Z} \equiv \mathbf{Z}_T = (Z_1, ..., Z_T)$	Aggregate batch measurement
$Z_t = \left(z_{t1}, ..., z_{t,n_t}\right)$	Scan measurement vector at time t, $Z_0 \equiv \varnothing$, $Z_t \neq \varnothing$ for $t \geq 1$
z_{tr}	Measurement r in scan t for $t \geq 1$
$\zeta_m(z_{tr}|x_{tm})$	Single measurement PDF, conditioned on target m at time t, $t \geq 1$
$\Pi \equiv \Pi_T = \left(\pi_1, ..., \pi_T\right)$	Batch target measurement probabilities
$\pi^m = \left(\pi_{1m}, ..., \pi_{Tm}\right)$	Measurement probability vector for target m, $1 \leq m \leq M$
$\pi_t = \left(\pi_{t1}, ..., \pi_{tM}\right)$	Target measurement probability vector for scan t, $t \geq 1$
π_{tm}	Probability of measurements in scan t originating from target m, $t \geq 1$
$\varphi_m(x_{t,m}|x_{t-1,m})$	Process model PDF of target m at time t, $t \geq 1$
$\Psi = (\Psi_X, \Psi_K)$	Full process model of the scan observer
Ψ_K	Discrete PDF process model of the scan observer's discrete component
Ψ_X	PDF of the process model of the scan observer's continuous component
$(\cdot)'$	Matrix/vector transpose operator

PROBABILISTIC MULTI-HYPOTHESIS TRACKING

1. INTRODUCTION

In a multitarget, multimeasurement environment, knowledge of the measurement-to-track assignments is typically unavailable to the tracking algorithm. This report is a probabilistic approach to the measurement-to-track assignment problem; that is, measurement assignments are modeled as discrete random variables. Measurements are not assigned to specific tracks as in traditional multi-hypothesis tracking (MHT) algorithms; instead, the probability that each measurement belongs to each track is estimated using an empirical Bayesian algorithm. The probabilistic multi-hypothesis tracking (PMHT) approach proposed in this report treats target states and measurement assignments as continuous and discrete random variables, respectively, and defines an appropriate joint density on these variables. The PMHT estimation algorithm is in the class of so-called empirical Bayesian methods (reference 1); that is, it is a hybrid maximum *a posteriori* (MAP) and maximum likelihood (ML) algorithm. The PMHT estimates are joint estimates of target states and measurement-to-track assignment probabilities. This report expands earlier work. (See references 2 and 3.)

The PMHT algorithm requires neither enumeration of measurement-to-track assignments nor pruning because all measurements are assigned to all tracks. A measurement is weighted with the estimated measurement-to-track assignment probability before it is assigned to a target track. The same measurement receives different weights for different tracks. Because pruning is not required, the PMHT algorithm is an optimal empirical Bayesian multitarget tracking algorithm under idealized assumptions.

The computational complexity of the PMHT algorithm is much less than the exponential complexity of MHT algorithms, which use exhaustive enumeration. Current MHT algorithms combine iterative methods for enumeration with careful scoring and pruning (e.g., measurement gating and branch elimination) to reduce computational complexity and computer storage requirements to manageable levels. The PMHT algorithm presented in this report is potentially more advantageous than the MHT algorithms because the PMHT algorithm is amenable to parallel computation on high-performance computer architectures. Specifically, PMHT measurement-to-track assignment probabilities greatly reduce the need for branching (i.e., algorithm flow is smooth and relatively uninterrupted).

One of the key ideas behind the PMHT algorithm is the avoidance of "hard" measurement-to-track assignment decisions. It may be argued that hard assignments should be avoided because they are equivalent to statistical decisions, that all statistical decisions are opportunities for error, and that erroneous decisions, even if rarely committed, necessarily increase estimation error. However, as will be demonstrated, the formulation of the PMHT probability structure does not support this argument because the optimal joint MAP estimate of target states and measurement assignments comprise state estimates and hard assignments. One of the primary technical contributions of the PMHT approach is that it avoids the computational complexity of MAP estimation (i.e., the theoretical basis of MHT algorithms) and, yet, is optimal in the empirical

1

Bayesian sense. The change in the definition of the likelihood function that results from treating assignments as discrete random variables thus has a dramatic effect on the resulting estimation algorithm and its computational complexity.

This report consists of ten numbered sections, the contents of which are indicated below:

- Section 1 introduces the PMHT algorithm.

- Section 2 introduces the PMHT observer and presents a theoretical overview of the PMHT estimation method.

- Section 3 states the likelihood structure of the PMHT observer. Bayesian inference networks (BINs) are introduced as a road map for representing the conditional independence assumptions that are central to the methods of this study. The validity of the PMHT approach for nonlinear non-Gaussian systems is attributed to the fact that these conditional independence assumptions, displayed graphically by the BIN, are nonparametric assumptions and, therefore, are not limited to linear-Gaussian parameterizations.

- Section 4 gives a derivation of the PMHT algorithm using the expectation-maximization (EM) method. The derivation assumes familiarity with the EM method.

- Section 5 states the PMHT algorithm in recursive form. The linear-Gaussian case was first presented (without proof) in reference 2.

- Section 6 discusses Fisher information matrices (FIMs) and multitarget observability. FIMs relate to PMHT estimation error, and the multitarget observability relates to PMHT algorithm convergence.

- Section 7 discusses several variations and extensions to the PMHT approach, including adaptive covariance estimation.

- Section 8 compares the PMHT system with other multitarget tracking algorithms.

- Section 9 presents a PMHT example.

- Section 10 provides concluding remarks and recommendations for further study.

2. THEORETICAL OVERVIEW

Section 2 uses an abbreviated form of the notation used in the rest of this report. A complete description of the issues mentioned in this section is presented in sections 3 and 4.

Measurements are outputs from a sensor signal processor, and the task of the postprocessor is to compute target tracks from the available measurements. PMHT is an algorithm for the postprocessor. Adopting the observer concept is useful when studying the multitarget tracking problem. The observer is not particularly interesting for single-target tracking, but its value for multitarget tracking becomes clear with use.

The PMHT algorithm assumes independent targets. In the special case of linear-Gaussian statistics, each target has a state motion model of the form

$$x_{t+1} = F_t x_t + G_t w_t,$$

where w_t is white Gaussian process noise with known covariance matrix Q_t. Measurements z_t within each scan Z_t are assumed conditionally independent, when conditioned on the collection of target states. Measurements in different scans are also assumed conditionally independent. Different scans may have different numbers of measurements.

The PMHT observer is defined by the state of the postprocessor. Since measurements are not presented to the postprocessor with labels identifying the correct measurement-to-track assignments, the appropriate observer state, denoted O_t, for measurement scan Z_t comprises the collection of target states together with the collection of measurement-to-track assignments. Therefore, the observer state has both continuous and discrete components that must be estimated. See the last paragraph in this section for a brief description of the PMHT estimation methodology (EM).

Suppose, momentarily, that the observer state is defined so that it does not comprise measurement assignments. Then, even if the target states are known exactly, this modified observer is unable to determine the correct measurement-to-track assignments. Hence, this modified observer cannot be an observer in the strict sense (see reference 4, section 9.2). Consequently, the measurement-to-track assignments must be part of the observer state definition. Incorporating assignments explicitly into the observer state appears to be novel to this study.

The probability density function (PDF) of the measurement scan Z_t is conditioned solely on the scan observer state O_t. The PDF of a measurement z_t in Z_t, conditioned on O_t, is assumed known. In the linear-Gaussian case, the measurement PDF has the form

$$p(z_t | O_t) = \mathcal{N}\left(z_t | H_{k_t} x_{k_t}, R_{k_t}\right),$$

where k_t is the discrete component of O_t corresponding to z_t; i.e., k_t denotes the target of origin for z_t. This is equivalent to writing

$$z_t = H_{k_t} x_{k_t} + v_{k_t},$$

where v_{k_t} is white Gaussian measurement noise with known covariance matrix R_{k_t}. Because the scan measurements are assumed to be independent, conditioned on the observer state, the scan likelihood function, denoted by $P(Z_t|O_t)$, is the product over all measurements z_t in Z_t of the individual measurement likelihood functions $p(z_t|O_t)$. It is important to emphasize that the discrete components of the observer O_t are essential to the definition of the scan-likelihood function.

PMHT is a batch algorithm that uses a finite number of successive measurement scans to estimate the batch observer state O. The batch observer is the collection of scan observers in the batch, and the batch observer state comprises the collection of the scan observer states, so that $O = \{O_t\}$. A batch PDF is formulated for the batch observer using the target motion models and the scan likelihood functions. The batch PDF is a joint function of the measurements, the continuous component X, and the discrete component K of the batch observer state O.

BINs are used in this report to display the conditional independence assumptions between measurements and the batch observer states in a graphical manner. These conditional independence assumptions are fundamental to the batch observer joint PDF and are easily understood if presented in a BIN graphical form. The BIN for PMHT is shown in figure 1. The nodes and edges of the BIN graph are interpreted in such a way that the graph is equivalent to a factorization of the batch observer PDF. (See the end of section 3 for a discussion of BINs.)
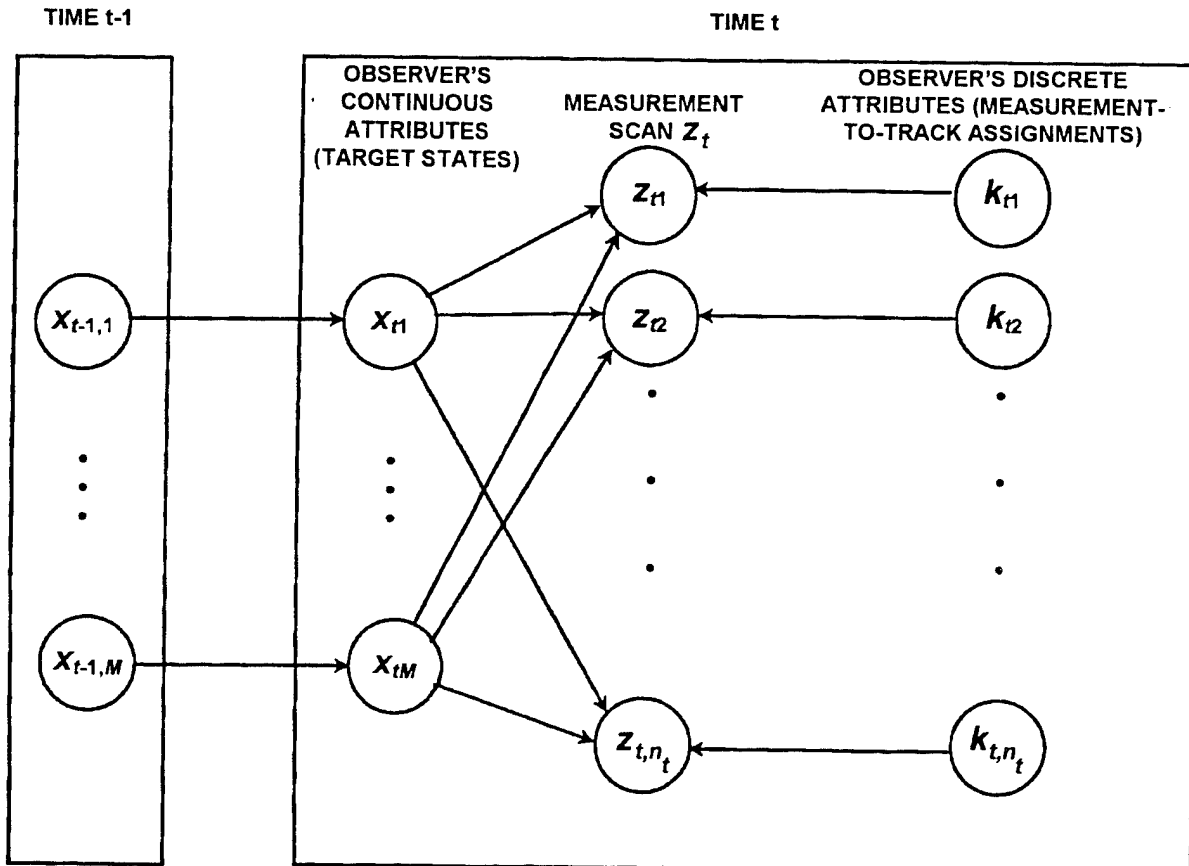

Figure 1. BIN Fragment for PMHT at Scan t

The MAP estimate of O is the maximum (over the observer's continuous and discrete components) of the PDF of O conditioned on the available batch measurements. However, in the absence of computationally efficient search techniques (e.g., dynamic programming), the MAP estimate is combinatorially hard to compute because it involves enumeration of all possible hard measurement-to-track assignments. Computation of such assignments is not required by PMHT.

The PMHT algorithm differs from the MAP point estimate just described. The PMHT estimates of the batch observer state are computed in three separate, but interrelated, steps. In the first step, the batch joint PDF is marginalized (summed) over the observer discrete component K, i.e., over all measurement-to-track assignments. The marginal density is the joint PDF of the batch measurement Z and the continuous component X of the batch observer. The second step estimates the target states x from the marginal PDF using an algorithm derived by the EM method. This computation yields the PMHT state estimate for each target at each scan in the batch. In the third step, the conditional density for the observer discrete component K is computed from Bayes theorem by conditioning on the measurements Z and on the PMHT target state estimates computed in the second step. This conditional PDF is the probability of measurement-to-track assignment for every measurement and track pair.

The coupling between the three PMHT steps is of fundamental importance. The coupling is caused by assignment interference because it must estimate the parameters defining the discrete conditional distribution of the observer discrete component K. These parameters are merely the measurement-to-track assignment probabilities. The PMHT algorithm is essentially a recursive algorithm for estimating assignment interference.

