Tracker Operating Characteristic for Integrated Probabilistic Data Association

I. Vaughan L. Clarkson*,†

* P.O. Box 920, Samford Village, Qld., 4520, AUSTRALIA
† Institute for Integrated and Intelligent Systems
Griffith University, 170 Kessels Rd., Nathan, Qld., 4011, AUSTRALIA
v.clarkson@ieee.org

Jason L. Williams‡,§

‡ National Security, Intelligence, Surveillance and Reconnaissance Division
Defence Science and Technology Group, AUSTRALIA
§ School of Electronic Engineering and Computer Science
Queensland University of Technology, Brisbane, AUSTRALIA
jason.williams@dst.defence.gov.au

***Abstract*—Integrated Probabilistic Data Association (IPDA) is a popular approach to target tracking and to track initiation and termination. In regards to track initiation, it may be said to belong to the "detect-then-track" class in which processing begins with sensor-level detections. This is in contrast to "track-before-detect" (TkBD) approaches in which un-thresholded sensor-level measurements are directly incorporated into a track-initiation statistic. In IPDA, the track-initiation statistic is the probability of existence (PoE). How should the confirmation threshold be set? We perform a statistical analysis that yields an accurate approximation of the false-track and track detection probabilities as a function of the threshold on the PoE. From these probabilities, IPDA's tracker operating characteristic (TOC) can be derived. We further propose a tracker using similar ideas but which belongs to the TkBD class. Its TOC is likewise derived and compared with IPDA. The results of numerical simulations are presented which support the theoretical derivations.**

I. INTRODUCTION

Target tracking is a core element of radar and sonar processing but reliable, automatic track initiation and termination continues to be a challenging research problem. The Integrated Probabilistic Data Association (IPDA) algorithm of Musicki, Evans and Stankovic [1] is popular because it is simple to compute, moderately robust to sensor-level false alarms and readily extended to multi-target tracking. Importantly, it also integrates a statistic, the track's probability of existence (PoE), which can be used to initiate and terminate a track. The PoE is updated after each sensor scan. When it exceeds a particular threshold, the track is confirmed. When it falls below another threshold, the track is terminated.

Most approaches to track initiation fall into one of two classes: track-before-detect (TkBD) or detect-then-track (DTTk) [2]. IPDA belongs to the latter class, in which the sensor performs hard thresholding on measurements before tracking or track initiation takes place. TkBD, on the other hand, makes use of un-thresholded returned amplitudes or energies directly in computing a track detection statistic.

As first proposed in [1] and in much of Musicki's later work on IPDA and its variants, there is little attempt to analyse thresholds. As well as a threshold, the PoE itself must be initialised. Initial values are assigned somewhat arbitrarily, based on experience. Thresholds are tweaked on the same basis or to achieve a level of parity with competing algorithms. For instance, the value 0.2 is proposed as an initial PoE in the original IPDA paper [1]. A value of 0.5 is rejected in [1], [3] with the suggestion in the former case that it is too high and in the latter that it promotes confirmation of false tracks. Without discussion, a value of 0.01 is used for simulations in [4] along with a confirmation threshold of 0.999.

We are interested in how the track-confirmation threshold affects system performance. System performance is gauged here by false-track and track detection probabilities. The false-track probability is the probability that a track is confirmed when no track exists, *i.e.*, where all detections available to the tracker are sensor false-alarms. The track detection probability is the probability that the measurements arising from a true track result in confirmation. The false-track and the track detection probabilities are implicitly dependent on one another through the track confirmation threshold. When one is plotted against the other or against the target signal-to-noise ratio (SNR), the curve is known as the tracker operating characteristic or TOC [5], similar to the receiver operating characteristic in detection theory and closely related to the system operating characteristic of Bar-Shalom *et al.* [6]. The TOC illustrates the discriminating power of a track initiation statistic. Our aim in this paper is to accurately approximate false-track and track detection probabilities for IPDA in closed form and thus to plot its TOC.

In analysing IPDA, we will decompose the probability-of-existence statistic into a number of terms, each of which themselves can be regarded as track-initiation statistics on individual sequences of measurements. These "sub-statistics" can be interpreted as generalised likelihood ratios (GLRs). Their form is reminiscent of maximum-likelihood PDA (ML-PDA) [7], [8]. Blanding *et al.* [9] has analysed ML-PDA in a similar context, in which ML-PDA is applied to the set of sequences formed from combinations of measurements in a batch of sensor scans. They use extreme-value theory to approximate the false-track probability in the batch. We will find that a simpler approach suffices for IPDA.

