PMHT: Problems and Some Solutions

PETER WILLETT, Senior Member, IEEE
YANHUA RUAN
University of Connecticut
ROY STREIT
Naval Undersea Warfare Center

The probabilistic multihypothesis tracker (PMHT) is a target tracking algorithm of considerable theoretical elegance. In practice, its performance turns out to be at best similar to that of the probabilistic data association filter (PDAF); and since the implementation of the PDAF is less intense numerically the PMHT has been having a hard time finding acceptance. The PMHT's problems of nonadaptivity, narcissism, and over-hospitality to clutter are elicited in this work. The PMHT's main selling-point is its flexible and easily modifiable model, which we use to develop the "homothetic" PMHT; maneuver-based PMHTs, including those with separate and joint homothetic measurement models; a modified PMHT whose measurement/target association model is more similar to that of the PDAF; and PMHTs with eccentric and/or estimated measurement models.

Ideally, "bottom line" would be a version of the PMHT with clear advantages over existing trackers. If the goal is of an accurate (in terms of mean square error (MSE)) track, then there are a number of versions for which this is available. In terms of lost tracks there is no clearly preferable PMHT, although several variants (e.g. convergence-aided via measurement-variance inflation, maneuvering, and homothetic) offer performance similar to that of the PDAF; also, in very adverse tracking situations, the PMHT is preferable. We further hope that our demonstration of the facility by which a new model "idea" is turned into a working algorithm is encouraging to other researchers.

Manuscript received March 5, 1999; revised October 18, 2000 and January 17, 2002; released for publication February 15, 2002.

IEEE Log No. T-AES/38/3/06424.

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

This research was supported by the Office of Naval Research under Contracts N66604-98-M-3735 and N00014-00-1-0678.

Authors' addresses: P. Willett and Y. Ruan, Dept. of Electrical and Systems Engineering, University of Connecticut, Rm. 312 Eng. III, U157, 260 Glenbrook Rd., Storrs, CT 06269-2157, E-mail: (willett@engr.uconn.edu); R. Streit, Naval Undersea Warfare Center, Newport, RI 02841.

0018-9251/02/$17.00 © 2002 IEEE

I. INTRODUCTION

Both the multihypothesis tracker (MHT) [4] and probabilistic data association filter (PDAF) [2, 3] track imperfectly detected targets in clutter via a *hard-association* model. That is, these algorithms enumerate the possible associations between measurements and target(s), and evaluate which is best. And because there are a great many such associations a full enumeration is computationally infeasible, hence in each case the search is suboptimal.

The probabilistic multihypothesis tracker (PMHT) [21] makes modification to the measurement model. The PDAF and MHT assume, quite rightly, that a target can generate at most one measurement per scan; the PMHT sacrifices this constraint, and posits the measurement/target association process as independent across measurements. By doing so, it is able to render a *fully optimal* (under the modified assumption) tracker. The associations become *soft*—in fact, governed by their posterior probabilities—and the integer-programming problem of target tracking is rendered continuous and amenable to an iterative "hill-climbing" method via the EM algorithm. The PMHT has much in common with the EM algorithm as applied to the estimation of parameters in a Gaussian mixture [22].

The PMHT is probabilistically-sound (all the "bad" parts of the PMHT are out in the open in the original assumptions), and it is easily extensible. This extensibility has been exploited in a number of ways, such as those dealing with multiple targets [17, 9, 7], multiple sensors [14] nonlinear models [8], and target-maneuver [15, 19]. It is very likely that this easy extensibility will be the aspect to the PMHT which provides it a wider audience and acceptance. However, as yet the PMHT has not managed to "beat" the simple PDAF in the game of lost tracks. Therefore, we propose to exploit the PMHT's easy extensibility to attempt its improvement. We develop a series of model-varied PMHTs and test them; since our goal is straightforward, we test in the most straightforward manner possible, by simulation in a linear Gaussian environment, with one target, imperfect detection, and clutter.

In the following section we go into some detail about the development of the PMHT. We then discuss some issues to do with its practical implementation, such as the manner of dealing with false alarms, the correct choice of prior probabilities, batching, and initialization. In the succeeding section we discuss in some detail our perception of the PMHT's *problems*: its lack of adaptivity to a track being lost, its inability to sense an impending track loss, and its tendency to welcome several clutter measurements as a single detection of high accuracy. We then provide details on a variety of different modifications on the PMHT; there are twelve PMHT variations in all. We test them extensively, and conclude.

It will be seen, unfortunately, that no modification on the PMHT is uniformly superior to the simple PDAF.[1] However, in very adverse tracking environments several PMHTs are much improved, offering, for example 50% lost tracks versus the PDAF's 80%. Further, certain PMHT variants come close to the PDAF in lost-track performance, and are of considerable preferability in tracking accuracy for those tracks not lost.

II. PMHT

A. EM Algorithm

Consider the following:

$\mathcal{Z}$ an observation,

$\mathcal{X}$ an unknown random quantity,

$\mathcal{K}$ an unknown random quantity,

in which any of the above quantities may be continuous or discrete, scalar, or vector. It is desired to maximize $p(\mathcal{X} \mid \mathcal{Z})$ over $\mathcal{X}$; that is, the maximum *a posteriori* (MAP) estimate of $\mathcal{X}$ is sought. What makes the problem interesting is the appearance of the "nuisance" random quantity $\mathcal{K}$. Estimation of $\mathcal{K}$ is not required; however, in problems for which the EM approach is useful it is generally straightforward to write the probabilities $p(\mathcal{K} \mid \mathcal{X})$ and $p(\mathcal{Z} \mid \mathcal{X},\mathcal{K})$, but the desired probability is obtainable only via removal of conditioning, which makes direct maximization a difficult chore indeed.

In the EM approach, the quantity

$$Q(\mathcal{X}^{n+1};\mathcal{X}^{n}) \equiv \int_{\mathcal{K}} \log(p(\mathcal{X}^{n+1},\mathcal{K} \mid \mathcal{Z}))p(\mathcal{K} \mid \mathcal{X}^{n},\mathcal{Z}) \tag{1}$$

(since $\mathcal{K}$ may be discrete, we define integration in its most general sense) is iteratively maximized over $\mathcal{X}^{n+1}$, and n incremented. It is straightforward to show that (1) may be written as

$$Q(\mathcal{X}^{n+1};\mathcal{X}^{n}) \equiv \log(p(\mathcal{X}^{n+1} \mid \mathcal{Z})) + \int_{\mathcal{K}} \log(p(\mathcal{K} \mid \mathcal{X}^{n+1},\mathcal{Z}))p(\mathcal{K} \mid \mathcal{X}^{n},\mathcal{Z}). \tag{2}$$

Since it is well known (e.g. [5]) that for any probability measures p and $\hat{p}$ we have

$$\int \log(\hat{p})p \leq \int \log(p)p \tag{3}$$

with equality iff $p \equiv \hat{p}$, then it is clear that with $Q(\mathcal{X}^{n+1};\mathcal{X}^{n}) > Q(\mathcal{X}^{n};\mathcal{X}^{n})$ we must have $p(\mathcal{X}^{n+1} \mid \mathcal{Z}) > p(\mathcal{X}^{n} \mid \mathcal{Z})$. Convergence, at least to a local maximum, is assured [6] provided that the likelihood of interest is bounded.

[1]The PMHT could be compared to a number of other tracking approaches, such as MHT [4], SME [13], or assignment [3]. It is not the purpose here to contrast those algorithms, however, and we compare only to the PDAF, with the idea that the relative performance versus other approaches can be inferred from that.

B. Basic PMHT Algorithm Description

In a recent paper [21] a new approach to multitarget tracking and data association was proposed. The approach, called the PMHT, uses a recursive EM optimization method to compute in an optimal way the associations between measurements and targets. It shares some features with the results of [1], derived independently, but whereas in this latter paper a maximum likelihood estimator (MLE) is sought, in the former MAP estimation is the goal.

Under the (admittedly realistic) constraint that each target can generate at most one measurement at each time, multitarget tracking is inherently a combinatorial optimization problem; practical algorithms to solve it (the pruned MHTs and the joint PDAF (JPDAF) [2, 3]) are intelligent suboptimal procedures. With the constraint relaxed, however, an optimal procedure can be derived. The resulting PMHT uses "soft" posterior-probability associations between measurements and targets, and its implementation is a straightforward iterative application of the Kalman smoother operating on "synthetic" (i.e., modified) measurements. These soft associations can be considered as mapping the problem from discrete (i.e., of combinatorial complexity) to continuous (i.e., amenable to iterative algorithms).

In the scenario to which the PMHT is to be applied, there are assumed to be M targets, the sth of which moves according to the discrete-time linear model:

$$\begin{aligned} \mathbf{x}_s(t+1) &= \mathbf{F}_s(t)\mathbf{x}_s(t) + \mathbf{G}_s(t)\mathbf{u}_s(t) + \mathbf{v}_s(t) \\ \mathbf{y}_s(t) &= \mathbf{H}_s(t)\mathbf{x}_s(t) + \mathbf{w}_s(t) \end{aligned} \tag{4}$$

for $t = 1,2,\ldots,T$. Here, as usual, $\mathbf{x}_s(t)$ represents the trajectory of the sth model at time t, and $\mathbf{y}_s(t)$ its corresponding observation; the matrix sequences $\{\mathbf{F}_s(t),\mathbf{G}_s(t),\mathbf{H}_s(t)\}$ are known, and are assumed to represent an observable and controllable system. The random sequences $\{\mathbf{v}_s(t),\mathbf{w}_s(t)\}$ are assumed white, zero-mean, Gaussian, and mutually independent, with $E\{\mathbf{v}_s(t)\mathbf{v}_s^T(t)\} = \mathbf{Q}_s(t)$, $E\{\mathbf{w}_s(t)\mathbf{w}_s^T(t)\} = \mathbf{R}_s(t)$. The control sequences $\{\mathbf{u}_s(t)\}$ are known, and since these contribute in the form of "ownship" platform motion in a straightforward but notationally inconvenient way, we take them as zero with no loss of generality.

Now, naturally in a multitarget tracking scenario the thorniest problem is of *data-association*; that is, how to determine which measurements come from which targets. We pose that here in terms of EM algorithm parameters [21], in that for each t we define $\{k_r(t),\mathbf{z}_r(t)\}_{r=1}^{n_t}$ such that

$$\mathbf{z}_r(t) = \mathbf{y}_{k_r(t)}(t) \tag{5}$$

meaning that the rth measurement (out of n_t) at time t comes from model $k_r(t)$—and, of course, $k_r(t)$ is unknown.

In EM algorithm terms, then, we have

$\mathcal{X} = \{\mathbf{x}_s(t)\}$, where $\mathbf{x}_s(t)$ is the state of target s at time t,

$\mathcal{Z} = \{\mathbf{z}_r(t)\}$, where $\mathbf{z}_r(t)$ is the rth measurement vector at time t,

$\mathcal{K} = \{k_r(t)\}$, where $k_r(t)$ is the target from which the rth measurement at time t arises.