The second PMHT step yields, upon convergence, a conditional error covariance matrix for each target. It will be shown that the inverse covariance matrix for target s, computed in the second step, is the expected FIM, where the expectation is, with respect to the conditional density, on the observer discrete component K.

PMHT reliance on marginalization is unusual from the perspective of multitarget tracking, but it is natural from a probabilistic perspective. Marginalization is also a traditional method of treating so-called nuisance variables in statistical problems. Marginalization removes dependence on knowledge of particular outcomes of the variables marginalized, but the dependence on the distributional parameters of these variables remains. In the multitarget tracking problem, the measurement-to-target assignments are the nuisance variables, and they are marginalized out. The parameters of the nuisance variable distributions are estimated from the marginal distribution using the EM method.

3. PMHT OBSERVER LIKELIHOOD STRUCTURE

Let $M \geq 1$ denote the assumed number of independent target motion models. The integer M should be at least as large as the number of targets present in the measurement sequence. Choosing M greater than the number of targets may have practical utility for background noise or clutter normalization purposes.

The PMHT algorithm is a batch algorithm. Let $T \geq 1$ denote the length of the current batch. The measurement scans are numbered so that the batch comprises scans from time $t = 1$ to time $t = T$. The PMHT observer is a batch observer, so it is convenient to define the batch observer as a collection of scan observers. For $t \geq 1$, the PMHT scan observer state is denoted by $O_t \equiv \left(X_t, K_t \right)$, and it comprises the continuous component vector X_t and the discrete component vector K_t. The continuous component X_t comprises the M target states. Explicitly, X_t is given by

$$X_t = \left(x_{t1}, \ldots, x_{tM} \right), \qquad t = 1, \ldots, T,$$

where x_{ts} denotes the state of target s at time t. Component states x_{ts} and x_{tm} of X_t are independent, for $s \neq m$ because the targets are assumed to be independent. The *a priori* PMHT scan observer state is denoted by $O_0 \equiv \left(X_0, K_0 \right)$, where the discrete component $K_0 = \varnothing$ (the empty set) because, by definition, measurements are unavailable at $t = 0$. The continuous component $X_0 = \left(x_{01}, \ldots, x_{0M} \right)$ is discussed below. The target state vector X^m comprises the state of target m at each time in the batch, that is,

$$X^m = \left(x_{0m}, x_{1m}, \cdots, x_{Tm} \right), \qquad m = 1, \ldots, M.$$

Let N_{Xs} denote the dimension of the target state for model s. The state dimension N_{Xs} may vary from target model to target model, but it is assumed constant from scan to scan.

Let $\varphi_s \left(x_{ts} \middle| x_{t-1,s} \right)$ denote the process model for target s, $1 \leq s \leq M$. Target model s assumes that x_{0s} is a realization of the *a priori* distribution denoted by $\varphi_s \left(x_{0s} \right)$. The time dependence of the functions $\left\{ \varphi_s \right\}_{s=1}^{M}$ is indicated implicitly by the function argument. In the special case of linear-Gaussian target process models,

$$\varphi_s \left(x_{ts} \middle| x_{t-1,s} \right) = \mathcal{N} \left(x_{ts} \middle| F_{t-1,s} x_{t-1,s}, \, G_{t-1,s} Q_{t-1,s} G'_{t-1,s} \right), \qquad t = 1, \ldots, T, \tag{1}$$

where $\mathcal{N} \left(\cdot \middle| \mu, \Sigma \right)$ denotes a multivariate Gaussian PDF with mean vector μ and covariance matrix Σ. Equation (1) is equivalent to the conventional form

$$x_{ts} = F_{t-1,s} x_{t-1,s} + G_{t-1,s} w_{t-1,s}, \qquad t = 1, 2, \ldots, T, \tag{2}$$

where $w_{t-1,s}$ is a white Gaussian process with (positive definite) covariance matrix $Q_{t-1,s}$. Noise processes $w_{t-1,s}$ and $w_{t-1,m}$ are assumed to be independent because $s \neq m$. The matrices $F_{t-1,s}$, $G_{t-1,s}$, and $Q_{t-1,s}$, are assumed known for all $t > 0$ and s.

The target process models hold for $t = 0$, so state x_{1s} is a realization of the distribution

$$\varphi_s(x_{1s}) = \int_{-\infty}^{+\infty} \varphi_s(x_{1s}|x_{0s}) \, \varphi_s(x_{0s}) dx_{0s}. \tag{3}$$

The initial target states $\{x_{0s}\}$ can therefore be eliminated without altering the batch likelihood structure. In this report, however, target states are initialized at time $t = 0$ for notational convenience. For the linear-Gaussian case, x_{0s} is a realization of the *a priori* Gaussian PDF

$$\varphi_s(x_{0s}) = \mathcal{N}\left(x_{0s}\middle|\overline{x}_{0s}, \overline{\Sigma}_{0s}\right), \tag{4}$$

where the mean $\overline{x}_{0s}$ and covariance matrix $\overline{\Sigma}_{0s}$ are given, and the distribution (3) is given by

$$\varphi_s(x_{1s}) = N\left(x_{1s}\middle| F_{0s}\overline{x}_{0s}, \; F_{0s}\overline{\Sigma}_{0s}F_{0s}' + G_{0s}Q_{0s}G_{0s}'\right),$$

where the mean and covariance of $\varphi_s(x_{1s})$ are the predictive statistics.

Let Ψ_x denote the PDF of the scan observer continuous component X_t. The time dependence of Ψ_x will be indicated implicitly by its argument. For $t \geq 1$, the PDF of X_t is conditioned on X_{t-1}. From the assumption of independent targets, it follows that

$$\Psi_X(X_t|X_{t-1}) = \prod_{s=1}^{M} \varphi_s(x_{ts}|x_{t-1,s}). \tag{5}$$

The *a priori* PDF of the continuous component X_0 defining the initial observer is given by

$$\Psi_X(X_0) = \prod_{v=1}^{M} \varphi_v(x_{0v}), \tag{6}$$

where independent target initialization has been assumed.

8

For $t \geq 1$, the discrete component K_t of the scan observer state O_t comprises a unique measurement-to-track assignment for every measurement in scan Z_t. The discrete component K_t and the continuous component X_t are assumed to be statistically independent. Let n_t denote the number of measurements in scan Z_t. For notational simplicity it is assumed that $n_t \geq 1$; however, no theoretical difficulty arises if $n_t = 0$. Recall that $K_0 = \varnothing$. Let z_{tr} denote measurement r in scan $Z_t \equiv \left(z_{t1}, z_{t2}, \ldots, z_{t,n_t} \right)$. The discrete component K_t is defined by

$$K_t = \left(k_{t1}, k_{t2}, \ldots, k_{t,n_t} \right),$$

where $1 \leq k_{tr} \leq M$ for all $r = 1, 2, \ldots, n_t$. Thus, measurement z_{tr} is assigned by K_t to the track motion model with index $k_{tr} \in K_t$. It is assumed that the discrete components $\{ k_{tr} \}$ are independent, that is, k_{tr} and $k_{tr'}$ are independent if $r \neq r'$. Finally, the discrete component vectors $\{ k_t \}$ are assumed statistically independent from scan to scan, that is, K_t is independent of $K_{t'}$ because $t \neq t'$.

The PMHT measurement PDF is defined for measurement z_{tr} by

$$p\left(z_{tr} | O_t \right) \equiv p\left(z_{tr} | X_t, K_t \right) \equiv \left. \zeta_m \left(z_{tr} | x_{tm} \right) \right|_{m = k_{tr}}, \qquad t \geq 1.$$

All measurements in all scans are assumed (for simplicity) to have the same dimension, denoted by N_z. Under linear-Gaussian assumptions,

$$\zeta_m \left(z_{tr} | x_{tm} \right) = \mathcal{N} \left(z_{tr} | H_{tm} x_{tm}, R_{tm} \right). \tag{7}$$

The linear-Gaussian case can also be written in the equivalent form

$$z_{tr} = H_{tm} x_{tm} + v_{tm},$$

where v_{tm} is an additive white Gaussian noise process with covariance R_{tm}. The PMHT scan likelihood function is a PDF defined over the measurements in scan Z_t, and it is conditioned on the observer state O_t. Explicitly, because the measurements are conditionally independent,

$$P\left(Z_t | O_t \right) \equiv P\left(Z_t | X_t, K_t \right) \equiv \prod_{r=1}^{n_t} \left. \zeta_m \left(z_{tr} | H_{tm} x_{tm}, R_{tm} \right) \right|_{m = k_{tr}}. \tag{8}$$

The discrete component k_{tr} of the vector K_t serves as a pointer to the appropriate target model and is thus an integral component of measurement conditioning. The parameters defining probability mass function of the discrete components of the measurements in scan K_t are not part of the observer state. (Reference to the BIN of figure 1 will clarify this important distinction.)

Let π_{tm} represent the probability that a measurement in scan Z_t is assigned to target motion model m, where $t \geq 1$. The probability π_{tm} reflects the fraction of scan measurements assigned to

target m at time t. These target measurement probabilities are needed because some targets may produce more measurements per scan than others because of individual target characteristics (e.g., signal-to-noise ratio), environmental effects, sensor properties, and other application considerations. Denote the "within-scan" measurement probability vector by

$$\pi_t \equiv \left(\pi_{t1}, \pi_{t2}, \ldots, \pi_{tM} \right). \tag{9}$$

The vector π_t parameterizes the distribution of the discrete component k_{tr} for the measurement z_{tr} in scan Z_t. Because measurements within a scan are assumed conditionally independent and identically distributed, the distribution π_t is assumed to be the same for all measurements made at time t. Thus, $\pi_{tm} = \mathrm{Prob}\left[k_{tr} = m \right]$ for all measurement assignments k_{tr}. The batch target measurement probabilities $\Pi \equiv \Pi_T \equiv \left(\pi_1, \ldots, \pi_T \right)$ are estimated by the PMHT algorithm from the batch measurement data.. Alternative assumptions concerning Π are possible. For instance, a Bayesian *a priori* density can be assumed for π_t if desired. A Bayesian *a priori* density for K_t is analogous to the initialization (3) for X_t. This and other alternatives are described in section 7.1.

Let Ψ_K denote the discrete PDF (or, probability mass function) of the discrete component K_t of the scan observer O_t. The time dependence of Ψ_K is indicated implicitly by its argument. The PDF of K_t takes the form

$$\Psi_K \left(K_t \right) \equiv \mathrm{Prob}\left[k_{t1}, \ldots, k_{t,n_t} \right] \equiv \prod_{r=1}^{n_t} \pi_{tm} \Big|_{m=k_{tr}}. \tag{10}$$

The product in (10) follows from the assumption that different measurements within a scan have statistically independent and identically distributed discrete components k_{tr}. It turns out (see equation (22) and the ensuing discussion) that the measurement independence assumptions imply that π_{tm} are mixing proportions of a mixture density modeling the scan measurements. For the linear-Gaussian case, the means of the Gaussian components of the mixture are the target state estimates for the scan.

The batch observer state, denoted $O \equiv O_T \equiv \left(X_T, K_T \right)$, comprises the continuous and discrete components of the T-scan observers in the current batch. The dependence of the observer and its components on batch length T is suppressed throughout the remainder of this report. Thus, the continuous and discrete components of the batch observer O are

$$X \equiv X_T \equiv \left(X_0, X_1, \ldots, X_T \right) \tag{11}$$

and

$$K \equiv K_T \equiv \left(K_0, K_1, \ldots, K_T \right), \tag{12}$$

10

respectively. K_0 is a place holder in (12) because $K_0 = \varnothing$. The batch measurement is denoted

$$Z \equiv Z_T \equiv \left(Z_1, \ldots, Z_T\right). \tag{13}$$

It is assumed that measurement scans and measurements within scans are statistically independent, conditioned on the batch observer state O.

The PMHT batch observer PDF is a joint function of the measurements Z and the batch observer state O. Because no measurements are made at time $t = 0$, the *a priori* PDF of the scan observer at time $t = 0$ is defined to be

$$\Psi\left(O_0\right) = \prod_{v=1}^{M} \varphi_v\left(x_{0v}\right). \tag{14}$$

The scan observer at time $t \geq 1$ is conditionally dependent on the scan observer at time $t - 1$. The scan observer continuous and discrete components are independent, so the observer conditional PDF is given by

$$\Psi\left(O_t\middle|O_{t-1}\right) = \Psi\left(X_t, K_t\middle|X_{t-1}, K_{t-1}\right),$$

$$= \Psi_x\left(X_t\middle|X_{t-1}\right)\Psi_K\left(K_t\middle|K_{t-1}\right),$$

$$= \Psi_x\left(X_t\middle|X_{t-1}\right)\Psi_K\left(K_t\right).$$

Substituting the expressions (5) and (10) gives for $t \geq 1$

$$\Psi\left(O_t\middle|O_{t-1}\right) \equiv \left\{\prod_{s=1}^{M} \varphi_s\left(x_{ts}\middle|x_{t-1,s}\right)\right\}\left\{\prod_{r=1}^{n_t} \pi_{tm}\bigg|_{m=k_{tr}}\right\}. \tag{15}$$

The batch observer PDF is, from the conditional independence assumptions,

$$P(Z, O) \equiv P(Z, X, K) = \Psi\left(O_0\right)\prod_{t=1}^{T} \Psi\left(O_t\middle|O_{t-1}\right)P\left(Z_t\middle|O_t\right). \tag{16}$$

Substituting (8) and (15) into this expression gives the batch observer joint PDF in the form

$$P(Z, X, K) = \left\{\prod_{v=t}^{M} \varphi_v\left(x_{0v}\right)\right\}\prod_{t=1}^{T}\left\{\left[\prod_{s=1}^{M} \varphi_s\left(x_{ts}\middle|x_{t-1,s}\right)\right]\prod_{r=1}^{n_t}\left[\pi_{tm}\zeta_m\left(z_{tr}\middle|x_{tm}\right)\bigg|_{m=k_{tr}}\right]\right\}. \tag{17}$$

When the parametric dependence on the probabilities Π is made explicit, the joint PDF is written $P(Z,X,K : \Pi)$. The batch PDF (17) is the fundamental probabilistic structure of the PMHT algorithm. The derivation of the PMHT algorithm is given in section 4.