Before analysing IPDA, we propose and analyse a related tracker which belongs instead to the TkBD class. Developing the key idea we use for IPDA, we express it as an amalgamation of sub-statistics, each of which is a GLR on candidate sequences

II. SYSTEM MODEL

A. Sensor Model

The sensor performs scans at regular intervals t_{scan}. We will process data from a batch of N consecutive scans, using n to index the scans. The sensor scans a k_{m}-dimensional surveillance volume $\mathcal{S} \subset \mathbb{R}^{k_{\text{m}}}$ of volume $\operatorname{vol}\mathcal{S}$ partitioned into C contiguous resolution cells. A cell measures a complex amplitude $b_{c,n}$ and a position $\mathbf{z}_{c,n} \in \mathcal{S}$ for $c = 1, \ldots, C$ and $n = 1, \ldots, N$. In TkBD, the cell measurements are used directly for track initiation. Amplitude is always independent of position.

In DTTk, the amplitude measurements are thresholded before being used for track initiation. A threshold ζ is applied to the squared modulus of each amplitude measurement, *i.e.*, it is received power or energy that is thresholded. At the nth scan, f_n cell energies exceed the threshold. The corresponding positions are copied to the sensor output. These outputs are labelled $\mathbf{y}_{m,n}$, with $\{\mathbf{y}_{m,n}\} \subseteq \{\mathbf{z}_{c,n}\}$.

In the absence of a target, amplitude measurements are independent and identically distributed (i.i.d.) standard circularly-symmetric complex normal random variables (r.v.s). The probability that received energy exceeds the threshold ζ is the sensor's false-alarm probability, P_{FA}. The spatial intensity of false alarms is denoted $\lambda \triangleq P_{\text{FA}} C / \operatorname{vol} \mathcal{S}$. The thresholded position measurements are i.i.d. uniform r.v.s in $\mathcal{S}$.

B. Target Model

If a target exists, it is observable in all N scans in the batch. A target has a k_{s}-dimensional state $\mathbf{x}$ which evolves from scan to scan according to a standard linear, Gaussian model in which $\mathbf{x}_n = \mathbf{F}\mathbf{x}_{n-1} + \boldsymbol{\xi}_n$ where $\mathbf{F}$ is the state update matrix and $\boldsymbol{\xi}_n$ is plant noise. The plant noise we assume to be i.i.d. multivariate normal with zero mean and covariance $\mathbf{Q}$. It follows that the state is itself a r.v. except for the first state, $\mathbf{x}_1$, which we regard as a parameter.

At the sensor, the target return is seen in only one cell in each scan. Its amplitude measurement is complex normal and has the same unit variance as in the noise-only case but the mean becomes μ_n. For TkBD, we do not specify a distribution on the μ_n. We regard them as nuisance parameters. However, for DTTk, we assume that the probability of detection, P_{D}, is known and independent of n.

The measured position of the target is $\mathbf{y}_n = \mathbf{H}\mathbf{x}_n + \boldsymbol{\eta}_n$ where $\mathbf{H}$ is the measurement matrix and $\boldsymbol{\eta}_n$ is measurement noise. The $\boldsymbol{\eta}_n$ i.i.d. multivariate normal with zero mean and covariance $\mathbf{R}$. Each $\mathbf{y}_n$ appears amongst the $\mathbf{z}_{c,n}$ and, if detected, amongst the $\mathbf{y}_{m,n}$.

III. Example Scenario

To illustrate our results, we introduce an example scenario. The scenario is inspired by and, in the DTTk case, closely matches the original scenario illustrating IPDA in Musicki *et al.* [1].

The sensor performs one scan per second, *i.e.*, $t_{\text{scan}} = 1\,\text{s}$. It measures spatial positions only, not Doppler. The surveillance area is a rectangle whose dimensions are approximately $h \times w$ where $h = 1000\,\text{m}$ and $w = 400\,\text{m}$. The area is divided into $58 \times 23 = 1334$ square cells. Each cell has an area of $300\,\text{m}^2$. The energy threshold applied to each cell is $\zeta = 7.01$ which gives $P_{\text{FA}} = 0.03$ and $\lambda \approx 1 \times 10^{-4}\,\text{m}^{-2}$. The standard deviation on position measurement is $5\,\text{m}$ in each direction, consistent with a uniform distribution over the cell but assumed to be Gaussian when a target is present.

We assume a Swerling Type I scintillating target with an SNR of $15.09\,\text{dB}$, giving $P_{\text{D}} = 0.9$. The target state is four-dimensional, consisting of target position and velocity. It has state-update and plant-noise covariance matrices

$$\mathbf{F} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \otimes \mathbf{I}_2 \quad \text{and} \quad \mathbf{Q} = q \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} \otimes \mathbf{I}_2$$

where $q = 0.75$, $\otimes$ denotes the Kronecker product and $\mathbf{I}_2$ is the 2×2 identity matrix.