What remains is to provide a probabilistic structure for $\mathcal{K}$. In fact, our assumption is that

$$\Pr(k_r(t) = s) = \pi_s \tag{6}$$

and that all are independent random variables. It should be noted that the generally accepted probabilistic structure of each target having associated at most one measurement at each time is here abandoned. It is a perfectly feasible event that *all* measurements come from the same target.

With this, then, we have

$$\begin{aligned} p(\mathcal{Z},\mathcal{X},\mathcal{K}) &= \prod_{s=1}^{M} p(\mathbf{x}_s(1)) \prod_{t=2}^{T}\prod_{s=1}^{M} p(\mathbf{x}_s(t) \mid \mathbf{x}_s(t-1)) \\ &\quad \times \prod_{t=1}^{T}\prod_{r=1}^{n_t} \pi_{k_r(t)} \mathcal{N}\{\mathbf{z}_r(t); \hat{\mathbf{y}}_{k_r(t)}(t), \mathbf{R}_{k_r(t)}(t)\} \end{aligned} \tag{7}$$

where we have used

n_t the number of measurements at time t,

T the number of time samples in the batch,

M the number of targets,

$\mathcal{N}\{\chi;\boldsymbol{\mu},\boldsymbol{\Sigma}\}$ Gaussian density in variable χ, with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$,

$\hat{\mathbf{y}}_{k_r(t)}(t) \equiv \mathbf{H}_{k_r(t)}(t)\mathbf{x}_{k_r(t)}(t)$.

This is illustrated in Fig. 1. Equation (7), or rather its logarithm, specifies one element of the Q-function to be maximized. The other element is $p(\mathcal{K} \mid \mathcal{Z},\mathcal{X})$, for which we need $p(\mathcal{Z},\mathcal{X})$. This can be obtained by summation of (7) and appropriate gathering of terms as

$$\begin{aligned} p(\mathcal{Z};\mathcal{X}) &= \prod_{s=1}^{M} p(\mathbf{x}_s(1)) \prod_{t=2}^{T}\prod_{s=1}^{M} p(\mathbf{x}_s(t) \mid \mathbf{x}_s(t-1)) \\ &\quad \times \prod_{t=1}^{T}\prod_{r=1}^{n_t}\left[\sum_{l=1}^{M} \pi_l \mathcal{N}\{\mathbf{z}_r(t); \hat{\mathbf{y}}_l(t), \mathbf{R}_l(t)\}\right]. \end{aligned} \tag{8}$$

With this, we may factor

$$p(\mathcal{K} \mid \mathcal{Z},\mathcal{X}) = \frac{p(\mathcal{K},\mathcal{Z},\mathcal{X})}{p(\mathcal{Z},\mathcal{X})} = \prod_{t=1}^{T}\prod_{r=1}^{n_t} w^n_{k_r(t),r}(t) \tag{9}$$

where

$$w^n_{l,r}(t) = \frac{\pi_l \mathcal{N}\{\mathbf{z}_r(t); \hat{\mathbf{y}}^n_l(t), \mathbf{R}_l(t)\}}{\sum_{p=1}^{M}[\pi_p \mathcal{N}\{\mathbf{z}_r(t); \hat{\mathbf{y}}^n_p(t), \mathbf{R}_p(t)\}]}. \tag{10}$$

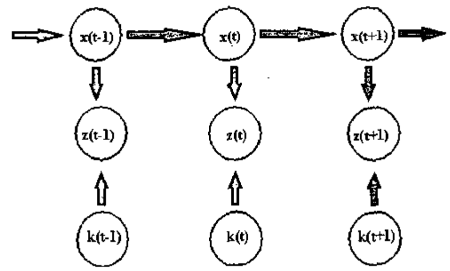

Fig. 1. The influence diagram of the basic PMHT.

From its definition in (10), $w^n_{k_r(t),r}(t)$ has the interpretation that it is the posterior probability (conditioned on $\mathcal{Z}$ and $\mathcal{X}$) that the rth measurement at time t arises from target $k_r(t)$. The superscript n amplifies the fact that quantities are computed from the nth EM iteration. In the EM approach, we have [21]

$$\begin{aligned} Q(\mathcal{X}^{n+1};\mathcal{X}^n) &= \sum_{\mathcal{K}} \log(p(\mathcal{X}^{n+1},\mathcal{K} \mid \mathcal{Z})) \prod_{t=1}^{T}\prod_{r=1}^{n_t} w^n_{k_r(t),r}(t) \\ &= \log\left(\prod_{s=1}^{M} p(\mathbf{x}^{n+1}_s(1)) \prod_{t=2}^{T}\prod_{s=1}^{M} p(\mathbf{x}^{n+1}_s(t) \mid \mathbf{x}^{n+1}_s(t-1))\right) \\ &\quad + \sum_{\mathcal{K},t,r} \log(\pi_{k_r(t)} \mathcal{N}\{\mathbf{z}_r(t); \hat{\mathbf{y}}^{n+1}_{k_r(t)}(t), \mathbf{R}_{k_r(t)}(t)\}) w^n_{k_r(t),r}(t). \end{aligned} \tag{11}$$

We notice from (11) that

$$\nabla_{\{\mathcal{X}^{n+1}\}} Q(\mathcal{X}^{n+1};\mathcal{X}^n) = \nabla_{\{\mathcal{X}^{n+1}\}} \hat{Q}(\mathcal{X}^{n+1};\mathcal{X}^n)$$

where

$$\begin{aligned} \hat{Q}(\mathcal{X}^{n+1};\mathcal{X}^n) &= \log\left(\prod_{s=1}^{M} p(\mathbf{x}^{n+1}_s(1)) \prod_{t=2}^{T}\prod_{s=1}^{M} p(\mathbf{x}^{n+1}_s(t) \mid \mathbf{x}^{n+1}_s(t-1))\right) \\ &\quad - \frac{1}{2}\sum_{s=1}^{M}\sum_{t=1}^{T}([\tilde{\mathbf{z}}_s(t) - \mathbf{H}_s(t)\mathbf{x}^{n+1}_s(t)]^T \tilde{\mathbf{R}}_s(t)^{-1} \\ &\qquad \times [\tilde{\mathbf{z}}_s(t) - \mathbf{H}_s(t)\mathbf{x}^{n+1}_s(t)]) \end{aligned} \tag{12}$$

where

$$\tilde{\mathbf{z}}_s(t) \equiv \frac{\sum_{r=1}^{n_t} w^n_{s,r}(t)\mathbf{z}_r(t)}{\sum_{r=1}^{n_t} w^n_{s,r}(t)} \tag{13}$$

$$\tilde{\mathbf{R}}_s(t) \equiv \frac{\mathbf{R}_s(t)}{\sum_{r=1}^{n_t} w^n_{s,r}(t)}. \tag{14}$$

Equation (12) is the logarithm of the joint likelihood function of M target models *for which there is no data association uncertainty* and for which measurements and corresponding measurement covariances are given by their *synthetic* values $\{\tilde{\mathbf{z}}\}$ and $\{\tilde{\mathbf{R}}\}$. As such, maximization of (12) is a (relatively) straightforward matter of applying the appropriate Kalman smoother (e.g. [2]). Thus, we have the algorithm:

1) calculate the ws based upon the measurements and the current track estimate

2) form the "synthetic" measurements and covariances from these ws

3) update the track using a Kalman smoother

to be iterated upon these three steps. The convergence of the PMHT is proven in [21]. It is also, perhaps, worth noting that the PMHT's computational load is to a first approximation linear in the number of targets and of measurements. This compares favorably to the complexity of most other target tracking approaches, and [20] suggests that any eventual acceptance of the PMHT may be on those terms.

C. PMHT Implementation Issues

Implementation of the PMHT algorithm just discussed appears straightforward; however, there are a number of issues which must first be dealt with at the planning stage. Note that these "issues" are not "problems" with the PMHT that require modification or fixing—those are discussed in the next subsection. Instead, these issues are more in the nature of design choices, and here we present our views.

1) *Initialization For Batch PMHT*: The PMHT as presented in the previous section is an inherently "batch" algorithm; that is, it is presented with T scans of data and finds the (locally) MAP track estimate for the entire T scans. From the algorithm description we note that the PMHT must be *initialized*[2] to a "guess" $\{\mathbf{x}_s(t)\}_{t=1}^{T}$. In general the PMHT (or any tracker) is provided with time-initial data in the form of knowledge that $\mathbf{x}_s(1)$ is Gaussian with specific mean $\bar{\mathbf{x}}_s(1)$ and covariance $\mathbf{P}_s(1)$. To form the track-initial value we simply use

$$\mathbf{x}_s^1(t) = \prod_{\tau=2}^{t} \mathbf{F}_s(\tau)\mathbf{x}_s(1) \qquad (15)$$

which is the MAP estimate of the track based upon the time-initial data and no measurements. There are other ideas, for instance in [10].

Now, with a small batch-length T (certainly $T < 10$), or with an unusually benign tracking environment (low clutter and high P_d), there is little problem in treating the batch in its entirety. However, with MAP track initialization as described above, errors in the important initial guess about the track accumulate with T. In many cases the PMHT will find an attractive false alarm near some point of its initialized track, and will never recover. In many cases, therefore, we prefer the approach of the following section.

2) *Sliding Batch*: One solution to a poor initialization is a *sliding*-batch method. The idea is relatively straightforward. Standard track-growth (by T_g) is used until the track reaches a length T_b, and from then estimation proceeds over a batch of T_b time samples which "slides" T_g samples after PMHT estimation on the current batch is finished. Since it is important not to use "future" (with respect to the beginning of the next batch) data to estimate the initial covariance (of the next batch), the final step for each batch should be a PMHT over the first $T_g + 1$ samples of the current batch only. In our simulations we use values $T_b = 6$ and $T_g = 3$; hence we give for reference:

$$1\;2\;3\;\rlap{$\displaystyle\underbrace{\underbrace{4\;5\;6\;7}_{b}\;8\;9}_{a}$}\;\overbrace{8\;9\;10\;11\;12}^{c}\;13\;14\;15$$

to show the transition between second and third batches. Here the numbers refer to the time sample number, a and c denote those samples from the second and third batches, respectively, and b shows those samples over which a PMHT should be run between the second and third batches in order to avoid contamination of the third batch's initial covariance estimated by "future" data.

3) *Prior Probabilities*: In the PMHT, we need to compute the prior probabilities $\{\pi_s(t)\}$, the (prior) probability that a measurement at time t comes from target s. The method we adopt is based on a valid statistical model.[3] That is, we fix the $\pi_s(t)$ a priori to the correct value conditioned on the number of measurements received in the current scan, and parametrized by the clutter density, the observation volume, and the probability of detection of the true measurement.