Insight and understanding of the structure of $P(Z,X,K)$ is enhanced by displaying its conditional independence assumptions in a mathematically equivalent graphical format. A BIN is a very effective representational technique designed for just such a purpose. A fragment of the fundamental graphical structure of the PMHT batch PDF for measurements at scan t is portrayed as a BIN in figure 1. The BIN of figure 1 is a directed graph, and its nodes represent random variables in the batch observer PDF $P(Z,X,K)$. The directed edges depict the conditioning of these random variables in the following way: Each node is conditioned jointly on all of its "parents." The overall likelihood structure of the BIN is defined to be the product of all the conditional PDFs. There are precisely as many conditional PDF factors as there are nodes in the graph. Following this procedure for the BIN of figure 1 shows that its joint likelihood function is identical, factor by factor, to the factorization of $P(Z,X,K)$ given by equation (17). The precise mathematical forms of the individual conditional PDF factors are not defined by the BIN, but they are stipulated by the requirements of the specific application. The Gauss-Markov equation (2), the Gaussian initialization (3), and the measurement conditioning equation (7) are specific to PMHT. The directed graph provides global likelihood structure, while the specific conditional PDFs provide local structure. The graph may not have "directed cycles" because they result in improper factorizations of the global joint likelihood function.

The special case $M = 1$ of the PMHT batch PDF (17) is identical to the PDF of a fixed-interval Kalman filter formulated for multiple measurements. The reason for these identical PDFs is that, in this case, the batch discrete component K and the within-scan target measurement probabilities Π are trivial because there is only one target and all batch measurements are necessarily assigned to it. In addition, if each scan has precisely one measurement, the PMHT batch likelihood function is identical to the likelihood function of the usual textbook Kalman filter (see reference 5).

4. DERIVATION OF THE PMHT ALGORITHM

In this discussion, the authors assume the reader is familiar with the EM method. For background material on this very general technique, see reference 6 and the references mentioned therein. For PMHT, the "missing data" in the sense of EM is the discrete component K of the batch observer, and the "complete data" comprises Z, X, and K. This confusing nomenclature is not used in this report. In the special case of linear-Gaussian statistics with known covariances, the target measurement probabilities Π and the states X of the continuous component are batch parameters to be estimated. This case is especially simple because the parameters defining the target process and measurement PDFs are linear functions of the state means, and the estimated means are the target state estimates. In general, however, it is necessary to distinguish between state parameter estimates and state estimates. For nonlinear, non-Gaussian processes, therefore, it is assumed that X is the parameter set defining the target process and measurement PDF. This slight abuse of notation should not cause confusion.

The discussion in this section derives the expectation step (E-step) of the EM method for the general case because the E-step is done as easily in general as it is for linear-Gaussian assumptions. As is seen below, the measurement probabilities Π and target parameters X are treated separately in the maximization step (M-step) even in the general case. The M-step treats the probabilities Π for the general case. The M-step for the parameter estimates X is significantly easier to understand for linear-Gaussian statistics, so this special case is treated separately first. Subsequently, the general case for X is treated.

The E-step begins by defining a PDF on the discrete component K. Let $\mathcal{K}$ denote this PDF. On convergence of the PMHT algorithm, $\mathcal{K}$ gives the measurement-to-track assignment probabilities for all possible measurement and target pairs. From Bayes theorem, the conditional PDF on K is defined by

$$\mathcal{K}(K|Z,X) \equiv \frac{P(Z, X, K : \Pi)}{P(Z, X : \Pi)},$$

(18)

where the denominator of (18) is the marginal distribution of the PMHT likelihood function over the discrete component K. When the parametric dependence on the probabilities Π is made explicit, the conditional PDF (18) is written $\mathcal{K}(K|Z,X:\Pi)$. The marginal distribution of $P(Z,X,K)$ over K is defined by

$$P(Z, X) \equiv P(Z, X : \Pi) \equiv \sum_K P(Z, X, K),$$

(19)

where the summation over the batch component K is defined by

$$\sum_K = \sum_{t=1}^{T} \sum_{K_t} \equiv \sum_{K_1} \cdots \sum_{K_T},$$

(20)

13

and where the sum over the scan component K_t is defined by

$$\sum_{K_t} = \sum_{r=1}^{n_t} \sum_{k_{tr}=1}^{M} \equiv \sum_{k_{t1}=1}^{M} \cdots \sum_{k_{t,n_t}=1}^{M}.$$

(21)

Using the most expanded of the summation forms (20) and (21), it is straightforward to verify the important algebraic identity

$$P(Z, X) = \left\{ \prod_{v=1}^{M} \varphi_v\left(x_{0v}\right) \right\} \prod_{t=1}^{T} \left\{ \left[\prod_{s=1}^{M} \varphi_s\left(x_{ts} \big| x_{t-1,s}\right) \right] \prod_{r=1}^{n_t} \left[\sum_{m=1}^{M} \pi_{tm} \, \zeta_m\left(z_{tr} \big| x_{tm}\right) \right] \right\}.$$

(22)

PMHT computes an ML estimate of the continuous component X and the target measurement probabilities Π from the marginal PDF (22). The distribution $P(Z,X)$ is, thus, the appropriate PDF for a batch observer whose state definition does not explicitly include discrete components. The marginal PDF is fundamental to multitarget observability questions, and observability is related to PMHT algorithm convergence. (See section 6 for additional information.)

The marginal PDF (22) exhibits clearly the underlying assumption that the measurements within scan Z_t are conditionally independent. It also provides the useful interpretation that the components of the batch target measurement probability vector π_t are the mixing proportions of the measurement PDF mixture

$$\sum_{m=1}^{M} \pi_{tm} \, \zeta_m\left(z_{tr} \big| x_{tm}\right),$$

from which the measurements in scan Z_t are drawn. Substituting (17) and (22) into definition (18) gives

$$\mathcal{K}(K|Z, X) = \prod_{t=1}^{T} \prod_{r=1}^{n_t} w_{str} \Big|_{s=k_{tr}},$$

(23)

where the weight $w_{str} > 0$ is a function of Z, X, and Π and is given by

$$w_{str} = \frac{\pi_{ts} \zeta_s\left(z_{tr} \big| x_{ts}\right)}{\sum_{m=1}^{M} \pi_{tm} \zeta_m\left(z_{tr} \big| x_{tm}\right)}.$$

(24)

14

The expression (23) does not include a product over the state models because terms involving $\{\varphi_s\}$ cancel out of the ratio (18). The weight w_{str} is interpreted, using Bayes theorem, as the conditional probability that measurement z_{tr} is assigned to the target model s, conditioned on the continuous component X and the measurements Z; that is,

$$w_{str} = \text{Prob}\left[k_{tr} = s \middle| Z, X\right].$$

From the definition (24), it is straight forward to verify the algebraic identities

$$\sum_K \mathcal{K}(K|Z,X) = 1,$$

(25)

and

$$\sum_{K \backslash k_{tr}} \mathcal{K}(K|Z,X) = w_{str}\Bigg|_{s=k_{tr}},$$

(26)

where, in (26), the sum over $K \backslash k_{tr}$ is the sum over all indices in K except k_{tr}.

The E-step is concluded by evaluating the expectation of the logarithm of the PMHT batch PDF (17), where the expectation is with respect to the conditional PDF (23). Let X' denote a given value for the continuous component of the batch observer, and let Π' denote a given value for the target measurement probabilities. The primes on these variables do not denote the matrix/vector transpose operator. The expectation required by the E-step is written explicitly as

$$Q \equiv Q(\Pi, X|\Pi', X') = \sum_K \left\{\log P(Z, X, K:\Pi)\right\} \mathcal{K}(K|Z, X':\Pi').$$

(27)

The function Q is called the auxiliary function in reference 6, and it is closely related to cross-entropy and the Kullback-Leibler distance. In this report function Q is called the cross-entropy function. Taking the logarithm of the PMHT batch PDF (17) gives

$$\log P(Z, X, K:\Pi) = \sum_{v=1}^{M} \log \varphi_v(x_{0v}) + \sum_{t=1}^{T}\sum_{s=1}^{M} \log \varphi_s(x_{ts}|x_{t-1,s})$$
$$+ \sum_{t=1}^{T}\sum_{r=1}^{n_t} \left[\log \pi_{tm} + \log \zeta_m(z_{tr}|x_{tm})\right]\Bigg|_{m=k_{tr}}.$$

Substituting this expression into the definition (27), interchanging the summation order so that, for example,

$$\sum_K \sum_t \sum_r = \sum_t \sum_r \sum_{k_{tr}} \sum_{K \backslash k_{tr}},$$

15

and then using the identities (25) and (26) gives the cross-entropy function in the form

$$Q = \sum_{t=1}^{T} Q_{t,\Pi} + \sum_{m=1}^{M} Q_{m,X},$$

(28)

where

$$Q_{t,\Pi} \equiv Q_{t,\Pi}(\pi_t | Z, X_t', \pi_t') = \sum_{r=1}^{n_t} \sum_{m=1}^{M} w_{mtr}' \log \pi_{tm},$$

(29)

$$Q_{m,X} \equiv Q_{m,X}\left[X^m | Z, (X^m)', (\pi^m)' \right]$$

$$= \log \varphi_m(x_{0m}) + \sum_{t=1}^{T} \left\{ \log \varphi_m(x_{tm} | x_{t-1,m}) + \sum_{r=1}^{n_t} w_{mtr}' \log \zeta_m(z_{tr} | x_{tm}) \right\},$$

(30)

and where the weights $w_{mtr}' \equiv w_{mtr}'(Z, X', \Pi')$ are defined as in (24). In equations (29) and (30), the notations

$$X' = (X_0', X_1', \dots, X_T') = [(X^1)', \dots, (X^M)']$$

and

$$\Pi' = (\pi_1', \dots, \pi_T') = [(\pi^1)', \dots, (\pi^M)'],$$

have been used. Again, the primes on these variables do not denote transpose.

The M-step is a maximization problem. Explicitly, given values Π' and X', the M-step requires computing values $\hat{\Pi}$ and $\hat{X}$ for which

$$Q(\hat{\Pi}, \hat{X} | \Pi', X') = \max_{\Pi, X} Q(\Pi, X | \Pi', X').$$

(31)

The given variables Π' and X' comprise the current values of the PMHT algorithm, and the variables $\hat{\Pi}$ and $\hat{X}$ comprise updated values of the algorithm. The PMHT recursion is stated explicitly in section 5. On convergence of the PMHT algorithm, $\hat{\Pi}$ and $\hat{X}$ are ML-parameter estimates. Equations (28), (29), and (30) demonstrate that the maximization problem (31) decouples into a maximization problem for each of the T-probability vectors π_t and a maximization problem for each of the M-target state sequences X^m. The weights $\{w_{mtr}'\}$ change from M-step to M-step, but within each M-step, the weights are fixed.

16

The maximization problem for π_t is constrained by the requirement

$$\sum_{m=1}^{M} \pi_{tm} = 1,$$

(32)

so the appropriate Lagrangian for this problem is

$$L_t \equiv Q_{t,\Pi} + \gamma_t \left(1 - \sum_{m=1}^{M} \pi_{tm} \right),$$

where γ_t is the Lagrange multiplier. Differentiating the Lagrangian with respect to π_{tm} and setting the result to zero gives the necessary condition

$$\hat{\pi}_{tm} = \frac{1}{\gamma_t} \sum_{r=1}^{n_t} w'_{mtr}.$$

Summing these equations from $m = 1$ to M and using the constraint (32) gives

$$\gamma_t = \sum_{r=1}^{n_t} \sum_{m=1}^{M} w'_{mtr}.$$

From the expression (24) for the weights, it follows easily that $\gamma_t = n_t$, so the unique stationary point of the Lagrangian L_t is

$$\hat{\pi}_{tm} = \frac{1}{n_t} \sum_{r=1}^{n_t} w'_{mtr}.$$

(33)

By lemma 2 of reference 7, it follows that $\hat{\pi} = \left(\hat{\pi}_{t1}, \hat{\pi}_{t2}, \ldots, \hat{\pi}_{tM} \right)$ is the unique global maximum of the cross-entropy function $Q_{t,\Pi}$. It follows from (33) and (24) that $\hat{\pi}_{tm} = 0$, if and only if $\hat{\pi}_{tm} = 0$. Because $Q_{t,\Pi} = -\infty$ and, hence, $Q = -\infty$ if $\hat{\pi}_{tm} = 0$, it is assumed without loss of generality that the initial probabilities Π' are chosen to be strictly positive.

The solution for the state sequence X^m for target m is derived first for the special case of linear-Gaussian statistics. The insight provided by this special case is helpful in understanding the general case, which is discussed beginning with equation (39). Recall the general gradient identity

$$\nabla_X (Fx - \mu)' \Sigma^{-1} (Fx - \mu) = 2 F' \Sigma^{-1} (Fx - \mu).$$

(34)

17

Substituting the linear-Gaussian models (1), (4), and (7) into the definition of the cross-entropy function $Q_{m,x}$ and taking the gradient of $Q_{m,x}$ with respect to the state vector x_{tm} for $t = 0,1,...,T$ gives a symmetric block tridiagonal system of linear equations for the state sequence for target model m. This system is written

$$\left(A_{0m} + D_{0m}\right)x_{0m} - B_{0m}x_{1m} = A_{0m}\bar{x}_{0m},$$

$$-B'_{tm}x_{t-1,m} + \left(A_{tm} + D_{tm}\right)x_{tm} - B_{tm}x_{t+1,m} = n_t\hat{\pi}_{tm}H'_{tm}R^{-1}_{tm}\tilde{z}_{tm}, \qquad t = 1,...,T-1, \tag{35}$$

$$-B'_{T,m}x_{T-1,m} + A_{Tm}x_{Tm} = n_T\hat{\pi}_{Tm}H'_{Tm}R^{-1}_{Tm}\tilde{z}_{Tm},$$

where the block matrices A_{tm}, D_{tm}, and B_{tm} are given by

$$A_{0m}\bar{\Sigma}^{-1}_{0m}$$

$$A_{tm} = \left(G_{t-1,m}Q_{t-1,m}G'_{t-1,m}\right)^{-1} + n_t\hat{\pi}_{tm}H'_{tm}R^{-1}_{tm}H_{tm}, \qquad 1 \leq t \leq T,$$

$$D_{tm} = F'_{tm}\left(G_{tm}Q_{tm}G'_{tm}\right)^{-1}F_{tm}, \qquad\qquad 0 \leq t \leq T-1, \tag{36}$$

$$B_{tm} = F'_{tm}\left(G_{tm}Q_{tm}G'_{tm}\right)^{-1}, \qquad\qquad 0 \leq t \leq T-1,$$

and the synthetic measurement $\tilde{z}_{tm}$ is defined by

$$\tilde{z}_{tm} = \frac{1}{n_t\hat{\pi}_{tm}} \sum_{r=1}^{n_t} w'_{mtr}z_{tr}, \qquad\qquad 1 \leq t \leq T. \tag{37}$$

The solution of the system (35) is the updated M-step state sequence $\hat{X}^m$. This completes the algorithm derivation for the linear-Gaussian case.

The number $n_t\hat{\pi}_{tm}$ in (36) and (37) represents the expected number of measurements in scan Z_t that are assigned to target m. From (33), it follows that the synthetic measurement $\tilde{z}_{tm}$ is the probabilistic centroid for target m of the measurements in scan Z_t; that is, $\tilde{z}_{tm}$ is the expected measurement for target m at time t. The probabilistic centroid $\tilde{z}_{tm}$ always lies in the convex hull of the scan measurements. (The convex hull of a given set is, intuitively, the smallest convex set that contains the given set as a subset. By definition, the convex hull is the intersection of all convex sets containing the given set.)

The solution of the block tridiagonal system (35) can proceed along strictly algebraic lines, following the methods suggested in section 5.5 of reference 8. Unfortunately, this procedure provides little insight into the resulting algorithm, and it does not yield state error-covariance matrices. A direct connection with available Kalman filtering techniques overcomes both these deficiencies.

18

The connection is established by noting that $\exp\left(Q_{m,X}\right)$ is the PDF of a linear-Gaussian Kalman filter. From (30),

$$\exp\left(Q_{m,X}\right) \propto \varphi_m\left(x_{0m}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}\middle|x_{t-1,m}\right)\prod_{r=1}^{n_t}\mathcal{N}\left[z_{tr}\middle|H_{tm}x_{tm},\left(w'_{mtr}\right)^{-1}R_{tm}\right]\right\}$$

$$\propto \varphi_m\left(x_{0m}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}\middle|x_{t-1,m}\right)\mathcal{N}\left[\tilde{z}_{tm}\middle|H_{tm}x_{tm},\left(n_t\hat{\pi}_{tm}\right)^{-1}R_{tm}\right]\right\},$$

(38)

where the target measurement centroid $\tilde{z}_{tm}$ is defined by (37). These proportional expressions are derived by algebraically manipulating the state-dependent terms of the Gaussian exponents. The right-hand side of (38) is the PDF of a Kalman filter whose plant is identical to the target m process model and whose measurement is the target measurement centroid $\tilde{z}_{tm}$. The conditional independence relationships implicit in likelihood function (38) are displayed graphically in the BIN of figure 2. The directed graph of figure 2 is a subgraph of the full PMHT graph, where the full graph comprises all the BIN fragments depicted in figure 1. Well-known, fixed-interval Kalman filter-smoothing recursions (reference 5, section 7.4) can therefore be used to compute the ML-state sequence. These recursions are used in the PMHT algorithm summary described below in equations (49) through (56). This approach shows that the solution of the block tridiagonal system (35) is identical to the state estimates of a fixed-interval Kalman-smoothing filter.

If the solution of (35) or (37) exists and is unique, then it maximizes the cross-entropy function $Q_{m,X}$ as required by the M-step. The fact that such a solution must be a maximum for $Q_{m,X}$ is not evident from the strictly algebraic development, but it follows immediately from equation PDF (38). In general, however, the solution may not exist and, if it exists, it may not be unique. Significant insight into these possibilities is provided by the PDF (38) because they are related to the observability problems associated with the Kalman filter. Further discussion of convergence issues are given in section 6.
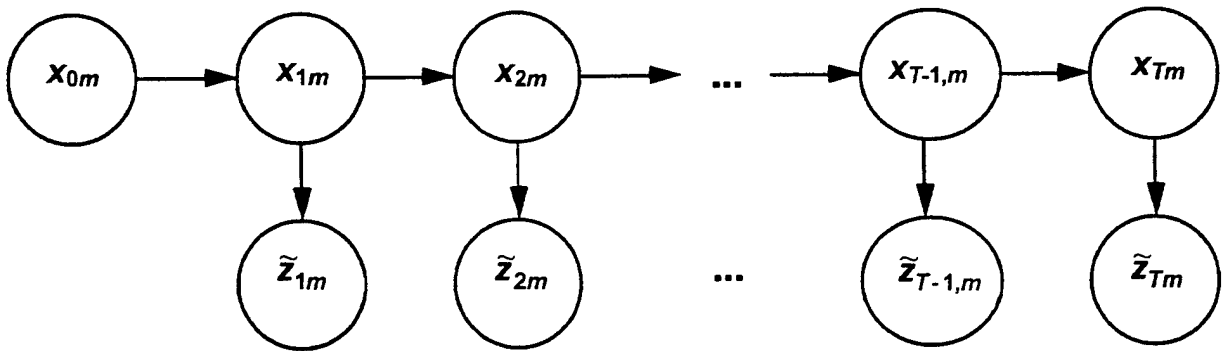

Figure 2. BIN Structure of an Interval Kalman Filter

Error covariances corresponding to the state sequence $\hat{X}^m$ are readily computed using established fixed-interval recursions for the Kalman-smoothing filter (38). These recursions are given by (57) and (58). For further discussion and derivation of these recursions, see page 189 of reference 5 and the references therein. It is stressed, however, that the M-step of the PMHT algorithm does not explicitly use error covariances.

It is intuitively clear that the PMHT error covariances computed from (38) for $\hat{X}^m$ are important in the multitarget tracking application. However, the cross-entropy function does not directly provide a statistical interpretation for error covariances. Section 7 gives such an interpretation in terms of a randomized decision rule applied to the measurement-to-track assignments. As will be seen, this interpretation also holds in the nonlinear non-Gaussian case.

The M-step for the PMHT algorithm is now easily treated for the general case. As discussed in the beginning of this section, for nonlinear non-Gaussian processes, X denotes a parameter set defining target and measurement distributions, and not state estimates as in the linear-Gaussian case. Instead of (38), one obtains from (30) the expression

$$\exp\left(Q_{m,X}\right) = \varphi_m\left(x_{0m}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}\middle|x_{t-1,m}\right)\prod_{r=1}^{n_t}\left\{\zeta_m\left(z_{tr}\middle|x_{tm}\right)\right\}^{w'_{mtr}}\right\}$$

(39)

$$\propto \varphi_m\left(x_{0m}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}\middle|x_{t-1,m}\right)\Xi_m\left(Z_t\middle|x_{tm}\right)\right\},$$

(40)

where, for an appropriate normalization constant c,

$$\Xi_m\left(Z_t\middle|x_{tm}\right) = c\prod_{r=1}^{n_t}\left\{\zeta_m(z_{tr}|x_{tm})\right\}^{w'_{mtr}}$$

(41)

defines a conditional PDF on the full-scan measurement. (The conditional PDF (41) is different from the observer scan PDF (8).) The parameter sequence $\hat{X}^m$ that maximizes (40) is the updated parameter set required by the M-step. The computation of $\hat{X}^m$ requires, in general, using an iterative numerical algorithm. This numerical procedure is conceptually equivalent to a single-target MAP tracker, and its availability is assumed. Whether the MAP tracker is computationally efficient is irrelevant to the M-step: All that is necessary is that it compute the MAP parameter estimate $\hat{X}^m$. The computation of $\hat{X}^m$ using the MAP tracker completes the M-step in the general case.

The conditional PDF (41) is assumed known because it is either known explicitly or because it can be derived from quantities that are given. In the most general case, it is possible to write only

$$z_{tr} = h_{tm}\left(x_{tm}, v_{tm}\right),$$

(42)

20

where the measurement function h_{tm} is given, and the PDF of the measurement noise v_{tm} is assumed known. Theoretically, the PDF $\zeta_m\left(z_{tr}|x_{tm}\right)$ can be derived from (42) and the noise PDF. The conditional PDF $\Xi_m\left(Z_t|x_{tm}\right)$ can then be derived directly from definition (41).

Expressions simpler than (41) are not available in general. If sufficient statistics for the samples Z_t are known for the family of distributions $\left\{\zeta_m^{w'_{mtr}}\right\}$, then it is possible to write the product (41) succinctly. A special case is that of additive Gaussian measurement noise, that is,

$$z_{tr} = h_{tm}\left(x_{tm}\right) + v_{tm},$$
(43)

where v_{tm} denotes white Gaussian noise with covariance R_{tm}. Because (43) is equivalent to

$$\zeta_m\left(z_{tr}\right) = \mathcal{N}\left[z_{tr}|h_{tm}\left(x_{tm}\right), R_{tm}\right],$$
(44)

the PDF (41) can be written

$$\Xi_m\left(Z_t|x_{tm}\right) = \mathcal{N}\left[\tilde{z}_{tm}\left(Z_t\right)|h_{tm}\left(x_{tm}\right), \left(n_t\hat{\pi}_{tm}\right)^{-1}R_{tm}\right],$$
(45)

where the centroid $\tilde{z}_{tm}\left(Z_t\right)$ is given by equation (37). The sufficient statistics are the measurement centroids $\left\{\tilde{z}_{tm}\right\}$ and the weighted covariance matrices $\left\{\left(n_t\hat{\pi}_{tm}\right)^{-1}R_{tm}\right\}$.

The posterior state PDF corresponding to the parameter sequence $\hat{X}^m$ can be derived from the right-hand side of (40) and, subsequently, state estimates derived from the posterior PDF. Error covariances may also be derived from the posterior PDF. However, useful expressions for the state estimates in the general case are unavailable.

5. EXPLICIT STATEMENT OF THE PMHT ALGORITHM IN RECURSIVE FORM

5.1 LINEAR-GAUSSIAN CASE

The PMHT algorithm is summarized in this section for the special case of linear-Gaussian statistics. The batch measurement Z is assumed given. Initialize the target measurement probabilities $\Pi^{(0)} \equiv \left\{ \pi_{tm}^{(0)} \right\}$ so that $\pi_{tm}^{(0)} > 0$. Initialize a target state sequence $\left(x_{0m}^{(0)}, x_{1m}^{(0)}, ..., x_{Tm}^{(0)} \right)$ for each of the M target models. Let $i \geq 0$ denote the PMHT iteration index.