To be specific, suppose there are $n_t > 0$ measurements at scan t, the probability of detection of target s is $P_d(s)$, and the number of false alarms in the surveillance region has a Poisson distribution with mean λV (V is the "volume" of the surveillance region, and λ is the expected number of clutter returns per unit volume). Then if there is one target being tracked ($M = 1$) we have

$$\Pr(\text{a particular measurement is that from target } s = 1) \qquad (16)$$

$$= \frac{\dfrac{P_d(1)}{n_t}\dfrac{e^{-\lambda V}(\lambda V)^{n_t-1}}{(n_t-1)!}}{P_d(1)\dfrac{e^{-\lambda V}(\lambda V)^{n_t-1}}{(n_t-1)!} + (1-P_d(1))\dfrac{e^{-\lambda V}(\lambda V)^{n_t}}{n_t!}}$$

$$= \frac{P_d(1)}{n_t P_d(1) + (1-P_d(1))(\lambda V)}. \qquad (17)$$

[2] In most discussions of tracking *initialization* refers to the first point in a track estimate $\mathbf{x}_s(1)$. In discussion of the PMHT we refer to this as *time*-initialization; there is also *track*-initialization, corresponding to the first step in the algorithm description, to be considered.

[3] Since the original PMHT model allows *more* than one measurement per target in each scan, selection of the πs as follows is really an attempt to "match" this to the model that a target can generate *at most* one.

Particularly if V is large there is little lost by approximating (17) as $\pi_1(t) = P_d(1)/n_t$. Expressions such as (17) for $M > 1$ are straightforward to derive but are somewhat more messy.

4) *False Alarms*: In the PMHT as given in the previous section there is no mechanism for false alarms whatever: each measurement comes from a target being tracked. This is not an issue to be ignored, since if there are false alarms the unforewarned PMHT *will* treat each of these as a valid detection, and will synthesize a measurement $\tilde{\mathbf{z}}$ having little relevance to the true trajectory but possessing, via the denominator of (14), very high apparent accuracy.

The solution is simple: a spurious *false alarm* target model must be introduced, its role to absorb what are presumably clutter-generated detections far from the currently estimated trajectories through the posterior probability ws in (10), but whose trajectory need not be tracked. This false alarm "target" may be a uniform distribution over the surveillance region. To understand the consistency of this uniform model, note that the conditional prior probabilities of the previous section are (loosely) inversely proportional to the surveillance volume V; but as the value of the false alarm distribution is also inversely proportional to V, V has little effect.

5) *EM Convergence*: The PMHT is an iterative algorithm, and of concern is when to stop this iteration. One approach is to iterate a fixed number of times. If this is adopted, we have observed that good results are usually obtained if this number is between 5 and 10. However, there is often a great waste of computational resource since many later iterations engender little movement in the trajectory estimate. An adaptive scheme is sought, and there can be little argument that the most correct is based on an examination of the current likelihood function, directly computable from (8). However, the track likelihood is expensive to compute, and we prefer to monitor the difference between iterated track estimates. Specifically, we examine the measure

$$\frac{1}{T}\sum_{t=1}^{T}(\mathbf{x}_s^n(t)-\mathbf{x}_s^{n-1}(t))^T\mathbf{Q}^{-1}(\mathbf{x}_s^n(t)-\mathbf{x}_s^{n-1}(t)) \quad (18)$$

and cease iteration when this drops below a set level, for example 1%.

D. PMHT Problems

The weakness of the PMHT is in its *dependence on a good track-initialization*. There are a number of reasons why the PMHT is sensitive in this regard, and while all could be explained away as poor initialization, it is instructive to consider them separately that modifications can be made. In the following Section III a number of PMHT variants

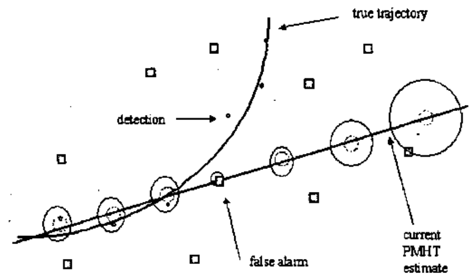

Fig. 2. Illustration of nonadaptivity of basic PMHT: true detections are shown as ○s, false alarms as □s, the PMHT's internal estimate of its state error covariance as the larger ellipses, and the measurement error covariances which control the ws are shown as smaller dashed ellipses. Note both that the PMHT's state error covariances have shape controlled by the measurement covariances (both are circular in this plot); and that while the state error covariances increase in size, the ws are sought over a fixed-size volume.

are presented, and these to some extent mitigate the effects of the problems.

1) *Nonadaptivity*: Before examining the PMHT, let us consider the behavior of the PDAF. The PDAF forms posterior probabilities of each measurement being associated with the current track (called the βs in [2]) by computing the probability of each detection under an assumption of Gaussianity, with mean the predicted value of the measurement, and with covariance that of the innovations, usually denoted $\mathbf{S}$. On an intuitive level, therefore, if the current track estimate from the PDAF is apparently poor, the resulting "large" $\mathbf{S}$ will force an adjustment in the posterior probabilities such that a wider field of view is encouraged, meaning that the PDAF looks further afield for a valid measurement. If this action occurs soon enough and there *is* a valid unmissed measurement nearby, then a "rescue" of the track is effected; if not, then $\mathbf{S}$ grows without bound and the track becomes lost.

Examination of (10) shows that each posterior probability of a measurement being track-generated is for the PMHT controlled by the measurement covariance $\mathbf{R}$, regardless of the quality of the current state estimate. There is no adaptivity whatever, and the PMHT is to a large extent incapable of rescuing itself from a currently poor track estimate. This is illustrated in Fig. 2.

A further aspect to this nonadaptivity is through the estimation covariance *shape*. Specifically, the spread of its validated measurements enters the PDAF update as its "spread of the innovations" term, and the internal PDAF state estimate covariance (the $\mathbf{nP}$) adjusts its shape to reflect it. On an intuitive level, if several scans of data all have validated measurements that are offset in the same direction from the a-priori track estimate, then the PDAF's $\mathbf{P}$ becomes elongated in that direction reflecting its true uncertainty. We contrast this with the behavior of the

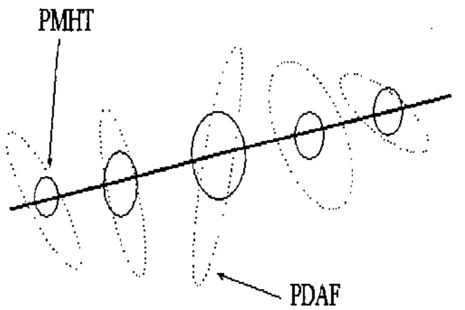

Fig. 3. Illustration of difference between PDAF and PMHT, in that the former's state estimation covariances are adjusted in shape by the data, while the latter's are adapted only in size, not shape.

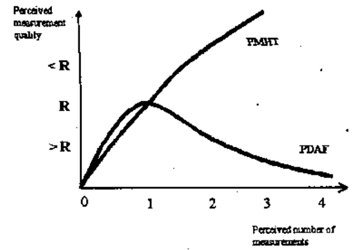

Fig. 4. Illustration of effect of multiple "valid" measurements on PDAF and PMHT. As this number increases beyond unity, the PMHT perceives a high-quality "multiple" measurement (i.e., $\tilde{\mathbf{R}} < \mathbf{R}$), a result which intuition suggests is incorrect.

PMHT, for which the implication of (14) is that the presence/position of measurements can adjust the *synthetic* measurement error covariance in a scalar way. That is, if we suppose that both the initial state estimation error covariance $\mathbf{P}_s(1)$, and all measurement error covariances $\mathbf{R}$ are spherical, then there is no mechanism at all for the PMHT to have other than spherical $\mathbf{P}$s. This is illustrated in Fig. 3.

2) *Narcissism*: Further examination of Fig. 2 illustrates another PMHT problem: even a single false alarm near an incorrect track estimate is satisfactory to the PMHT, and the lack of a reasonable number of validated detections does not faze it. This willingness to believe that its track is progressing normally in the face of overwhelming evidence to the contrary is certainly related to the "nonadaptivity" of the search for valid measurements discussed in the previous subsection.

3) *Hospitality*: In both the PDAF and PMHT the estimation covariance that would appear at scan t in the absence of any measurement at scan t is reduced by an amount corresponding to the perceived quality of the measurement(s) at scan t. In the case of the PMHT this amount is easy to quantify: it is the *synthetic* measurement covariance $\tilde{\mathbf{R}}(t)$. For the PDAF the situation is more complicated: but there is a reduction, weighted by the perceived probability that a true target-generated measurement is present, of the standard Kalman information factor $\mathbf{WSW}^T$; and a further *inflation* by a "spread-of-the-innovations" term whose value grows with the number of validated measurements.

If the posterior estimate of the number of valid measurements at time t is less than unity, it is easy to see from (14) that for the PMHT we have $\tilde{\mathbf{R}} > \mathbf{R}$, meaning that there may be a measurement, but it is of poor quality. There is a similar effect for the PDAF caused by the posterior valid-measurement probability weighting. If there is exactly one measurement near enough to its expected location to be called valid, then for the PMHT we have $\tilde{\mathbf{R}} \approx \mathbf{R}$; and the effect on the PDAF is similar to that on a Kalman filter.

Both of the above behaviors, for both algorithms, are as intuition suggests they should be. We now consider the case that there are several "valid" measurements. In the case of the PDAF the spread-of-the-innovations terms amounts to an admission of confusion by the data association step, and the perceived measurement quality is low. For the PMHT, however, the denominator in (14) can be greater than unity: the information perceived as of high quality, or $\tilde{\mathbf{R}} < \mathbf{R}$. This is natural from the PMHT model, which admits as feasible events zero, one, two, or even *all* measurements as having been target-generated, and the PMHT is hospitable enough to welcome extra measurements as providing extra accuracy. The situation is presented schematically in Fig. 4. Intuition suggests that the PDAF behavior is reasonable, while that of the PMHT is not, since the plurality of measurements hints that some must be clutter.

III. FIXING THE PMHT

The PMHT is an elegant approach to target tracking. However, the performance of the basic PMHT is often disappointing. The reasons for this, as we understand them, are given in the previous section: nonadaptivity of the "look" region to the present estimation uncertainty; unwillingness to admit that a detection-poor trajectory estimate is unlikely; and too great a hospitality to multiple measurements.

There are a number of approaches to fixing the PMHT. These may be grouped as follows.

Model-Modification: The basic PMHT model, as elaborated in the previous section, must be augmented or changed. Virtually all of the following "variety" of PMHTs are in this category, and an example is the *homothetic* PMHT.

Track-Rescue: When the PMHT, presumably via a low likelihood, senses a track that is in error, a reinitialization is commanded starting from a previous scan in which the PMHT confidence is high. While rescue approaches are useful for practical implementation, we do not pursue them in this study: comparison is murky, since a rescued PMHT must be compared with a rescued PDAF, and it is difficult to know how hard to "try" with a rescue.