For $i \geq 0$, compute the assignment weights

$$w_{mtr}^{(i+1)} = \frac{\pi_{tm}^{(i)} \; \mathscr{N}(z_{tr} | H_{tm} x_{tm}^{(i)}, R_{tm})}{\sum\limits_{s=1}^{M} \pi_{ts}^{(i)} \; \mathscr{N}(z_{tr} | H_{ts} x_{ts}^{(i)}, R_{ts})},$$

(46)

for $m = 1,...,M$, $t = 1,...,T$, and $r = 1,...,n_t$. Thus, an assignment weight is computed for every target and measurement combination at each scan in the batch. Update the target measurement probabilities

$$\pi_{tm}^{(i+1)} = \frac{1}{n_t} \sum_{r=1}^{n_t} w_{mtr}^{(i+1)}.$$

(47)

A target measurement probability is computed for every target and scan in the batch. Update the target measurement centroids

$$\tilde{z}_{tm}^{(i+1)} = \frac{1}{n_t \pi_{tm}^{(i+1)}} \sum_{r=1}^{n_t} w_{mtr}^{(i+1)} z_{tr}, \qquad t = 1,...,T.$$

(48)

A centroid is computed for each target at each scan in the batch. The effective covariance matrix for the centroid (48) is proportional to the covariance matrix R_{tm}; explicitly,

$$R_{tm}^{(i+1)} = [n_t \pi_{tm}^{(i+1)}]^{-1} R_{tm}, \qquad t = 1,...,T.$$

(49)

Target state sequences for each of the M targets are updated using fixed-interval Kalman smoothers whose inputs are the target measurement centroids. Let the output of the Kalman smoother for target model m at iteration $i + 1$ be denoted by $\left(x_{0m}^{(i+1)}, x_{1m}^{(i+1)}, ..., x_{Tm}^{(i+1)} \right)$. These estimates are computed via a forward and a backward recursion. Initialize intermediate variables of the forward recursion by

$$\hat{y}_{0|0} \equiv \bar{x}_{0m} \tag{50}$$

$$P_{0|0} \equiv \bar{\Sigma}_{0m}. \tag{51}$$

Intermediate variables are dummy variables, so their dependence on model m and iteration number i is suppressed for notational simplicity. The forward recursion is defined for $t = 0, 1, \ldots, T - 1$ by

$$P_{t+1|t} = F_{tm} P_{t|t} F'_{tm} + G_{tm} Q_{tm} G'_{tm}, \tag{52}$$

$$W_{t+1} = P_{t+1|t} H'_{t+1,m} \left\{ H_{t+1,m} P_{t+1|t} H'_{t+1,m} + R^{(i+1)}_{t+1,m} \right\}^{-1}, \tag{53}$$

$$P_{t+1|t+1} = P_{t+1|t} - W_{t+1} H_{t+1,m} P_{t+1|t}, \tag{54}$$

$$\hat{y}_{t+1|t+1} = F_{tm} \hat{y}_{t|t} + W_{t+1} \left\{ \tilde{z}^{(i+1)}_{t+1,m} - H_{t+1,m} F_{tm} \hat{y}_{t|t} \right\}. \tag{55}$$

The updated PMHT state estimate for model m at time T is given by

$$x^{(i+1)}_{Tm} = \hat{y}_{T|T}, \tag{56}$$

and the updated PMHT state estimates for $t = T - 1, \ldots, 1, 0$ are given by the backward recursion

$$x^{(i+1)}_{tm} = \hat{y}_{t|t} + P_{t|t} F'_{tm} P^{-1}_{t+1|t} \left[x^{(i+1)}_{t+1,m} - F_{tm} \hat{y}_{t|t} \right]. \tag{57}$$

Equations (50) through (57) comprise a bank of M Kalman smoothers that are run in parallel for each PMHT iteration; however, these smoothing filters are not independent because they are linked via the assignment weights $\left\{ w^{(i+1)}_{mtr} \right\}$ defined by (46).

Equations (46) through (57) comprise one step of the PMHT algorithm for the special case of linear-Gaussian statistics. The associated error covariance estimates are readily computed. Denote the PMHT error covariance matrix of target m at time t by Σ_{tm}. At time T, $\Sigma_{Tm} = P_{T|T}$. The covariances at other times in the batch are computed by a backward recursion. Explicitly, because $t = T - 1, \ldots, 0$,

$$\Sigma_{tm} = P_{t|t} + P_{t|t} F'_{tm} P^{-1}_{t+1|t} \left[\Sigma_{t+1,m} - P_{t+1|t} \right] P^{-1}_{t+1|t} F_{tm} P_{t|t}. \tag{58}$$

The covariances $P_{t|t}$ and $P_{t+1|t}$ used in (58) are the intermediate variables computed from (52) and (54) during the final PMHT iteration; hence, they depend on the target model m and are different

24

for different targets. Because of the computational complexity of the error covariances (58), it is fortunate that they must be computed only once, after the PMHT algorithm has converged.

Unlike the single-target Kalman filter, the error covariances (58) cannot be computed offline. This result is attributed to the fact that Kalman gains (53) are coupled by assignment interference; that is, the number of measurements each target receives is unknown and must be estimated. The ML estimate of this number is the reciprocal of the coefficient of R_{tm} in equation (49). It is possible to compute the covariances offline only if this coefficient is known. The single-target Kalman filter is thus a degenerate case; that is, the assumption of precisely one target implies perfect knowledge of all measurement assignments.

Numerical problems can arise if one or more of the target measurement probabilities $\pi_{tm}^{(i+1)}$ become small during PMHT iteration. This problem is easily eliminated by rewriting the equations to cancel the common factors of $\pi_{tm}^{(i)}$ implicit in (48). Explicitly, the modified weights are defined as

$$\omega_{mtr}^{(i+1)} \equiv \frac{\mathcal{N}(z_{tr}|H_{tm}x_{tm}^{(i)}, R_{tm})}{\sum_{s=1}^{M} \pi_{ts}^{(i)} \; \mathcal{N}(z_{tr}|H_{ts}x_{ts}^{(i)}, R_{ts})} .$$

(59)

The updated target measurement probability $\pi_{tm}^{(i+1)}$ takes the form

$$\pi_{tm}^{(i+1)} = \overline{\omega}_{mt}^{(i+1)} \pi_{tm}^{(i)},$$

(60)

where the mean measurement weight for target m at time t is defined by

$$\overline{\omega}_{mt}^{(i+1)} \equiv \frac{1}{n_t} \sum_{r=1}^{n_t} \omega_{mtr}^{(i+1)} .$$

(61)

The target measurement centroid defined by (48) is then rewritten in the form

$$\tilde{z}_{tm}^{(i+1)} = \frac{1}{n_t \overline{\omega}_{mt}^{(i+1)}} \sum_{r=1}^{n_t} \omega_{mtr}^{(i+1)} z_{tr} .$$

(62)

Similar modification of equation (52) gives the algebraically equivalent form

$$W_{t+1} = n_{t+1} \pi_{t+1,m}^{(i+1)} P_{t+1|t} H'_{t+1|m} \left\{ n_{t+1} \pi_{t+1,m}^{(i+1)} H_{t+1,m} P_{t+1|t} H'_{t+1,m} + R_{t+1,m} \right\}^{-1} .$$

(63)

The use of equations (59) through (63) eliminates numerical difficulties associated with small target measurement probabilities.

Numerical difficulties may be encountered with generic Kalman filtering if covariance matrices are computed explicitly. These kinds of difficulties are very closely related to the well-known numerical problems that can arise with linear least-squares computations if the normal equations are solved directly. Strictly numerical techniques such as QR factorization or singular value decomposition (SVD) (see reference 8) are very useful in practical computations; however, good numerical methods cannot prevent problems that arise because of observability limitations. Numerically stable forms of the Kalman filter have been studied. For example, see reference 9.

5.2 GENERAL CASE

Initialize the PMHT target measurement probabilities $\Pi^{(0)} > 0$ and a target parameter sequence $\left(x_{0m}^{(0)}, x_{1m}^{(0)}, \ldots, x_{Tm}^{(0)}\right)$ for each of the M-targets. For $i \geq 0$, compute the assignment weights

$$w_{mtr}^{(i+1)} = \frac{\pi_{tm}^{(i)} \, \zeta_m(z_{tr}|x_{tm}^{(i)})}{\sum\limits_{s=1}^{M} \pi_{ts}^{(i)} \, \zeta_s(z_{tr}|x_{ts}^{(i)})},$$

(64)

for $m = 1,\ldots,M$, $t = 1,\ldots,T$, and $r = 1,\ldots,n_t$. The target measurement probabilities are updated exactly as before using equation (47). Define the conditional PDF

$$\Xi_m^{(i+1)}(Z_t|x_{tm}) = c^{(i+1)} \prod_{r=1}^{n_t} \{\zeta_m(z_{tr}|x_{tm})\}^{w_{mtr}^{(i+1)}},$$

(65)

where $c^{(i+1)}$ is an irrelevant normalization constant. The parameters $\{x_{tm}\}$ are unknown variables in (65). Apply the assumed available single-target MAP tracker to compute updated parameters $\left(x_{om}^{(i+1)}, x_{1m}^{(i+1)}, \ldots, x_{Tm}^{(i+1)}\right)$ such that

$$\varphi_m\left(x_{0m}^{(i+1)}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}^{(i+1)}\Big|x_{t-1,m}^{(i+1)}\right)\Xi_m^{(i+1)}\left(Z_t\Big|x_{tm}^{(i+1)}\right)\right\}$$

$$= \max_{\{x_{0m},x_{1m},\ldots,x_{Tm}\}} \varphi_m\left(x_{0m}\right)\prod_{t=1}^{T}\left\{\varphi_m\left(x_{tm}\Big|x_{t-1,m}\right)\Xi_m^{(i+1)}\left(Z_t\Big|x_{tm}\right)\right\}.$$

(66)

The updated state parameters $\left\{x_{tm}^{(i+1)}\right\}$ are MAP estimates, and they depend on the current estimates (64) of the assignment interference. The PMHT algorithm is a multitarget tracking algorithm that uses the single-target MAP tracker as an essential building block; thus, the PMHT algorithm is conceptually equivalent to an iteratively reweighted bank of MAP trackers.

The computational complexity and numerical characteristics of the PMHT algorithm are closely related to the complexity and numerical accuracy of the single-target MAP tracker. The total computational effort is proportional to the product of number of targets M, the number of PMHT iterations needed to estimate the assignment interference to sufficient accuracy, and the effort needed to solve the single-target MAP parameter estimates from (66). Precise estimates of computational effort depend on the particular application.

In the special case of independent, additive white Gaussian state and measurement noises, the maximization problem (66) is equivalent to a nonlinear least-squares problem. This is easily seen using (45) and the fact that

$$x_{tm} = f_{t-1,m}(x_{t-1,m}) + w_{t-1,m} \tag{67}$$

is equivalent to

$$\varphi_m(x_{tm}|x_{t-1,m}) = \mathcal{N}\left[x_{tm}\Big| f_{t-1,m}(x_{t-1,m}), Q_{t-1,m}\right], \tag{68}$$

where $Q_{t-1,m}$ is the covariance of the white Gaussian noise $w_{t-1,m}$. The state evolution function $f_{t-1,m}$ and the *a priori* distribution at the initial batch time $t = 0$ are assumed given. Using the Gaussian *a priori* densities (4), the centroids (48), and the effective covariances (49) gives the following nonlinear least-squares problem equivalent to (66):

$$\min_{\{x_{0m}, x_{1m}, \ldots, x_{Tm}\}} \left\{ \begin{array}{l} (x_{0m} - \bar{x}_{0m})' \overline{\Sigma}_{0m}^{-1} (x_{0m} - \bar{x}_{0m}) \\ + \sum_{t=1}^{T} \left[x_{tm} - f_{t-1,m}(x_{t-1,m}) \right]' (Q_{t-1,m})^{-1} \left[x_{tm} - f_{t-1,m}(x_{t-1,m}) \right] \\ + \sum_{t=1}^{T} \left[\tilde{z}_{tm}^{(i+1)} - h_{tm}(x_{tm}) \right]' (R_{t-1,m}^{(i+1)})^{-1} \left[\tilde{z}_{tm}^{(i+1)} - h_{tm}(x_{tm}) \right] \end{array} \right\}. \tag{69}$$

The solution of (69) is the updated parameter vector $\left(x_{0m}^{(i+1)}, x_{1m}^{(i+1)}, \ldots, x_{Tm}^{(i+1)} \right)$. In this case, a special purpose algorithm such as Levenberg-Marquardt (reference 10, chapter 14.4) can be used to solve (69) and can serve as the single-target MAP tracker. Further consideration of these issues lies outside the scope of this report.

6. THEORETICAL ASPECTS OF THE PMHT ALGORITHM

6.1 CONVERGENCE PROPERTIES OF THE PMHT ALGORITHM

Algorithms derived from the EM method have two useful properties. These algorithms are guaranteed to converge under very mild assumptions, and they typically get very close to the limit point during their first few iterations, even when poorly initialized. It is known that the convergence rate of EM algorithms is only asymptotically linear; however, these results do not explain the practical observation of rapid early-convergence rate. Considering the various modeling approximations that are made in practical problems, it is unlikely that exact solutions are really necessary for the multitarget tracking application — a good approximation is typically sufficient.

If greater precision is required than is provided by computing a few steps of the EM method, it may be wiser to use a compound computational procedure: Begin with EM to get close to the solution quickly, and then switch to a more rapidly convergent algorithm (e.g., Newton-Raphson). The choice of optimization algorithm is theoretically irrelevant, as long as ML estimates are computed from the marginal PDF (22), the basis of the PMHT formulation. The EM method was chosen in this study for this estimation problem because it is structurally compatible with the notions of the single-target Kalman filter.

At every iteration, the PMHT algorithm is guaranteed to either increase the likelihood of the marginal PDF $P(Z,X)$, defined by equation (22), or to be converged to a stationary point of the marginal PDF. This property alone does not guarantee convergence, however. For example, if the marginal is unbounded above, the PMHT iterates may be attracted to infinite singularities of the likelihood function, and this can happen even though each iterate is well defined mathematically. The general theory (see reference 6) shows that boundedness from above guarantees convergence to either an ML point or a stationary point of the marginal PDF (22). Although stationary points may seem of little practical importance, convergence to local ML solutions instead of a global ML solution may sometimes cause difficulties. Section 9 discusses a meaningful local solution for an example involving two crossing targets. In a multitarget tracking application, PMHT requires sliding the batch as new measurement scans arrive; consequently, it is probable that careful initialization of the starting point of the PMHT algorithm for each new batch, using the solution from the previous batch, will result in convergence to global ML solutions. This aspect of PMHT requires further research.

Under linear-Gaussian statistics, an appropriate observability condition guarantees that the PMHT marginal PDF is bounded. One such condition is that the individual target-error-covariances are positive definite throughout the batch for all possible measurement-to-track assignments K. There is, however, an easily overlooked subtlety. Consider target m in isolation. If even one of the process gain matrices $\{G_{tm}\}$ (first defined in equation (1)) is less than full rank, then the PDF (38) is unbounded because at least one of the process PDFs $\{\varphi_m(x_{tm}|x_{t-1,m})\}$ is a degenerate Gaussian distribution. It is a simple matter to require all the matrices $\{G_{tm}\}$ to have

full rank to avoid the degeneracy; however, most kinematic process models have deficient rank, so this requirement is overly restrictive. (Increases in the marginal PDF are easily verified numerically under full-rank conditions, and this serves in practice to check the implementation of the PMHT algorithm.) Fortunately, the marginal PDF (22) is bounded above on the linear subspace in which the individual target processes are confined. Because the PMHT estimates are themselves confined to the same linear subspace, the PMHT algorithm is guaranteed to converge. Further study of this topic is outside the scope of this report.