Convergence-Aid: Due to the nature of its model's likelihood surface and the "hill-climbing" behavior of the EM algorithm, the PMHT often converges to a locally MAP estimate.[4] Two PMHT variants, the deflationary and homoscedastic, can be considered convergence aids.

For the most part we discuss PMHTs that use altered measurement/plant models. There is a somewhat fuzzy boundary between these and PMHTs with aided convergence, since the convergence aid usually relies on a modified measurement model. However, we draw the distinction that an aided-convergence PMHT uses an altered model only *temporarily*, for initial EM iterations.

In what follows we discuss the batch PMHT, using data from $t = 1,\ldots,T$, with the understanding that in each case the intention is to use a *sliding* batch. We also concentrate on the case that there is but $M = 1$ target and drop the target subscript s, allowing that extension to multiple targets is both straightforward and somewhat messier. Similarly, since allowing for a time-varying target model yields little except denser notation, we drop the argument t on $\mathbf{F}$, $\mathbf{H}$, $\mathbf{R}$, and $\mathbf{Q}$.

A. Homothetic PMHT

The homothetic PMHT, discussed previously in [17, 18], is a modification on the basic PMHT model such that measurements at scan t can come from any Gaussian density having mean $\mathbf{x}(t)$ and variance $\{\kappa_p^2\mathbf{R}\}_{p=1}^P$. Typical values used are $P = 2$, with $\kappa_1 = 1$ and $\kappa_2 = 3$. The "homothetic" nomenclature derives from "having the same mean"; that is, the model has been altered from having one target to P targets, and EM MAP estimation proceeds under the constraint that each of these models has the same track $\{\mathbf{x}(t)\}_{t=1}^T$.

Equation (8) is replaced by

$$p(\mathcal{Z},\mathcal{X}) = p(\mathbf{x}(1))\prod_{t=2}^{T} p(\mathbf{x}(t)\mid\mathbf{x}(t-1))\prod_{t=1}^{T}\prod_{r=1}^{n_t} \times\left[\frac{\pi_0}{V} + \sum_{p=1}^{P}\pi_p\mathcal{N}\{\mathbf{z}_r(t);\hat{\mathbf{y}}(t),\kappa_p^2\mathbf{R}(t)\}\right] \quad (19)$$

from which the obvious modification to (10) is

$$w_{p,r}(t) = \frac{\pi_p\mathcal{N}\{\mathbf{z}_r(t);\hat{\mathbf{y}}_p(t),\kappa_p^2\mathbf{R}(t)\}}{(\pi_0/V) + \sum_{l=1}^{P}[\pi_l\mathcal{N}\{\mathbf{z}_r(t);\hat{\mathbf{y}}_l(t),\kappa_l^2\mathbf{R}(t)\}]}. \quad (20)$$

[4]All practical trackers, including the PDAF, suffer from this: they lose track.

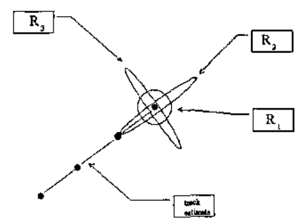

Fig. 5. Illustration of the eccentric measurement covariance ellipses assumed under the "Spirograph" homothetic PMHT model.

In accordance with our previous discussion, it is reasonable to predefine

$$\pi_p = \frac{P_d}{P[n_t P_d + (1-P_d)(\lambda V)]}$$

in practice. Some algebra yields that

$$\tilde{\mathbf{z}}(t) \equiv \frac{\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}(t)\mathbf{z}_r(t)/\kappa_p^2}{\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}(t)/\kappa_p^2} \quad (22)$$

$$\tilde{\mathbf{R}}(t) \equiv \frac{\mathbf{R}(t)}{\sum_{r=1}^{n_t}\sum_{l=1}^{P} w_{l,r}(t)/\kappa_l^2} \quad (23)$$

are the new *synthetic* measurements and associated covariances, to replace (13) and (14)—the PMHT iteration is otherwise intact.

B. Spirograph PMHT

Essentially this is a more-involved homothetic PMHT. Whereas the homothetic PMHT assumes that measurements at scan t can come from any Gaussian density having mean $\mathbf{x}(t)$ and variance $\{\kappa_p^2\mathbf{R}\}_{p=1}^P$, the variances in the case of the spirograph PMHT can be any of $\{\mathbf{R}_p\}_{p=1}^P$. As illustrated in Fig. 5 we use $P = 3$, with $\mathbf{R}_1 = \mathbf{R}$, and the other two eccentric ellipses with, respectively, major and minor axes lying in the direction of (currently estimated) PMHT motion. That is, we have

$$\mathbf{R}_2 = \frac{\sigma_m^2}{\sqrt{\xi}}\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \xi\end{pmatrix}\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix}$$

$$\mathbf{R}_3 = \frac{\sigma_m^2}{\sqrt{\xi}}\begin{pmatrix}\sin(\theta) & \cos(\theta)\\ -\cos(\theta) & \sin(\theta)\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \xi\end{pmatrix}\begin{pmatrix}\sin(\theta) & -\cos(\theta)\\ \cos(\theta) & \sin(\theta)\end{pmatrix} \quad (24)$$

in which $\theta \equiv \tan^{-1}(v_y/v_x)$, with v_x and v_y the (PMHT estimated) velocity in x and y coordinates. The intention is that the PMHT can be either sped up or slowed down appropriately ($\mathbf{R}_2$); or it can be moved to the side ($\mathbf{R}_3$); our view is that such errors account for many PMHT lost tracks. Note that to avoid too much *hospitality*, the volume of each covariance ellipse is the same.

The modification to the PMHT update is similar to, but slightly more complicated than, that of the homothetic PMHT, but with (22) and (23) replaced by

$$\tilde{\mathbf{z}}(t) \equiv \left(\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}^{n}(t)\mathbf{R}_p^{-1}\right)^{-1}\left(\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}^{n}(t)\mathbf{R}_p^{-1}\mathbf{z}_r(t)\right) \tag{25}$$

$$\tilde{\mathbf{R}}(t) \equiv \left(\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}^{n}(t)\mathbf{R}_p^{-1}\right)^{-1}. \tag{26}$$

One advantage to the use of the spirograph model is that it has the protection to overcome both aspects of the basic PMHTs *nonadaptivity*. Certainly there is encouragement to look further afield for valid measurements; but there is also a means for the PMHT to develop a state estimation covariance ellipse *shape* that is adaptive to the measurements obtained.

C. Adaptive Homothetic PMHT

In this case we have $P = 2$, with the first ($p = 1$) homothetic model the standard ($\mathbf{R}$) one, and the second ($p = 2$) a measurement noise model that is estimated from the data in the batch. That is, in EM terms, we insert $\mathbf{R}_2$ to the *desired* variables $\mathcal{X}$. After each Kalman smoother operation we therefore reestimate

$$\mathbf{R}_2^{n+1} = \frac{\sum_{t=1}^{T}\sum_{r=1}^{n_t} w_{2,r}(t)[\mathbf{z}_r(t) - \mathbf{H}\mathbf{x}^{n+1}(t)][\mathbf{z}_r(t) - \mathbf{H}\mathbf{x}^{n+1}(t)]^T}{\sum_{t=1}^{T}\sum_{r=1}^{n_t} w_{2,r}(t)}. \tag{27}$$

In practice we initialize $\mathbf{R}_2^1 = 2\mathbf{R}$. It is also necessary to avoid a catastrophic deflation of $\det\{\mathbf{R}_2^n\} \to 0$ —usually because only one $w_{2,r}(t)$ is significant, and thus $\mathbf{R}_2$ estimated therefrom accelerates to singular—hence $\mathbf{R}_2$ is estimated as $\mathbf{R}_2 = \mathbf{R}_1 + \breve{\mathbf{R}}$ for some positive-definite $\breve{\mathbf{R}}$.

If the PMHT's current trajectory is diverging from the true one, it is reasonable to expect there to be an agglomeration of detections offset from that current trajectory estimate in a consistent direction. The appeal of the adaptive homothetic PMHT is that this direction (ideally) should appear in $\mathbf{R}_2$, and measurements in that direction explicitly sought. This is sketched in Fig. 6. In practice, our results have indicated little difference between this PMHT and the previous spirograph model.

D. PMHT with PDAF Measurement Model

There is a legitimate complaint that the PMHT measurement model, in which measurement/track associations are independent across all measurements, is unrealistic. The events that zero, one, two, or even *all* measurements be target-generated are all feasible and their probabilities evaluated. The PDAF

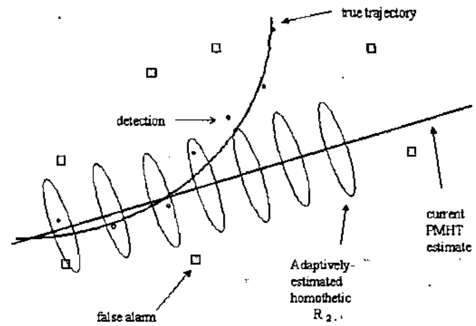

Fig. 6. Illustration of the eccentric measurement covariance ellipses encouraged by a track estimate that is diverging from the true trajectory, under the adaptive homothetic model.

(and MHT) measurement models do not permit this; instead, given that there are n_t measurements at scan t, the feasible events $\{k_r(t)\}_{r=0}^{n_t}$ are

$$k_r(t) = \begin{cases} \text{all measurements are false alarms} & r = 0 \\ \text{only measurement } r \text{ is a true detection} & 1 \le r \le n_t. \end{cases} \tag{28}$$

(For the basic one-target PMHT there are 2^{n_t} feasible events.) From this, and assuming a Poisson-distributed number of false alarms, we have

$$w_r(t) = \begin{cases} \dfrac{(1-P_d)\lambda}{(1-P_d)\lambda + P_d\sum_{l=1}^{n_t}\mathcal{N}\{\mathbf{z}_l(t);\mathbf{H}\hat{\mathbf{x}}(t),\mathbf{R}(t)\}} & r = 0 \\ \dfrac{P_d\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}\hat{\mathbf{x}}(t),\mathbf{R}(t)\}}{(1-P_d)\lambda + P_d\sum_{l=1}^{n_t}\mathcal{N}\{\mathbf{z}_l(t);\mathbf{H}\hat{\mathbf{x}}(t),\mathbf{R}(t)\}} & 1 \le r \le n_t. \end{cases} \tag{29}$$

The PMHT iteration is thus modified only in that the *w*s are calculated according to (29) rather than (10). In our view, there is not a great deal of difference between this and the basic PMHT model, although we admit that a number of researchers, ourselves included, have tried it or some variation [18, 15, 16], and have found positive features. One advantage, of course, is that the problem of too much *hospitality* is entirely avoided, since the *w*s sum to unity. A disadvantage is that in the multitarget situation the evaluation of the events (28) becomes combinatorially-hard.