6.2 FISHER INFORMATION MATRICES FOR PMHT

The PMHT error-covariance matrices, or inverse FIMs, for the target states appear incidentally as a by-product of the M-step of the PMHT algorithm, and it is unclear from the EM method context exactly how these matrices should be interpreted. The difficulty stems from the fact that the cross-entropy function is defined for its desirable analytical properties, and not derived from statistical assumptions concerning the measurements. Consequently, quantities, such as FIMs, that are computed from the cross-entropy function lack statistical interpretation in the general theory of the EM method. The purpose of this subsection is to present one such statistical interpretation for the FIMs defined for the PMHT algorithm. The interpretation presented is not theoretically completely satisfactory, and, therefore, further work on this topic is justified.

The PDF (40) is easily interpreted statistically in terms of a Markov-state sequence and the scan PDF $\Xi_m\left(Z_t \middle| x_m\right)$. The difficulty lies in the statistical interpretation of $\Xi_m\left(Z_t \middle| x_m\right)$. $\Xi_m\left(Z_t \middle| x_m\right)$ is not equivalent to a PDF that corresponds to making hard assignments because the weights (24) do not converge to hard assignment decisions; that is, no weight converges to either zero or one. The interpretation adopted is that $\Xi_m\left(Z_t \middle| x_m\right)$ is the PDF of a randomized decision rule defined over the ensemble of all possible batch measurements $\{Z\}$. Each member of the ensemble is assumed to have exactly the same number of measurements in every scan as the given batch measurement Z, the only measurement available in practice. For each member of the ensemble, hard measurement-to-track assignments are made; however, assignments are discrete random variables, and the probability of different assignments are given by the weights (24). The PMHT algorithm and the given batch measurement Z are used to obtain the weights (24), and these weights are applied to the ensemble $\{Z\}$. The appropriate PDF for a randomized decision rule conceptualized in this manner is the scan PDF $\Xi_m\left(Z_t \middle| x_m\right)$. The FIM corresponding to this interpretation of the measurements is computed by the PMHT algorithm from the function (40).

FIMs for unbiased estimators are derived directly from a given joint PDF. The joint PDF may incorporate nonrandom parameters, random parameters (see reference 11, pp. 72, 84-5), and mixed random and nonrandom parameters. Unfortunately, the FIM is undefined for the PMHT observer PDF (17) because the full parameter list comprises the random parameters X and K, and the gradient with respect to K does not exist because K is discrete. This difficulty is circumvented by deriving the FIM from the observer marginal PDF $P(Z,X;\Pi)$, defined by (22), by treating it as

a joint PDF whose full parameter list comprises the random parameter X and the nonrandom parameter Π. This FIM, called the marginal FIM, for X and Π is defined by

$$J(\overline{\Pi}) \equiv - \iint\limits_{Z\ X} \left[\nabla_{X,\Pi} \left\{ \nabla'_{X,\Pi} \ \log \ P(Z,X{:}\Pi) \right\} \right] P(Z,X{:}\Pi) \, dZ \, dX \Bigg|_{\Pi=\overline{\Pi}} , \qquad (70)$$

where $\overline{\Pi}$ is the true value of Π. The marginal FIM for X and Π is very difficult to evaluate explicitly in the multitarget case because the targets are coupled by assignment interference, as discussed in section 2. A study of the marginal FIM is outside the scope of this report.

7. VARIATIONS AND EXTENSIONS OF PMHT

7.1 TARGET MEASUREMENT PROBABILITIES, Π

Several interesting variations of the PMHT likelihood function have cross-entropy Q functions that are easily obtained. Because of the decoupling of the estimation steps for Π and X, these variations result in only slight changes in the PMHT algorithm and are readily implemented. This subsection discusses two alternatives for Π.

Some applications may require that the target measurement probability vectors π_t be identical across the batch, but not assumed known *a priori*. The BIN for this variation is unchanged from that of figure 1 because it is equivalent to a constraint on the parameterization of the joint PDF; that is, the conditional independence assumptions of the random variables are unaltered by this constraint, but are only reparameterized. In this case, the $\left\{Q_{t,\Pi}\right\}$ terms (29) are modified by substituting

$$\pi_s \equiv \pi_{1s} = \pi_{2s} = \cdots = \pi_{Ts}, \qquad s = 1,\ldots,M. \tag{71}$$

The probability π_s is used only in this section, and it is not to be confused with the probability vector π_t given by equation (9). The constraint (71) modifies the M-step slightly. Using the notational conventions of section 5, the updating recursions (46) through (49) are replaced by the recursions

$$w_{mtr}^{(i+1)} = \frac{\pi_m^{(i)}\, \mathcal{N}(z_{tr}|H_{tm}x_{tm}^{(i)}, R_{tm})}{\displaystyle\sum_{s=1}^{M} \pi_s^{(i)}\, \mathcal{N}(z_{tr}|H_{ts}x_{ts}^{(i)}, R_{ts})}, \tag{72}$$

$$\pi_m^{(i+1)} \equiv \frac{1}{N} \sum_{t=1}^{T}\sum_{r=1}^{n_t} w_{mtr}^{(i+1)}, \tag{73}$$

$$\tilde{z}_{tm}^{(i+1)} = \left[\sum_{r=1}^{n_t} w_{mtr}^{(i+1)}\right]^{-1} \sum_{r=1}^{n_t} w_{mtr}^{(i+1)} z_{tr}, \tag{74}$$

$$R_{tm}^{(i+1)} = \left[\sum_{r=1}^{n_t} w_{mtr}^{(i+1)}\right]^{-1} R_{tm}, \tag{75}$$

respectively, where the constant N in (73) is given by

$$N = n_1 + \cdots + n_T.$$

(76)

The recursions (50) through (57) are unchanged. The appropriate equations in the general case are derived similarly. This variation is not considered further.

If the target measurement probabilities Π are assumed known *a priori*, for example, $\pi_{tm} = 1/M$ for all t and m, then the Q function for this case comprises the terms $\{Q_{m,X}\}$ in equation (30), but it does not include the terms $\{Q_{t,\Pi}\}$. The M-step in this case updates X in exactly the same way as before, but it uses the specified *a priori* Π instead of the estimated $\hat{\Pi}$. More generally, the Q function is easily modified to incorporate a specified *a priori* distribution on Π, if one is available. These variations are not considered further in this report.

7.2 ADAPTIVE COVARIANCE ESTIMATION

The measurement and target-covariance matrices $\{R_{tm}\}$ and $\{Q_{tm}\}$ are assumed known in the above development for the linear-Gaussian case. When these matrices are unknown, they can be treated as parameters to be estimated from the batch measurement data. Although covariance estimation problems are outside the scope of this report, the necessary conditions are easily obtained and are presented below because of their potential use.

The EM approach to covariance estimation does not alter the likelihood structure (15) or the cross-entropy Q function, but it does significantly alter the M-step because the covariance matrix estimates are coupled to the state estimation equations. The Q function for this case is identical to (28), but its gradient with respect to each of the matrices R_{tm} and Q_{tm} must now be taken and set to 0. Taking the gradient of (30) with respect to R_{tm}, applying the identity (reference 5, section 2.14.2)

$$\nabla_\Sigma \log \mathcal{N}(x|\mu,\Sigma) = -\Sigma^{-1} + \Sigma^{-1}(x-\mu)(x-\mu)'\Sigma^{-1},$$

and setting the result to zero gives the estimate

$$\hat{R}_{tm} = \left(\sum_{r=1}^{n_t} w'_{mtr}\right)^{-1} \sum_{r=1}^{n_t} w'_{mtr}\left(\tilde{z}_{tm} - H_{tm}\hat{x}_{tm}\right)\left(\tilde{z}_{tm} - H_{tm}\hat{x}_{tm}\right)'.$$

(77)

The estimate (77) is full rank (with probability one) if the dimension N_z of the measurements is less than the number of measurements. Similarly, assuming the matrices G_{tm} are nonsingular, the gradient of (30) with respect to Q_{tm} gives the estimate

$$\hat{Q}_{tm} = \left[G_{tm}^{-1}\left(\hat{x}_{t+1,m} - F_{tm}\hat{x}_{tm}\right)\right]\left[G_{tm}^{-1}\left(\hat{x}_{t+1,m} - F_{tm}\hat{x}_{tm}\right)\right]' . \tag{78}$$

The estimate (78) is full rank only if the dimension of the target-state variable N_{Xm} is exactly 1.

Rank difficulties for covariance matrix estimation within a scan are significantly reduced by requiring the covariance matrices to be stationary throughout the batch, that is, $R_{tm} \equiv R_m$ and $Q_{tm} \equiv Q_m$ for all t. This additional requirement gives the measurement covariance estimate

$$\hat{R}_m = \left(\sum_{t=1}^{T}\sum_{r=1}^{n_t} w'_{mtr}\right)^{-1} \sum_{t=1}^{T}\sum_{r=1}^{n_t} w'_{mtr}\left(\tilde{z}_{tm} - H_{tm}\hat{x}_{tm}\right)\left(\tilde{z}_{tm} - H_{tm}\hat{x}_{tm}\right)' . \tag{79}$$

The covariance matrix estimate (79) is theoretically of full rank (with probability one) only if the the total number N of individual measurements is greater than the dimension N_z of the measurement variables. Similarly, assuming $G_{tm} \equiv G_m$, the estimate for the process covariance is given by

$$\hat{Q}_m = \frac{1}{T}\sum_{t=1}^{T}\left[G_m^{-1}\left(\hat{x}_{t+1,m} - F_{tm}\hat{x}_{tm}\right)\right]\left[G_m^{-1}\left(\hat{x}_{t+1,m} - F_{tm}\hat{x}_{tm}\right)\right]' . \tag{80}$$

The estimate (80) is the average of the estimates (78), and it is theoretically of full rank (with probability one) if T is greater than the dimension N_{Xm} of the state variables for target m.

Equations (77) through (80) are nonlinearly coupled with the M-step state estimates, but they are decoupled into M smaller systems, one for each target model. An explicit solution of these nonlinear equations is not available. However, it is straightforward to show that the estimates (79) and (80) are the basic equations for the generalized EM algorithm (reference 6) for estimating covariances.

Implementing the covariance estimates in the mathematical forms given above will unnecessarily square the numerical condition number of the underlying "data matrix." Numerical ill-conditioning arising from this source is completely avoided by the use of QR methods (reference 12). QR algorithms are stable numerically and require very little additional computational effort, so they are highly recommended for avoiding sample covariance matrix formation. Another source of numerical ill-conditioning for the estimate (79) is the reduction of the effective numerical rank that occurs when one or more of the summand coefficients becomes extremely small. QR algorithms can limit, but not fully overcome, ill-conditioning due to the dynamic range of these coefficients.

8. COMPARISONS WITH OTHER METHODS

8.1 MAXIMUM LIKELIHOOD DATA ASSOCIATION

Avitzour, in reference 13, considers data association using an ML algorithm that is based on the EM method. Several of the ideas described are closely related to the ideas developed in this report; however, there are also significant differences. The points of similarity are discussed first. Both methods consider data association for multitarget tracking from a probabilistic perspective, and both consider batch measurements. Both also use the EM method to estimate target states and assignment probabilities. Moreover, both recognize the significance in the multitarget tracking application of the assignment probabilities (24). Avitzour is "not concerned with finding the correct association of measurements to targets/clutter" because assignment probabilities are "obtained as a by-product" of state estimation. This view of measurement assignments is identical to the view expressed in this report. Finally, both discuss data association using hard assignments as joint MAP estimation of states and discrete random variables called assignments and derive MAP state estimates from a marginal distribution over the assignments using the EM method.

One of the important differences between reference 13 and this report is that Avitzour does not fully interpret the EM method in the context of the multitarget tracking application. In particular, an observer whose state includes the assignments is not defined explicitly in reference 13. It is this important feature of the observer that distinguishes PMHT from joint probabilistic data association (JPDA) and MHT, enables the EM method to be applied effectively, and provides a unified cogent exposition of the probabilistic issues.

Error-covariance estimates are estimated and discussed in this report, but are absent from reference 13. The primary reason this study can estimate covariances is the recognition that the exponential of the cross-entropy function is proportional to the likelihood structure of a Kalman filter (compare with equations (38) and (40)) that treats an entire measurement scan Z_t as a single measurement. This important relationship between the M-step and the Kalman filter is not recognized in reference 13.

The central role played by conditional independence in multivariate statistical problems such as multitarget tracking is not emphasized by Avitzour. The BIN graphical representation of conditional independence structure of the PMHT approach is novel to this study and is not used in reference 13. The BIN graph helps clarify the differences between PMHT and JPDA.

In reference 13, Avitzour does not allow process noise, although he does allow moving targets. This is not a serious difference, however, as it is possible to extend his method to include process noise. Finally, Avitzour presents a potentially useful termination criterion for the EM method. This termination criterion may be useful in this study also, but it is not used here.

8.2 JOINT PROBABILISTIC DATA ASSOCIATION

In the case of linear-Gaussian statistics, the PMHT algorithm resembles the JPDA algorithm; however, PMHT is fundamentally different from JPDA. The resemblance is closest for a unit batch of size ($T = 1$), which is the batch size assumed for PMHT in the present discussion. Linear-Gaussian statistics are assumed thoughout this discussion.