E. Detection-Oriented PMHT

The basic PMHT's *narcissism* discussed in Section IID2 stems at least partly from its inability to realize that a string of missed detections is inappropriate. We therefore consider modifying the PMHT such that there must be at least one detection in P scans of data—or at least by making the event that all P scans of data contain nothing but false alarms infeasible. We have, therefore, that

$$p(\{\{\mathbf{z}_r(t)\}_{r=1}^{n_t}\}_{t=t_0}^{t_1} \mid \mathcal{X}) = c\left[\prod_{t=t_0}^{t_1}\prod_{r=1}^{n_t}\left(\pi_0\frac{1}{V} + \pi_1\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}\hat{\mathbf{x}}(t),\mathbf{R}(t)\}\right) - \left(\frac{\pi_0}{V}\right)^{\sum_{t=t_0}^{t_1} n_t}\right] \tag{30}$$

in which c is a normalizing constant. Naturally introduction of an infeasible association creates dependency between the $2^{\sum_{t=t_0}^{t_1} n_t}$ measurement events, as does the PDAF measurement-model above. However, computation is quite straightforward: we first calculate the $2(\sum_{t=t_0}^{t_1} n_t)$ probabilities

$$\tilde{w}_{s,r}(t) = \begin{cases} \pi_0\frac{1}{V} & s = 0 \\ \pi_1\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}\hat{\mathbf{x}}(t),\mathbf{R}\} & s = 1 \end{cases} \tag{31}$$

and then form the actual posteriors

$$w_{1,r}(t) = \frac{\tilde{w}_{1,r}(t)}{[\tilde{w}_{0,r}(t) + \tilde{w}_{1,r}(t)]\left[1 - \prod_{t=t_0}^{t_1}\prod_{r=1}^{n_t}\frac{\tilde{w}_{0,r}(t)}{\tilde{w}_{0,r}(t) + \tilde{w}_{1,r}(t)}\right]} \tag{32}$$

and $w_{0,r}(t) = 1 - w_{1,r}(t)$. As usual $w_{0,r}(t)$ is the posterior probability that measurement r at time t is clutter, and $w_{1,r}(t)$ the posterior probability that it is target-generated. It would appear that extensibility to the multiple-target situation is straightforward.

The question remains as to how to choose $P = t_1 - t_0 + 1$. In our simulations we have tried both the value $P = 1$ and the value $P = 3$. In the latter case the model is essentially that no more than two missed detections in a row is acceptable; performance is reasonable here, and we report only those results.

F. Maneuvering PMHT

It turns out that the PMHT model easily incorporates an underlying Markov maneuver process—naturally, this appears as part of the augmented "nuisance" EM variables $\mathcal{K}$ into $\{k^1\}$, the original measurement/target association variables, and $\{k^2\}$, where $k^2(t) = l$ means that the (Markov) maneuver model is in state l at time t. Please see Fig. 7.

Thus, we have (for a single target)

$$p(\mathcal{Z},\mathcal{X},\mathcal{K}) = P(\mathbf{x}(1) \mid k^2(1))\prod_{t=2}^{T}\Bigg[p(\mathbf{x}(t) \mid \mathbf{x}(t-1),k^2(t))p(k^2(t) \mid k^2(t-1)) \times\prod_{r=1}^{n_t}\left(\left\{\begin{matrix} \pi_0\frac{1}{V} & k_r^1(t) = 0 \\ \pi_{k_r^1(t)}\mathcal{N}(\mathbf{z}_r(t);\mathbf{H}\mathbf{x}_{k_r^1(t)}(t),\mathbf{R}) & k_r^1(t) = 1 \end{matrix}\right\}\right)\Bigg]. \tag{33}$$

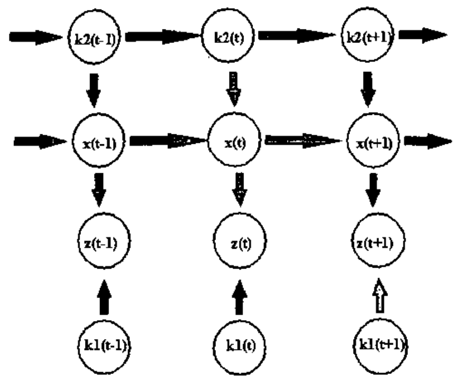

Fig. 7. Influence diagram of PMHT with maneuver. Difference between this and earlier influence diagram of basic PMHT in Fig. 1 is augmentation of "hidden" (nuisance) association variables $\{k^1\}$ by Markov maneuver process $\{k^2\}$.

Clearly, the difference between this and the original PMHT is the incorporation of the Markov model-transition process.

We denote $p(k^2(t) = l \mid \mathcal{Z},X) \equiv v_l(t)$, then we have

$$\begin{aligned} Q = &\sum_{l=1}^{M}[\log(p(\mathbf{x}(1) \mid k^2(1) = l))]v_l(1) \\ &+ \sum_{t=2}^{T}\sum_{l=1}^{M}[\log(p(\mathbf{x}(t) \mid k^2(t) = l,\mathbf{x}(t-1)))]v_l(t) \\ &+ \sum_{t=2}^{T}\sum_{r=1}^{n_t}[\log(p(\mathbf{z}_r(t) \mid \mathbf{x}(t),k_r^1(t) = 1))]w_r(t) \end{aligned} \tag{34}$$

where M denotes number of maneuver models.

In [15] the maneuver process controls which of a set of control inputs $\mathbf{u}(t)$ is active, and these presumably correspond to constant-velocity motion and a variety of turns and straight-line accelerations, a clever idea.[5] In [19] the models differ in their process-noise covariance matrices $\{\mathbf{Q}_m\}_{m=1}^{M}$. In [19] it was observed that the maneuver tracker was an improvement over the basic PMHT even when no maneuver was present—presumably this relates to a reduced degree of *narcissism.*

At any rate, the implementation of the maneuvering PMHT is somewhat more cumbersome than that of other models in that, as in the Baum–Welch re-estimation procedure, a forward-backward probability recursion must be performed in each EM iteration to recover the posterior probabilities (labeled v) of the Markov maneuver process. We define

$$\begin{aligned} \alpha_i(t) &\equiv p(\mathcal{X}_1^t, k^2(t) = i) \\ \beta_i(t+1) &\equiv p(\mathcal{X}_{t+1}^T \mid k^2(t) = i,\mathbf{x}(t)) \end{aligned} \tag{35}$$

[5]In this paper the PDAF measurement model was also used, as opposed to that of the basic PMHT.

and hence obtain the forward-backward recursion

$$\alpha_j(t+1) = \left[\sum_{i=1}^{M} A(i \to j)\alpha_i(t)\right] \times \frac{e^{-1/2[\mathbf{x}(t+1)-\mathbf{F}\mathbf{x}(t)]^T \mathbf{Q}_j^{-1}[\mathbf{x}(t+1)-\mathbf{F}\mathbf{x}(t)]}}{\sqrt{|2\pi \mathbf{Q}_j|}} \tag{36}$$

and

$$\beta_j(t+1) = \sum_{i=1}^{M} A(j \to i)\beta_i(t+2) \times \frac{e^{-1/2[\mathbf{x}(t+1)-\mathbf{F}\mathbf{x}(t)]^T \mathbf{Q}_i^{-1}[\mathbf{x}(t+1)-\mathbf{F}\mathbf{x}(t)]}}{\sqrt{|2\pi \mathbf{Q}_i|}}. \tag{37}$$

In (36) and (37) we have used $A(j \to i)$ to denote the transition probability matrix of the Markov maneuver process. We thus obtain

$$v_l(t) = \frac{\alpha_l(t)\beta_l(t+1)}{\sum_{m=1}^{M} \alpha_m(t)\beta_m(t+1)}. \tag{38}$$

Assuming that $k^2(t) = l$ implies that the process-noise covariance active at scan t is $\mathbf{Q}_l = \sigma_l^2 \mathbf{Q}_0$, we get

$$\tilde{\mathbf{Q}} = \frac{\mathbf{Q}_0}{\left[\sum_{l=1}^{M} (v_l(t)/\sigma_l^2)\right]}. \tag{39}$$

This $\tilde{\mathbf{Q}}$ is a *synthetic* process-noise covariance, exactly analogous to the synthetic measurement noise covariance of the basic PMHT, and it is inserted directly to the Kalman smoother in place of the fixed $\mathbf{Q}$. (For details of the above derivation, please see [19].)

G. Maneuvering/Homothetic PMHT

The maneuvering PMHT is less *narcissistic* than the basic PMHT, but problems with its lack of *adaptivity* remain. This is ameliorated somewhat by the addition of an *independent* homothetic measurement model. With reference to the influence diagram of Fig. 7 the association variables $\{k_r^1(t)\}$ in the original single-target maneuver model took on two values, $\in \{0,1\}$, corresponding to a measurement having arisen from clutter or being target-generated. With the augmentation of this to a homothetic measurement model, we now have $k_r^1(t) \in \{0,1,\ldots,P\}$. Since the two sets of "nuisance" variables $\{k_r^1(t)\}$ and $\{k^2(t)\}$ are insulated by $\{\mathbf{x}(t)\}$, and since the E-step in the PMHT is performed *conditioned* on the current track estimate $\{\mathbf{x}(t)\}_{t=1}^{T}$, there is no interaction. Thus, the maneuvering/homothetic PMHT is implemented as a simple combination of the two techniques in subsections IIIA and IIIF.

H. Joint Maneuver/Homothetic PMHT

In the previous variation on the PMHT the homothetic measurement model was combined with target maneuver. It was noted that the split of the "nuisance" variables $\mathcal{K}$ into $\{k_r^1(t)\}$, controlling measurement/target associations, and $\{k^2(t)\}$, the Markov maneuver process, was easily accomplished since these were *separated* probabilistically by the trajectory $\{\mathbf{x}(t)\}$. A variation, therefore, is to dissolve this separation: the same process is to control both maneuver and the homothetic aspects to the associations. More specifically, we have

$$\{k_r^1(t) = 0\} \Longleftrightarrow \{\text{measurement } r \text{ is a false alarm}\}$$
$$\{k_r^1(t) = 1\} \Longleftrightarrow \{\text{measurement } r \text{ is target-generated}\} \tag{40}$$

as in the basic PMHT; and

$$\{k^2(t) = p\} \Longleftrightarrow \{\text{all measurements at scan have } t \text{ measurement covariance } \mathbf{R}_p\}$$
$$\text{AND}$$
$$\{\text{the process noise between scans } t-1 \text{ and } t \text{ has covariance } \mathbf{Q}_p\}. \tag{41}$$

More detail of this variation can be found in [19] (in which this was suggested as a possible extension, but had not been tried). It appears to us that the both the basic maneuvering PMHT and the PMHT with separate maneuver/homothetic measurement models, while defying the basic PMHT's *nonadaptivity*, are still to some extent *narcissistic*. More specifically, the maneuver process seldom recognizes that a maneuver is present (through its $\{v_l(t)\}$s) unless a large deviation between $\mathbf{F}\mathbf{x}^n(t)$ and $\mathbf{x}^n(t+1)$ is forced upon it; and it is difficult (although not impossible) for the measurements to gather sufficiently convincing information to force such a large deviation since they are calculated assuming $\{\mathbf{x}^n(t)\}$. Through use of the joint model it is observed that an ongoing loss of track produces both an excitation of a homothetic model with high measurement covariance and a willingness to change direction via maneuver.

I. Deflationary PMHT

It has been stated previously that the major problem with the basic PMHT is of *track-initialization*; that is, the likelihood surface over which the basic PMHT performs its maximization is often so "busy" with local maxima that convergence to a good trajectory is unlikely. Thus, a means to smoothen the likelihood surface is sought. An approach which has been commonly applied (e.g. [12]) is to inflate the measurement covariance. It is hoped that if the initial EM iteration is performed with a large $\mathbf{R}$, and subsequent iterations use smaller and smaller (*deflating*) **R**s, then the eventual convergence to the global maximum will be encouraged. The idea is similar to that of simulated annealing, although "downhill" steps are not used.

There are many ways to select the inflated **R**s and the schedule of their decrease to the eventual "correct" measurement covariances. We have adopted the approach:

$$\mathbf{R}^n(t) = (1 - n/N)\mathbf{S}^n(t) + n/N\mathbf{R} = \mathbf{R} + (1 - n/N)\mathbf{H}\mathbf{P}^n(t)\mathbf{H}^T \quad (42)$$

in which n is the EM iteration number, N is the (fixed) total number of EM iterations, and $\mathbf{S}^n(t)$ is the innovations covariance at scan t and EM iteration n. We have selected this dependence on **S** as an attempt to recover some of the PDAF's *adaptivity*; but it should be noted that this is a *convergence-aided* PMHT. To avoid the problem of over-*hospitality*, we have forced the denominator of in the computation of $\tilde{\mathbf{R}}$ to be no larger than unity (see (14)) while an expanded measurement covariance is being used.

J. Homoscedastic PMHT

Whereas "homothetic" refers to measurement noise models having the same mean and different covariances, "homoscedastic" corresponds to identical covariances and different means. If a set of P offsets $\{\delta\mathbf{x}_p(t)\}_{p=1}^{P}$ is specified, then (19) (for the homothetic case) can be revised to

$$p(\mathcal{Z},\mathcal{X}) = p(\mathbf{x}(1))\prod_{t=2}^{T} p(\mathbf{x}(t) \mid \mathbf{x}(t-1)) \times \prod_{t=1}^{T}\prod_{r=1}^{n_t}\left[\frac{\pi_0}{V} + \sum_{p=1}^{P}\pi_p\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}[\mathbf{x}(t)+\delta\mathbf{x}_p(t)],\mathbf{R}(t)\}\right]. \quad (43)$$

This means that measurements can come from a variety of points offset from the true trajectory, and the intention is that an increased "look" region can be so obtained. To retain some degree of faith to the original model, we take $\delta\mathbf{x}_1(t) = 0$.

It is straightforward to modify the PMHT for the above model. Calculation of the ws proceeds in the usual way, with

$$w_{p,r}(t) = \frac{\pi_p\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}[\mathbf{x}(t)+\delta\mathbf{x}_p(t)],\mathbf{R}(t)\}}{(\pi_0/V) + \sum_{l=1}^{P}[\pi_l\mathcal{N}\{\mathbf{z}_r(t);\mathbf{H}[\mathbf{x}(t)+\delta\mathbf{x}_l(t)],\mathbf{R}(t)\}]}. \quad (44)$$

Reinsertion of these to the PMHT is accomplished via

$$\tilde{\mathbf{z}}(t) = \frac{\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}(t)[\mathbf{z}_r(t) - \mathbf{H}\delta\mathbf{x}_p(t)]}{\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}(t)}$$
$$\tilde{\mathbf{R}} = \frac{\mathbf{R}}{\sum_{r=1}^{n_t}\sum_{p=1}^{P} w_{p,r}(t)}. \quad (45)$$

Selection of the $\{\delta\mathbf{x}(t)\}_{p=2}^{P}$ is a matter of choice: these can be fixed a priori, presumably into a simplex pattern; or they can be chosen randomly, but remain constant with t. It appears that random scattering produces a more catholic look region, and that the greater randomness of the second option is the most favorable. We use, for $p = 2,\ldots,P$,

$$\delta\mathbf{x}_p(t) = \sqrt{\mathbf{P}(t)}\zeta_p(t) \quad (46)$$

in which $\{\{\zeta_p(t)\}_{t=1}^{T}\}_{p=2}^{P}$ are independent unit-normal random vectors.

While the homoscedastic PMHT is, from a mathematical point of view, a valid model, it does not make a great deal of physical sense. We therefore use it as a *convergence aid*, with the final EM iteration under the basic PMHT.

IV. EXPERIMENT

A. Model and Parameter Values

For our simulations we choose a two-dimensionally kinematic model with direct discrete-time process noise, that is, with reference to the kinematic model (4), we have

$$\mathbf{F} = \begin{Bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{Bmatrix}$$

$$\mathbf{Q} = \sigma_p^2 \begin{Bmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} & 0 & 0 \\ \frac{\Delta t^3}{2} & \Delta t^2 & 0 & 0 \\ 0 & 0 & \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} \\ 0 & 0 & \frac{\Delta t^3}{2} & \Delta t^2 \end{Bmatrix}$$

$$\mathbf{H} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{Bmatrix}$$

$$\mathbf{R} = \sigma_m^2 \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix}.$$

Measurements are of position-only, and the model is in all respects linear. Detections may be missed, independently from scan to scan with probability $1 - P_d$, and there are at each scan false alarms whose number is Poisson distributed according to clutter density λ (in events per meter^2).

Some notes and parameter follow.

1) We have chosen $\Delta t = 30$ s, a fast but not-unreasonable scan rate for active sonar.

2) For the most part we have $\sigma_m = 100$ m, corresponding to a constant-frequency pulse with

length of the order of 1 s. We also examine the case $\sigma_m = 10$ m.

3) We explore varying process noises: $\sigma_p = 0.0001$ meter/s^2, $\sigma_p = 0.001$ meter/s^2, and $\sigma_p = 0.01$ meter/s^2. According to [2], we thus have

$$\lambda_{\text{man}} = \sigma_p(\Delta t)^2/\sigma_m \in \{0.09, 0.009, 0.0009\} \quad (47)$$

as the *maneuvering index* of the target.

4) We explore various probabilities of detection: $P_d \in \{50\%, 70\%, 90\%\}$.

5) We explore various clutter densities, $\lambda \in \{10^{-5.5}, 10^{-6}, 10^{-6.5}, 10^{-7}\}$. If we imagine a circular PMHT "gate" of radius $4\sigma_m$, then these values imply that such a gate would contain, on average, 1.6, 0.5, 0.16, and 0.05 false alarms per scan under the respective λs. Thus these clutter values range from reasonably heavy to very light.

6) We choose a track length $T = 30$ scans of data.

For the most part, targets begin their trajectories at scan $t = 0$ (explained shortly) with position coordinates (0,0) and velocity coordinates (5,5) m/s, corresponding to 13.8 knots. We also examine the cases in which the initial velocities are (1,1) and (20,20) m/s.

All simulations are on the basis of 200 Monte Carlo simulation runs. The exception to this is the PDAF, for which 500 runs are used. We adopt the approach, common to many tracking studies, of *two-point* time-initialization [2] and we declare a track lost if

$$(\mathbf{x}_1^n(T) - \mathbf{x}_1^{\text{true}}(T))^2 + (\mathbf{x}_3^n(T) - \mathbf{x}_3^{\text{true}}(T))^2 > 2(4\sigma_m)^2 \quad (48)$$

in which the subscripts 1 and 3 refer to the two position coordinates. We have usually found that the precise value of the threshold is not important over a fairly wide range: if a track is lost, this value is orders of magnitude too high.

B. Results

We test the various PMHTs described earlier, and also the PDAF, in several situations. These are:

$\lambda = 10^{-5.5}$, initial speed 13.8 knots, $\sigma_m = 100$ (Fig. 8 with $P_d = 90\%$ and $\sigma_p = 0.0001$);

$\lambda = 10^{-6}$, initial speed 13.8 knots, $\sigma_m = 100$;

$\lambda = 10^{-6.5}$, initial speed 13.8 knots, $\sigma_m = 100$ (Fig. 9 with $P_d = 90\%$ and $\sigma_p = 0.01$);

$\lambda = 10^{-7}$, initial speed 13.8 knots, $\sigma_m = 100$;

$\lambda = 10^{-6}$, initial speed 2.8 knots, $\sigma_m = 100$;

$\lambda = 10^{-6}$, initial speed 55.2 knots, $\sigma_m = 100$;

$\lambda = 10^{-5}$, initial speed 13.8 knots, $\sigma_m = 10$.

In each of the above situations we use $P_d \in \{50\%, 70\%, 90\%\}$, and $\sigma_p \in \{0.01, 0.001, 0.0001\}$; that is, there are 7 situations above, and each of these is explored for 9 subcases. The fifth and sixth situations are included to show the effect of the relative angle of the time-initialized track. In the fifth slow-moving case, the heading error is large compared with the speed error; and vice-versa for the sixth fast-moving target. It was anticipated that the PMHT would "prefer" the fast-moving target; little evidence of this was found. The seventh situation corresponds to a shorter sonar pulse length, or to pulse-compression, in that the measurement error is comparatively small.

We list only selected results here, from the first, third, and seventh situations above; more extensive tables are given in [23]. The results in terms of in-track percentage are given in respective Tables II, III, and IV. Further results in terms of root-mean-square error (RMSE) for nonlost tracks are given in Table V. We attempt to digest our results in Table I, which lists the average rank (rank 1 is best, rank 13 is worst) for each of the thirteen algorithms tested, in each of the 63 situations in which the tests were run, in terms of lost tracks and of RMSE for nonlost tracks.

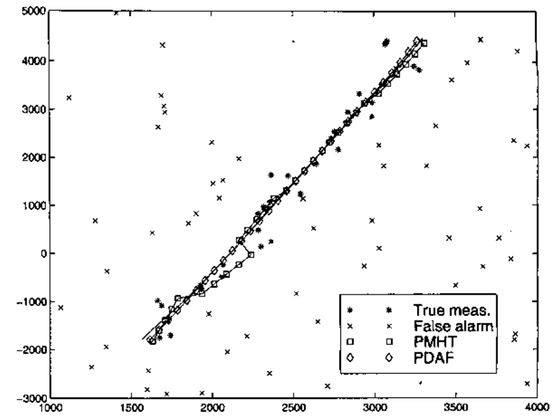

Fig. 8. The true, PMHT, and PDAF tracks with $\lambda = 10^{-5.5}$, $P_d = 90\%$, $\sigma_p = .0001$, $\sigma_m = 100$, and initial speed 13.8 knots. The *final scan only* of clutter (i.e. $t = T$) is shown, with clutter returns denoted by $\times$.