Both PMHT and JPDA assume that the measurements within a scan are independent samples of a random variable with a mixture Gaussian PDF. JPDA assumes *a priori* that the scan measurement mixture is uniform-Gaussian, that is, the mixing proportions of the Gaussian components are equal and their means are the predicted measurement means. JPDA use of predicted states is one source of track bias. In contrast to JPDA, PMHT uses ML estimates of the mixing proportions and the means of the Gaussian components in the scan measurement mixture. Because targets may not be equally represented by the measurements, PMHT use of nonuniform mixing proportions permits improved modeling of scan measurements. The ML estimates of the component means are the estimated target measurement means for the current scan. Because PMHT uses estimated – as opposed to predicted (as in JPDA) – target measurement means, PMHT target state estimates should often have smaller biases, smaller critical separation distances at which tracks coalesce, and better tolerance to target maneuvers than JPDA.

JPDA first tracks measurements and then smoothes the resulting state estimates using convex combinations. The convex combinations yield the output state estimates and their corresponding error covariances. Convex combinations occur in the JPDA state space because of the way in which JPDA treats assignments as events. In contrast, PMHT first smoothes measurements using convex combinations, and then tracks the smoothed measurement. The convex combinations of measurements are the probabilistic measurement centroids (62), and the centroids constitute synthetic measurements to be tracked. Convex combinations occur in the PMHT measurement space because assignments are incorporated into the observer (compare with equation (22)).

PMHT error covariances are different from those of JPDA. Covariances computed by JPDA are readily interpreted statistically because JPDA treats assignments as events. Unfortunately, a similar interpretation is not applicable to PMHT covariances. The PMHT covariances are tied implicitly to the use of the EM method to compute ML state estimates from the marginal PDF (22). For further discussion see section 6.2.

Measurement gates are used to censor measurements that are too far from predicted measurements to be associated with targets. Gates will probably always be needed to protect theoretical mathematical models from the vagaries of real data as well as to improve computation time in practical systems. JPDA is explicitly formulated using gates, but PMHT does not use gating in its formulation. However, gates are readily derived for PMHT by thresholding the assignment weights (24).

38

Gates are equivalent to applying zero weights to certain measurements. JPDA and PMHT discount zero-weighted measurements in very different ways. In JPDA, the convex combination of state estimates is unaffected by the addition of zero-weighted estimates. In PMHT, zero-weighted measurements contribute a multiplicative factor of +1 to the conditional PDF defined by equation (41). Consequently, zero-weighted measurements do not affect PMHT state estimates. The robustness of the PMHT and JPDA state estimates to measurements with small weights (i.e., outliers) remains to be studied.

8.3 GAUSSIAN SUM TRACKING FILTERS

Sorenson and Alspach (references 14 and 15) describe a Gaussian sum tracking filter for nonlinear estimation. This approach is also discussed in reference 5, section 8.4. The problem they studied was restricted to single targets – multiple targets are not mentioned. The main reason to mention multiple targets in this report is that the PMHT marginal PDF defined by equation (22) can be thought of as a Gaussian sum approximation of a nonlinear measurement PDF. This interpretation suggests that the PMHT approach may be applied to general nonlinear estimation problems.

The general nonlinear estimation problem is investigated in references 14 and 15 as a problem in the approximation of the posterior PDF. The form of PDF approximation used is that of a Gaussian mixture, so this approach is fundamentally different from the extended Kalman filter. In reference 14, the main problem discussed is that of a system with linear process and measurement equations with additive non-Gaussian noises. A Gaussian mixture approximates the prior density, and the additive process and measurement noises have PDFs that are approximated by Gaussian mixtures. Under these assumptions, the posterior density is shown to be exactly a Gaussian mixture. In reference 15, nonlinear process models are approximated by Gaussian mixtures by linearizing the process model about the mean of each Gaussian in the mixture. The primary conceptual drawback to this approach is that the number of Gaussian components in the mixture representing the posterior PDF grows rapidly as time evolves; therefore, pruning terms from the Gaussian mixture is necessary in practice.

Sorenson and Alspach estimated the parameters of the approximating mixtures by applying gradient-based numerical methods because, at the time their papers were written, the method of EM had not been developed in generality and was largely unknown outside the statistical community. The primary similarity with PMHT lies in their recognition that the state estimate produced by the Gaussian sum tracking filter is obtained as a convex combination of several linear-Gaussian filters. The filters Sorenson and Alspach refer to, however, are independent filters that are not coupled by assignment interference.

8.4 CENTROID GROUP TRACKING

Blackman, in section 11.2 of reference 16, discusses tracking groups of closely spaced targets by tracking the centroid of the group. This application is a specialized multitarget tracking problem, and the approach described is quite different from PMHT. Nonetheless, there are

similarities in certain details. In particular, the centroids of the measurements are used as synthetic measurements for updating the group track centroid. It is necessary in this approach to gate the measurements before forming the measurement centroids, and such gating is tantamount to making hard assignments of measurements to different groups. The PMHT equivalent of hard assignments and gates is to threshold the PMHT weights (24). There are also similarities between determining the number of targets in a group and the choice of the parameter M in the PMHT algorithm, a model-order selection problem that falls outside the scope of this report.

Formation group tracking is also discussed by Blackman in reference 16, section 11.3. In formation tracking the targets are not assumed to be independent, and the availability of information concerning their coordinated behavior is exploited to advantage. Although the PMHT approach is not limited to independent targets, this topic also falls outside the scope of this report.

9. CROSSING TARGETS WITH LINEAR-GAUSSIAN STATISTICS

Section 9 provides an example that demonstrates PMHT ability to estimate measurement-to-track assignments. It is given that two targets ($M = 2$) are present and constrained to move in the xy plane. For simplicity, positions are given in meters. Both target tracks begin at $t = 0$, are simulated without process noise, and have constant velocity. Target 1 is heading 45° from the +x-axis, has speed +1 m/s, and begins its track at the origin (0, 0) of the x-y-plane. Target 2 is heading +90° from the +x-axis, has speed $\sqrt{2} / 2 \approx 0.707$ m/s, and its track begins at +12.5 m on the x-axis. Given this geometry, the targets become superposed at $12.5\sqrt{2} \approx 17.68$ s. Measurement scans are taken at 1-second intervals, and the batch length is $T = 25$ s. One measurement is simulated from each target for $t = 1,2,...,25$ s. No measurements are simulated at $t = 0$. For simplicity, false alarms are not simulated, and each target generates exactly one measurement in the simulation. The number of measurements per scan is always $n_t = 2$, so that the total number of measurements in the batch is 50. The state vector of each target comprises the x- and y-components of position and velocity. For estimation, the PMHT target state models are

$$
\begin{bmatrix} x_{m,t+1} \\ \dot{x}_{m,t+1} \\ y_{m,t+1} \\ \dot{y}_{m,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{mt} \\ \dot{x}_{mt} \\ y_{mt} \\ \dot{y}_{mt} \end{bmatrix} + \sigma^2_{state} \begin{bmatrix} w^1_{mt} \\ w^2_{mt} \\ w^3_{mt} \\ w^4_{mt} \end{bmatrix}, \qquad m = 1,2,
$$

(81)

where the process noise is white Gaussian with covariance matrix given by the 4 x 4-identity matrix scaled by the dimensionless quantity $\sigma^2_{state} = 10^{-2}$. The Gaussian *a priori* distributions (compare with equation (4)) for targets 1 and 2 have mean vectors

Target 1: (0 m, a m/s, 0 m, a m/s)
Target 2: (12.5 m, 0 m/s, 0 m, a m/s),

where the dimensionless constant $a = \sqrt{2} / 2$. The covariances of the a priori distributions are equal to the 4 x 4-diagonal matrix $diag[16 \text{ m}^2, 9 \text{ m}^2 / \text{s}^2, 16 \text{ m}^2, 9 \text{ m}^2 / \text{s}^2]$ for both targets.

Prior distributions are centered on the simulated target state at $t = 0$, but they have large uncertainties in position and velocity. Because the simulation is defined in such a way that each target generates one and only one measurement, the measurement equations can be written in the form

$$z_{tr} \equiv \begin{bmatrix} z_{tr}^1 \\ z_{tr}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{rt} \\ \dot{x}_{rt} \\ y_{rt} \\ \dot{y}_{rt} \end{bmatrix} + \sigma_z^2 \begin{bmatrix} v_{rt}^1 \\ v_{rt}^2 \end{bmatrix}, \qquad r = 1,2,$$

(82)

where the measurement noise is white Gaussian with covariance matrix given by the 2 x 2-identity matrix scaled by the dimensionless quantity σ_z^2. Equation (82) is obtained by substituting $r = m$ in equation (7); however, the target labeling implicit in (82) is not used by the PMHT algorithm. It is unnecessary to randomly shuffle the measurements $\{z_{t1}, z_{t2}\}$ in the data file before calling the PMHT algorithm because data order does not alter the algorithm, as can be seen from equations (46) through (48).

The measurement scale σ_z^2 in equation (82) is chosen to control the "effective" duration of the track-crossing time. The duration of the track-crossing event is defined to be the total time the two targets lie within three measurement noise standard deviations in the x-y-plane. Given the above geometry, it is readily seen that the crossing duration, denoted by t_c, is linearly proportional to the measurement standard error; specifically, the $t_c = 6\sqrt{2} \ \sigma_z \approx 8.485 \ \sigma_z$. The measurement scale σ_z^2 is selected so that $t_c = 2 \ s$; hence, $\sigma_z = 0.2357$.

Targets 1 and 2 are initialized in the correct states for all time $t = 0,1,...,25$. This no-target-error initialization is a best-case scenario for determining the number of iterations to convergence. Using an absolute error convergence criterion on the sequence of overall likelihood function iterates, the PMHT algorithm converged in 38 iterations when the convergence error $\varepsilon = 0.0001$. The final value of the log likelihood, $\log_e P(Z,X)$, was +392.92. Experience with the PMHT algorithm reveals that it is very robust to poor target initialization; therefore, perfect initialization is unimportant in this example. (With clutter, however, sensitivity to poor target initialization is an important topic. The form of the PMHT algorithm presented here assumes no clutter in order to focus clearly and solely on the assignment problem. Incorporating clutter models into the PMHT approach is straightforward and is the subject of an ongoing investigation.)

Figure 3a depicts the true track positions overlaid with the measurements, and figure 3b depicts the true track positions overlaid on the PMHT-estimated track positions. Figure 4 depicts the root-mean-square error in the estimated track position and speed components for each target. The PMHT assignment weights are given in figure 5 for both targets and each measurement. The estimated weights are clearly indicative of the correct assignments (known only via simulation) everywhere except in the crossover regime. During crossover, the evidence from the data is inconclusive concerning correct assignments, and this uncertainty is reflected by the nearly equal weights given measurements in this regime.
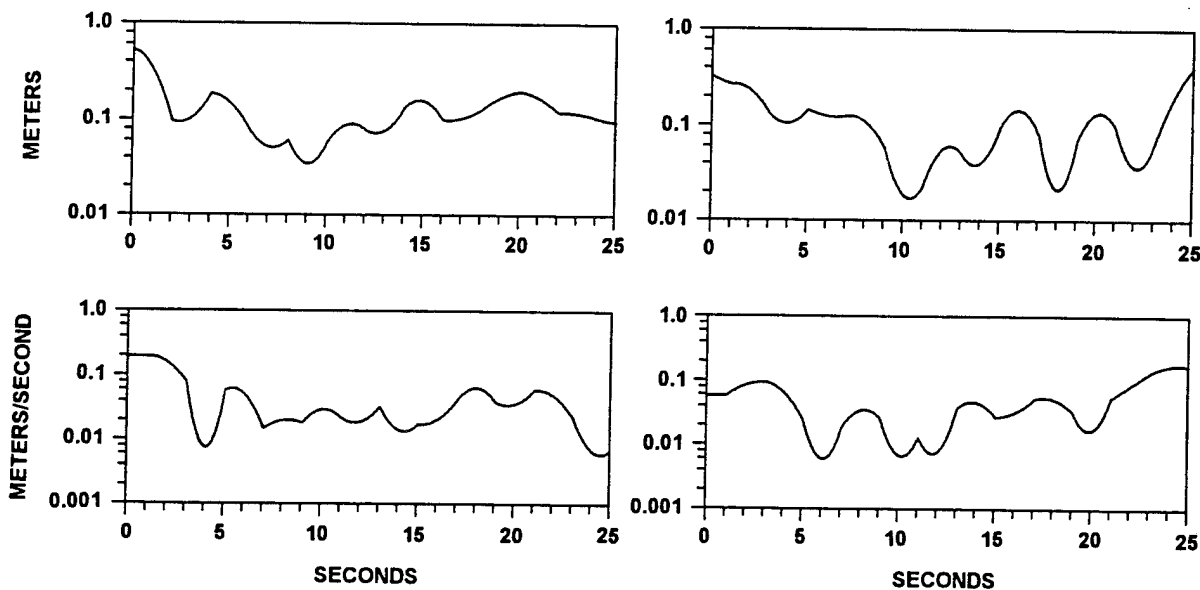
42
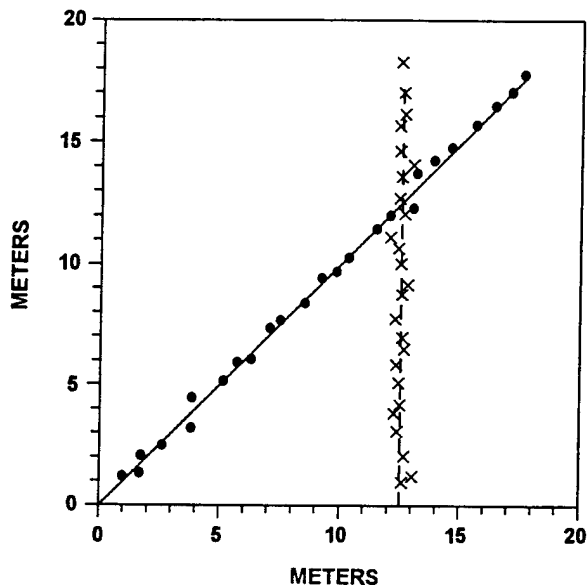
a. Simulated Target Tracks and Measurements
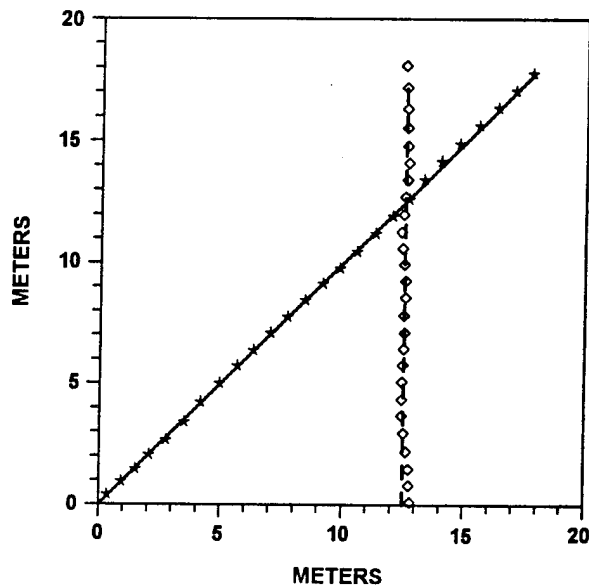
b. Simulated and Estimated Target Tracks

Figure 3. Two Target Crossing Track Example with $t_c = 2$

a. Target 1 Position Error (Top)
and Speed Error (Bottom)

b. Target 2 Position Error (Top)
and Speed Error (Bottom)

Figure 4. PMHT Position and Speed Estimation Errors for Each Target

a. Weights for Target 1 with z_{t1} (Top)
and z_{t2} (Bottom)

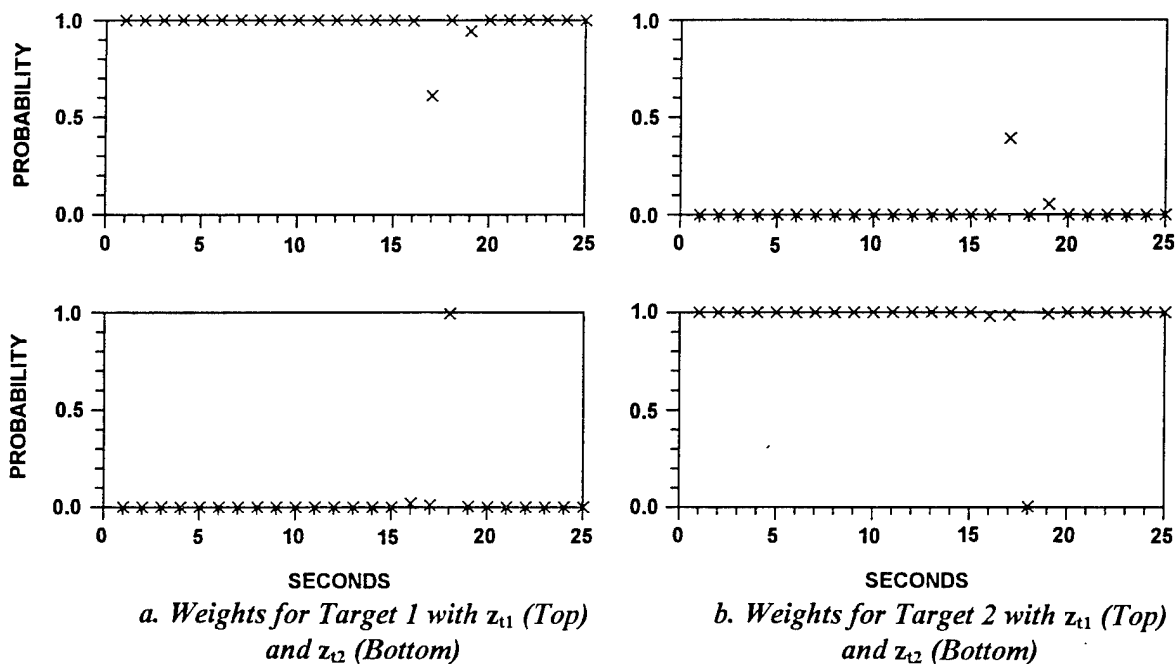b. Weights for Target 2 with z_{t1} (Top)
and z_{t2} (Bottom)

Figure 5. Converged PMHT Assignment Weights (18) for Two Measurements (82)

Poor PMHT initialization may result in convergence to suboptimal solutions; i.e., the PMHT algorithm may converge to only a local ML solution – not the global solution – to the tracking problem. In the multitarget tracking application, these suboptimal solutions may be meaningful because more than one interpretation of the measurements is often possible. A multiplicity of meaningful local solutions appears to be an important feature of good probabilistic models of the problem: If the observer PDF has only one peak, namely the global peak, then it does not contain within it alternative interpretations of the measurements. Alternative interpretations are conceived as local ML solutions whose likelihoods, determined by the PMHT algorithm, are significantly lower than the likelihood of the global solution. This view is discussed quantitatively for the two target-crossing examples presented above.

Two ML solutions, one global and the other local, are given in table 1. These solutions were obtained using identical *a priori* distributions (for all values of t_c) and measurement sets (for each value of t_c), but different initializations of the PMHT algorithm. The initializations used to obtain these solutions were obtained by swapping portions of the true tracks (known from simulation) in the obvious manner. The solution likelihoods given in table 1 are estimated using the PMHT algorithm and are values of the marginal PDF (22). The number of iterations required for PMHT convergence is also given in table 1. The three columns in table 1 correspond to easy ($t_c = 1$), moderately difficult ($t_c = 4$), and difficult ($t_c = 8$) crossing-trajectory problems, where problem difficulty is quantified by the crossing-time duration t_c. As seen from table 1, for a given level of difficulty, the solutions, ranked by decreasing likelihood, are in the same rank order as was anticipated intuitively. The likelihood ratio shows that the crossing-track solution is significantly more likely than the switching-track solution. The likelihood ratio also shows that the dynamic range of the likelihoods of the two solutions decreases with increasing problem

44

difficulty. Interpreting solution likelihood as solution explanatory power, the decrease in dynamic range means that the local solutions provide less differentiated explanations of the data as the tracking problem becomes more difficult. In the limit as $t_c \rightarrow \infty$, all local solutions provide equally good explanations of the data.

Table 1. Local ML Tracking Solutions Ranked by Total Likelihood and Crossing-Track Duration, t_c

Solution	$\text{Log}_e\, P(Z,X)$ and Number of PMHT Iterations		
Track description	$t_c = 1$ (easy)	$t_c = 4$ (moderately difficult)	$t_c = 8$ (difficult)
Crossing tracks	469.20 3 iterations	327.54 29 iterations	258.22 34 iterations
Switching tracks	459.62 3 iterations	318.88 52 iterations	252.88 30 iterations
Likelihood ratio	1.45×10^4	5.77×10^3	2.09×10^2

The solutions listed in table 1 are not the only solutions to the problem. In particular, because the observer PDF is unchanged by interchanging the roles of targets 1 and 2, two other solutions are readily obtained. These alternative solutions, however, cannot be considered different from those listed in table 1. The interchangeability of certain observer PDF parameters is related to the more general issue of parameter "identifiability," as it is known in the general statistical literature. This statistical terminology is unrelated to the system identifiability terminology in the Kalman filtering literature.

10. CONCLUSIONS AND RECOMMENDATIONS

10.1 CONCLUSIONS

The study presented in this report is a probabilistic approach to the measurement-to-track assignment problem. Measurements were not assigned to tracks as in traditional multi-hypothesis tracking (MHT) algorithms; instead, the probability that each measurement belongs to each track was estimated using a maximum *a posteriori* (MAP) method. These measurement-to-track probability estimates are intrinsic to the multitarget tracker called the probabilistic multi-hypothesis tracking (PMHT) algorithm. The PMHT algorithm is computationally practical because it requires neither enumeration of measurement-to-track assignments nor pruning. The PMHT algorithm is an optimal MAP multitarget tracking algorithm.

10.2 TOPICS OF PRACTICAL IMPORTANCE

The following topics of practical importance related to the PMHT approach remain to be investigated.

1. The choice of the number of target models M is an important model-order selection problem that is of practical importance to a mature multitarget tracking algorithm. It is anticipated that PMHT will be robust to mismatch between the true number of targets and the number M assumed by PMHT, provided that M is at least as great as the true number of targets. The reason for suspecting such tolerance for PMHT is that "extra" target models can be used to model background clutter and noise levels.

2. Outputs from the previous batch can be used to initialize the PMHT algorithm for the current batch, but this procedure does not reduce computational complexity because it is not recursive. Development of a recursive PMHT algorithm that operates between successive measurement batches is an important subject for future work.

3. The robustness of PMHT algorithms against target coalescence is of particular importance. All tracking algorithms will coalesce sufficiently closely spaced targets. Because PMHT estimates are optimal empirical Bayesian estimates, it is anticipated that PMHT ability to resolve closely spaced targets will be at least as good as other currently available methods. (A closely related topic is track estimation bias.)

4. The sensitivity of the PMHT algorithm to target maneuvers must also be studied.

5. Bayesian priors for the assignments K can be used in the PMHT approach, if they are available. Such priors might be found by exploiting additional target information that is not resident in the sensor measurements Z; however, such priors have not been studied.

10.3 TOPICS OF THEORETICAL INTEREST

The following topics of theoretical interest should be studied further.

1. The sequence of error covariances generated by the PMHT algorithm is significant because multitarget FIMs are not readily definable using other approaches. (See section 6.)

2. Convergence of the PMHT algorithm is related to target observability. The multitarget observability condition for PMHT stated in section 6 essentially requires that all targets be individually observable. Conditions under which it may be possible to weaken this requirement would be very interesting. For instance, if the targets are not independent because they move in a coordinated formation, it may be possible to improve the effective target observability.

3. The relationship between the linear-Gaussian PMHT algorithm and Gaussian sum approximation methods (see references 14 and 15) for single-target tracking problems with nonlinear non-Gaussian state and measurement processes deserves investigation. In this case, PMHT approximates the exact nonlinear state distributions by a (convex) linear superposition of several linear-Gauss-Markov processes that are linked by assignment interference. The PMHT approach may result in efficient algorithms for these methods.

10.4 TOPICS OF BOTH THEORETICAL AND PRACTICAL INTEREST

The following topics of both theoretical and practical interest merit further study.

1. The adaptive estimation of target covariances (i.e., Q_{tm}) and the measurement covariances (i.e., R_{tm}) using the EM method need further study. The equations for adaptive ML covariance estimates are given in section 7. These adaptive equations are coupled with the PMHT target-state estimates because of assignment interference. Algorithms for solving these coupled equations efficiently remain to be investigated.

2. Non-EM methods for solving for the parameters of the marginal density (22) may be very useful in practice, but have not been considered in this report.

48

REFERENCES

1. H. Robbins, "An Empirical Bayes Approach to Statistics," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 1956, pp. 157-175.

2. R. L. Streit and T. E. Luginbuhl, "A Probabilistic Multi-Hypothesis Tracking Algorithm Without Enumeration and Pruning," *Proceedings of the Sixth Joint Data Fusion Symposium*, Laurel, MD, 14-18 June 1993.

3. R. L. Streit and T. E. Luginbuhl, "Maximum Likelihood Method for Probabilistic Multi-Hypothesis Tracking," *Proceedings of the SPIE International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, Orlando, FL, 1994.

4. A. Gelb, ed., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

5. B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979.

6. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm (with Discussion)," *Journal of the Royal Statistical Society*, series B, vol. 39, 1977, pp. 1-38.

7. L. R. Liporace, "Maximum Likelihood Estimation for Multivariate Observations of Markov Sources," *IEEE Transactions on Information Theory*, vol. IT-28, 1982, pp. 729-734.

8. G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.

9. G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977.

10. W. H. Press, B. P. Flannery, S.A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1986.

11. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, John Wiley & Sons, New York, 1968.

12. J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart, *LINPACK Users' Guide*, SIAM Journal for Applied Mathematics, Philadelphia, PA, 1979.

13. D. Avitzour, "A Maximum Likelihood Approach to Data Association," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-28, 1992, pp. 560-566.

14. H. W. Sorenson and D. L. Alspach, "Recursive Bayesian Estimation Using Gaussian Sums," *Automatica*, vol. 7, 1971, pp. 465-479.

15. D. L. Alspach and H. W. Sorenson, "Nonlinear Bayesian Estimation Using Gaussian Sum Approximations," *IEEE Transactions on Automatic Control*, vol. AC-17, August 1972, pp. 439-448.

16. S. S. Blackman, *Multiple Target Tracking with Radar Applications*, Artech House, Norwood, MA, 1986.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Chief of Naval Research (D. Johnson -- ONR-321SI)	3
Naval Sea Systems Command (PEO-SUB, ASTO, ASTO-E)	3
Navy Research Laboratory	1
Naval Postgraduate School	1
Naval Surface Warfare Center	1
Advanced Research Projects Agency	1
Defense Technical Information Center	12
Center for Naval Analyses	1