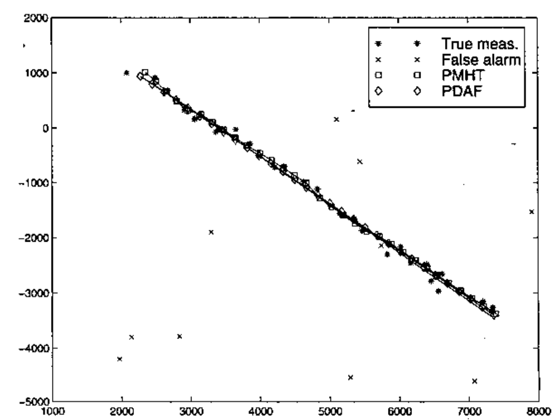

Fig. 9. The true, PMHT and PDAF tracks with $\lambda = 10^{-6.5}$, $P_d = 90\%$, $\sigma_p = .01$, $\sigma_m = 100$, and initial speed 13.8 knots. The *final scan only* of clutter (i.e. $t = T$) is shown, with clutter returns denoted by $\times$.

TABLE I
Average Rank of Each of the 13 Algorithms in Each of the 63 Situations Tested

	In-Track	RMSE
PDAF	3.9	6.1
Deflat. PMHT	4.5	3.5
Man. PMHT (hom)	4.5	8.4
Homothetic PMHT	4.9	11.3
Homosc. PMHT	5.0	9.4
D.O. PMHT ($P = 3$)	6.6	6.6
Maneuver PMHT	7.0	5.0
PMHT (sliding)	7.4	5.4
Man. PMHT (jnt)	7.4	10.8
PMHT (PDAF meas)	9.0	3.4
Spirograph PMHT	9.7	10.0
PMHT (no slide)	9.8	5.1
Adap./Hom. PMHT	11.4	12.8

Note: First column is in terms of in-track percentage, referring to the proportion of tracks that were not "lost" in all 30 scans of data. Second column refers to rank in terms of RMSE, where RMSE is computed only for those tracks that were not lost.

We have the following comments.

PDAF: This tracker, unfortunately from our point of view, is almost consistently the best. Exceptions appear only to be in the case of *very* heavy clutter and maneuvering targets, situations in which certain of the PMHT algorithms perform better but not excitingly.

PMHT without Sliding: This is the basic form of the PMHT, performed on batches of length 6 which overlap by exactly one sample, and tracks do not grow. Generally the performance of this is inferior to other PMHTs. However, it is seen that when tracking *is* working, the RMSE is low.

Sliding PMHT: This is the basic PMHT, performed on sliding batches of length 12, with a skip parameter of 3 scans. In most cases this is better than the basic PMHT with fixed batches; but other PMHTs have better performance. All PMHTs except the first use such a sliding window and batch length.

TABLE II
Illustration of Performance of Various Tracking Approaches, in Terms of In-Track Percentage

	$\sigma_p = 0.01$			$\sigma_p = 0.001$			$\sigma_p = 0.0001$		
	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$
PDAF	1	4	19	4	4	29	3	10	29
PMHT (no slide)	3	7	14	6	10	20	13	12	18
PMHT (sliding)	6	8	23	10	18	30	6	8	26
Homothetic PMHT	9	13	32	6	16	25	8	12	18
Spirograph PMHT	4	10	20	5	10	18	7	15	16
Adap./Hom. PMHT	0	0	2	0	2	2	1	1	3
PMHT (PDAF meas)	4	14	25	8	9	25	5	17	33
D.O. PMHT ($P = 3$)	6	7	21	10	17	29	5	8	25
Maneuver PMHT	19	32	50	25	34	50	24	35	58
Man. PMHT (hom)	16	41	50	19	37	59	20	33	59
Man. PMHT (jnt)	2	7	13	2	10	24	4	12	9
Deflat. PMHT	6	16	37	15	14	34	7	18	28
Homosc. PMHT	11	31	25	13	23	34	11	21	30

Note: The situation here is of initial speed 13.8 knots, clutter density $10^{-5.5}$ per square meter, and $\sigma_m = 100$ m.

TABLE III
Illustration of Performance of Various Tracking Approaches, in Terms of In-Track Percentage

	$\sigma_p = 0.01$			$\sigma_p = 0.001$			$\sigma_p = 0.0001$		
	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$
PDAF	60	86	94	66	90	95	64	89	94
PMHT (no slide)	32	46	67	45	62	69	40	65	64
PMHT (sliding)	33	57	75	47	67	72	46	71	76
Homothetic PMHT	53	70	91	51	76	87	49	68	81
Spirograph PMHT	42	56	76	42	58	76	43	67	69
Adap./Hom. PMHT	6	23	42	15	39	62	19	31	49
PMHT (PDAF meas)	34	64	70	44	54	74	38	65	74
D.O. PMHT ($P = 3$)	35	68	81	49	70	76	46	75	80
Maneuver PMHT	35	55	71	45	67	88	57	69	84
Man. PMHT (hom)	49	68	87	49	68	85	46	80	93
Man. PMHT (jnt)	40	68	91	43	73	89	40	71	86
Deflat. PMHT	40	61	85	44	72	90	52	74	85
Homosc. PMHT	43	71	88	54	72	82	37	67	81

Note: Situation here is of initial speed 13.8 knots, clutter density $10^{-6.5}$ per square meter, and $\sigma_m = 100$ m.

TABLE IV
Illustration of Performance of Various Tracking Approaches, in Terms of In-Track Percentage

	$\sigma_p = 0.01$			$\sigma_p = 0.001$			$\sigma_p = 0.0001$		
	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$
PDAF	16	65	87	87	92	97	88	96	97
PMHT (no slide)	18	23	37	63	78	75	61	74	87
PMHT (sliding)	27	34	43	74	78	88	69	79	92
Homothetic PMHT	30	56	71	76	91	97	81	93	97
Spirograph PMHT	2	15	35	52	79	92	63	76	86
Adap./Hom. PMHT	2	7	26	56	87	95	67	84	96
PMHT (PDAF meas)	5	12	33	40	73	90	46	72	86
D.O. PMHT ($P = 3$)	10	31	43	48	79	93	59	88	96
Maneuver PMHT	49	42	58	66	76	87	57	83	93
Man. PMHT (hom)	22	39	59	81	91	97	74	93	96
Man. PMHT (jnt)	14	28	64	74	88	91	75	91	93
Deflat. PMHT	14	53	76	66	87	93	64	86	95
Homosc. PMHT	31	38	71	32	48	65	34	45	65

Note: Situation here is of initial speed 13.8 knots, clutter density 10^{-5} per square meter, and $\sigma_m = 10$ m.

TABLE V
Illustration of Performance of Various Tracking Approaches, in Terms of RMSE at Final Scan for Those Tracks Not Considered Lost

	$\sigma_p = 0.01$			$\sigma_p = 0.001$			$\sigma_p = 0.0001$		
	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$	$P_d = .5$	$P_d = .7$	$P_d = .9$
PDAF	150	100	90	120	70	50	120	60	50
PMHT (no slide)	120	110	110	130	70	70	90	80	60
PMHT (sliding)	150	100	100	140	80	60	120	90	70
Homothetic PMHT	170	130	110	190	100	80	180	90	110
Spirograph PMHT	140	110	120	170	110	100	190	160	110
Adap./Hom. PMHT	330	330	320	370	390	310	360	370	360
PMHT (PDAF meas)	140	100	80	80	70	60	90	70	60
D.O. PMHT ($P = 3$)	130	120	100	130	90	70	140	90	70
Maneuver PMHT	110	130	110	80	90	80	120	90	80
Man. PMHT (hom)	160	170	110	150	130	90	160	130	100
Man. PMHT (jnt)	200	170	140	190	160	120	230	180	110
Deflat. PMHT	90	100	90	100	60	60	120	70	60
Homosc. PMHT	110	110	100	90	70	60	130	70	60

Note: Situation here is of initial speed 13.8 knots, clutter density $10^{-6.5}$ per square meter, and $\sigma_m = 100$ m.

Homothetic PMHT: This simple modification improves the PMHT enormously in almost all cases. Naturally, due to the increased size of the assumed measurement noise, the RMSE is comparatively large.

Spirograph PMHT: It is somewhat surprising that this version of the PMHT is not promising, and appears to be worse than the sliding PMHT.

Adaptive/Homothetic PMHT: This modification shows very little promise.

PMHT with PDAF Measurement Model: It was expected that there was little difference between the PMHT measurement model and that of the PDAF, and indeed that was found; our contention, therefore, is that PMHT problems are those described in Section IID, not the adulterated association model. One interesting point about this PMHT variant is that its mean square error (MSE) performance among tracks that are not lost is overall best.

Detection-Oriented PMHTs: This was a strong performer. Its results were similar to those of the homothetic PMHT. This is a promising PMHT.

Maneuver PMHT: This PMHT does well, particularly in those situations in which the PDAF is poor.

Independent Homothetic/Maneuver PMHT: This PMHT is better than the previous, and is a particularly strong performer.

Joint Homothetic/Maneuver PMHT: The performance of this PMHT was disappointing. It is felt that some degree of tuning is possible, hence it will not be abandoned.

Deflationary PMHT: This convergence-aided variant is the best overall PMHT, with in-track performance similar to that of the PDAF and MSE considerably better.

Homoscedastic PMHT: This convergence aid works well, but is inferior to the deflationary

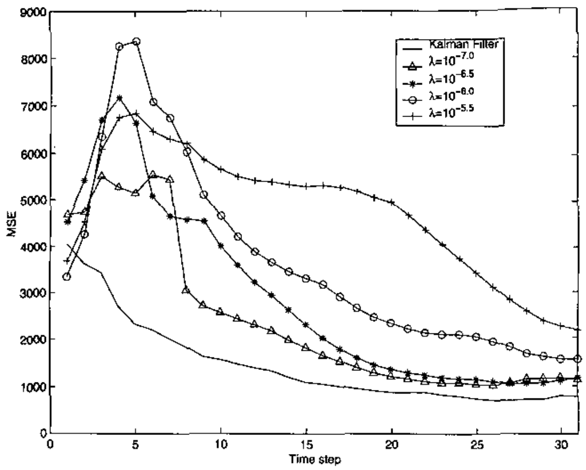

Fig. 10. MSE of Kalman filter with no association ambiguity and PMHT with $\lambda = 10^{-7.0}$; $\lambda = 10^{-6.5}$; $\lambda = 10^{-6.0}$; $\lambda = 10^{-5.5}$. We use $P_d = 99\%$, $\sigma_p = .01$, $\sigma_m = 100$, and initial speed 13.8 knots in the simulation.

PMHT. Improvement may be possible with some tuning.

In this paper we have compared the PMHT only with the PDAF, with the idea, as mentioned in the introduction, that its performance relative to other trackers could be inferred from that as measured against the PDAF. However, it is worth mentioning that the PDAF (and most trackers) are *filters*, whereas the PMHT performs *smoothing* over its batch of observations. It may at first seem that this gives an advantage to the PMHT in terms of its MSE performance. However, in all results shown, the MSE is that relating to its estimate at the *end* of its current batch: for example, given a batch length of 6, the MSE recorded for scan 11 is that from the PMHT's estimate based on observations from scans 6 through 11. Thus, since the estimate returned by a smoothing algorithm for its most-recent sample is identical to that reported by a filtering algorithm, the comparison is fair. Nonetheless, we also test the PMHT versus an idealized smoother—the Kalman smoother with no association ambiguity. The MSE is shown in Fig. 10.

V. SUMMARY

The PMHT is an elegant tracking algorithm, but has not up to now shown superiority to the simple PDAF. In this paper we have tried an extensive number of modifications to the PMHT, with the goal of finding those that are promising enough to pursue and "tune." The situation explored is of a single kinematic model in clutter, and is linear in both plant and measurement. We have attempted to keep parameters in line with active sonar practice.

Before describing the modifications the basic PMHT's "problems" were discussed in some detail. These include the PMHT's "nonadaptivity," meaning that unlike the PDAF its look-region is not data-adaptive; the PMHT's "narcissism," meaning that the PMHT is unwilling to believe its track lost even in the face of seemingly convincing evidence to the contrary in terms of a long string of missed detections; and the PMHT's "hospitality" to a perceived plurality of detections, many of which can be clutter.

We make a number of observations as follows.

1) The PMHT is improved by a sliding (as opposed to jumping) batch of scans.

2) The lost-track performance of the PMHT is much improved by a homothetic measurement model; the price paid is in terms of tracking accuracy (MSE).

3) Whatever problems the PMHT may have, they do not appear to have much to do with the measurement model: a PMHT with a PDAF-like measurement model (at most one measurement per target per scan) is little different from the standard PMHT in terms of lost-track performance. In terms of MSE this PMHT variant offers considerable improvement over the PDAF, presumably due to its use of multiple scans of data and consequent ability to smoothen.

4) A version of the PMHT which makes the event that several scans in a row are all detection free infeasible is a strong performer, and worthy of further study.

5) The use of a maneuver model, particularly a maneuver model with a homothetic measurement, seems to be a good idea even when the target does not maneuver. The ability of the track to execute a maneuver appears to encourage the PMHT to be less narcissistic.

6) Two convergence-enhancement PMHT variants are studied, one based on measurement noise inflation and the other on multiple initializations. The former performs particularly well: it is essentially as good as the PDAF in terms of lost tracks, and is much better in terms of MSE.

7) In fact, several PMHTs are consistently better than the PDAF in terms of RMSE for tracks that are not lost. In addition to that in the previous bullet, these include the standard PMHT, the PMHT using a PDAF measurement model, and the maneuvering PMHT.

8) In extremely adverse situations certain PMHT variants are much better than the PDAF in terms of lost tracks.

As complement to derivation of its variants, a number of practical hints for PMHT use—for example, in the modeling of clutter—are given in Section IIC.

Thus, if the environment is benign and a track is unlikely to be lost, a PMHT may be preferable to a PDAF in terms of tracking accuracy. In more interesting tracking situations in which lost tracks become of greater importance, the PDAF is preferable to most PMHT variants discussed. However, that version of the PMHT in which convergence is aided via an *inflated* measurement noise covariance almost matches the PDAF in terms of lost tracks, and appears preferable in terms of MSE. These are, we hope, reasonably positive results, and we additionally trust that we have in this paper illustrated the ease by which the PMHT can be modified to incorporate different modeling assumptions. We must also point out one remarkably simple PMHT "variant" can track multiple targets; this is in fact the PMHT as originally developed in [21].

The PMHT is based on an unusual measurement model: observations associate themselves to tracks (rather than being claimed by tracks), and the upshot is that zero, one, some or all measurements arise from a given target at a given scan are all feasible events. We do *not* in particular think that this assumption is better than the standard (at most one measurement per target). Possibly it is worse; but it does allow an elegant algorithm to be developed.

Part of the point of the paper is to show that the (possibly poor) assumption does not in fact yield poor performance. In fact, we can and do make a PMHT variant that is based on the more usual PDAF measurement model: it is, perhaps, a multiscan version of the PDAF. If the basic PMHT's measurement model were a problem, then presumably this PMHT variant would stand out; but in terms of lost tracks it is no better than a standard PMHT. We claim, in fact, that the greater concern with the PMHT is to avoid its convergence to a local maximum of likelihood, and in this paper we have attempted to suggest a number of ways this can be discouraged.

REFERENCES

[1] Avitzour, D. (1992)
A maximum likelihood approach to data association.
IEEE Transactions on Aerospace and Electronic Systems, **28**, 2 (Apr. 1992), 560–565.

[2] Bar-Shalom, Y., and Li, X. R. (1993)
Estimation and Tracking: Principles, Techniques and Software.
Boston: Artech House, 1993.

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

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

[5] Cover, T. M., Thomas, J. A. (1991)
Elements of Information Theory.
(series in telecommunications), New York: Wiley, 1991.

[6] Dempster, A., Laird, N., and Rubin, D. (1997)
Maximum likelihood from incomplete data via the EM algorithm.
Journal of the Royal Statistical Society, **39** (1977), 1–88.

[7] Dunham, D., and Hutchins, R. (1997)
Tracking multiple targets in cluttered environments with a probabilistic multi-hypothesis tracker.
In *Proceedings of SPIE Conference 3086*, Apr. 1997, 284–295.

[8] Giannopoulos, E., Streit, R., and Swaszek, P. (1996)
Probabilistic multi-hypothesis tracking in a multi-sensor, multi-target environment.
Presented at First Australian Data Fusion Symposium, Nov. 1996.

[9] Gauvrit, H., LeCadre, J-P., and Jauffret, C. (1997)
A formulation of multitarget tracking as an incomplete data problem.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 4 (Oct. 1997), 1242–1257.

[10] Gauvrit, H., LeCadre, J-P., and Jauffret, C. (1998)
Combinatorial optimization for initialization of probabilistic approaches.
In *Proceedings of Workshop Commun GdR ISIS (GT 1) and NUWC "Approches Probabilistes pour l'Extraction Multipistes,"* Nov. 1998.

[11] Hutchins, R. (1998)
Private Communication, Apr. 1998.

[12] Jauffret, C., and Bar-Shalom, Y. (1990)
Track formation with bearing and frequency measurements.
IEEE Transactions on Aerospace and Electronic Systems, **26**, 4 (Nov. 1990).

[13] Kamen, E. (1992)
Multiple target tracking based on symmetric measurement equations.
IEEE Transactions on Automatic Control, **37**, 3 (Mar. 1992), 371–374.

[14] Krieg, M., and Gray, D. (1997)
Multi-sensor probabilistic multi-hypothesis tracking using dissimilar sensors.
In *Proceedings of SPIE Conference 3086*, Apr. 1997, 129–138.

[15] Logothetis, A., Krishnamurthy, V., and Holst, J. (1997)
On maneuvering target tracking via the PMHT.
In *Proceedings of the Conference on Decision and Control*, Dec. 1997.

[16] Molnar, K., and Modestino, J. (1998)
Application of the EM algorithm for multi-target/multi-sensor tracking problem.
IEEE Transactions on Signal Processing, (Jan. 1998).

[17] Rago, C., Willett, P., and Streit, R. (1995)
Direct data fusion using the PMHT.
In *Proceedings of the 1995 American Control Conference*, June 1995.

[18] Rago, C. Willett, P., and Streit, R. (1995)
A modified PMHT.
In *Proceedings of the 1995 Conference on Information Sciences and Systems*, Mar. 1995.

[19] Ruan, Y., Willett, P., and Streit, R. (1998)
The PMHT for maneuvering targets.
In *Proceedings of the 1998 SPIE Conference on Signal and Data Processing of Small Targets*, Apr. 1998.

[20] Ruan, Y., Willett, p., and Streit, R. (1999)
Comparison of PMHT and SD-assignment trackers.
In *Proceedings of 1999 SPIE Conference 3692*, Orlando FL, Apr. 1999.

[21] Streit, R. L., and Luginbuhl, T. E. (1995)
Probabilistic multi-hypothesis tracking.
NUWC-NPT technical report 10,428, Feb. 1995.

[22] Titterington, D., Smith, A., and Makov, U. (1985)
Statistical Analysis of Finite Mixture Distributions.
New York: Wiley, 1985.

[23] Willett, P., Ruan, Y., and Streit, R. (1998)
A variety of PMHTs.
University of Connecticut technical report 1998-4, October 1998.

Peter K. Willett (S'83—M'86—SM'97) was born in Toronto, Ontario, Canada. He received the B.Sc. degree in engineering science from the University of Toronto in 1982. He received the M.E. and M.S. degrees in 1983 and 1984, respectively, and the Ph.D. degree in 1986, all in electrical engineering, from Princeton University, Princeton, NJ.

He is a Professor with the University of Connecticut, Storrs, where he has been since 1986. His interests are generally in detection theory, target tracking, and signal processing.

Dr. Willett is an Associate Editor for the *IEEE Transactions on Systems, Man and Cybernetics* and the *IEEE Transactions on Aerospace and Electronic Systems*.

Yanhua Ruan received B.S. and M.S. degrees in electrical engineering from Peking University, Beijing, China in 1994 and 1997, respectively.

He is currently a Ph.D. candidate and a research assistant in the Electrical and Computer Engineering Department of the University of Connecticut, Storrs. His research interests are in the areas of signal processing, detection, and estimation theory.

Roy L. Streit received the Ph.D. in mathematics from the University of Rhode Island, Kingston, in 1978.

He has been with the Naval Undersea Warfare Center since 1970, when he joined one of its predecessor organizations, the Navy Underwater Sound Laboratory in New London, CT. From 1981–1982, he was a Visiting Scholar with the Department of Operations Research, Stanford University. From 1982 to 1984, he was a Visiting Scientist with the Computer Science Department, Yale University. From 1987 to 1989, he was an Exchange Scientist with the Defence Science and Technology Organization in Adelaide, Australia. His current research interests include passive detection and localization of distributed targets in hyperspectral images, tracking in random media using low fidelity environmental wave propagation models, and numerical acoustic hull array design. He is currently a Senior Technologist for Acoustic Signal Processing.

Dr. Streit received the 1999 Solberg Award from the American Society of Naval Engineers for outstanding achievement in research and development related to naval engineering.