IV. Statistical Hypotheses

We next propose two hypothesis tests, one suitable for TkBD and the other for DTTk. Each applies to a candidate sequence of measurements, to decide whether the sequence results from a target or from sensor noise. The indices c_n and m_n that we introduce to select measurements from each scan are regarded as hidden parameters.

For TkBD, given a tentative track consisting of the sequence of cells c_n, $n = 1, \ldots, N$, we wish to test the hypotheses

$$\begin{aligned} H_0: &\quad \mathbf{z}_{c_n,n} \sim U(\mathcal{S}), b_{c_n,n} \sim CN(0,1), \\ H_1: &\quad \mathbf{z}_{c_n,n} \sim N(\mathbf{H}\mathbf{x}_n, \mathbf{R}), \mathbf{x}_n \sim N(\mathbf{F}\mathbf{x}_{n-1}, \mathbf{Q}), \\ &\quad b_{c_n,n} \sim CN(\mu_n, 1), \end{aligned} \tag{1}$$

for some suitable $\mathbf{x}_n$ and μ_n, $n = 1, \ldots, N$, and where $U(\cdot)$, $N(\cdot,\cdot)$ and $CN(\cdot,\cdot)$ are abbreviations for the uniform, normal and complex normal distributions discussed in Section II.

For DTTk, given a tentative track consisting of the sequence of detections m_n, $n = 1, \ldots, N$, the hypotheses are instead

$$\begin{aligned} H_0: &\quad \mathbf{y}_{m_n,n} \sim U(\mathcal{S}), \\ H_1: &\quad \mathbf{y}_{m_n,n} \sim (1 - P_{\text{D}}) U(\mathcal{S}) + P_{\text{D}} N(\mathbf{H}\mathbf{x}_n, \mathbf{R}), \\ &\quad \mathbf{x}_n \sim N(\mathbf{F}\mathbf{x}_{n-1}, \mathbf{Q}), \end{aligned} \tag{2}$$

where the use of addition in relation to the distribution in H_1 indicates a mixture distribution with the given weights.

V. Generalised Likelihood-Ratio Tests

A. Track Before Detect

Consider first the TkBD statistical hypotheses of (1). Since amplitudes and positions are mutually independent, the likelihood is separable under both H_0 and H_1. In the alternate hypothesis, the estimates, $\hat{\mu}_n$, can be chosen to fit the samples perfectly. Therefore, in the GLR test that we propose, the likelihood of the $b_{c_n,n}$ under H_1 can be neglected. The likelihood associated with the positions under the null hypothesis is constant throughout the surveillance volume, being uniform, and can likewise be ignored.

The GLR is therefore a ratio of the likelihood of the amplitudes under the null hypothesis to the likelihood of the positions under the alternative hypothesis. Writing $a_{c_n,n} = |b_{c_n,n}|^2$ as the received energy, maximising the likelihood of the $\mathbf{x}_n$ using the Kalman filter and neglecting constant terms, the GLR would appear to be

$$T_{\mathbf{c}} = \prod_{n=1}^{N} \frac{\exp(-a_n/2)}{\mathcal{N}(\mathbf{z}_n; \mathbf{H}\hat{\mathbf{x}}_{n|n-1}, \mathbf{S}_n)} \tag{3}$$

where we have dropped the c_n subscripts for clarity, except from T itself, where it is retained in vector form, and $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the likelihood of the multivariate normal with the given mean and covariance. The mean here is $\mathbf{H}\hat{\mathbf{x}}_{n|n-1}$ and the covariance is $\mathbf{S}_n = \mathbf{H}\mathbf{P}_{n|n-1}\mathbf{H}^T + \mathbf{R}$, where $\hat{\mathbf{x}}_{n|n-1}$ is the maximum-likelihood prediction for $\mathbf{x}_n$ given the position estimates up to scan $n-1$, obtained from the Kalman filter, and $\mathbf{P}_{n|n-1}$ is its covariance.

We must take care with the denominator of (3), however. The first state, $\mathbf{x}_1$, is an unknown parameter, lacking a distribution, so we are led to use the information filter, initialised with zero information, rather than the standard Kalman filter. We then need $I > 0$ initialising scans before the information matrix corresponding to $\mathbf{S}_n$ reaches full rank. We denote this matrix $\mathbf{S}_n^{-1}$ even though $|\mathbf{S}_n^{-1}| = 0$ for $n \leqslant I$. When $n \leqslant I$, the corresponding terms in the denominator of (3) are zero. Therefore, the product in the denominator is taken from $n = I + 1$ rather than from $n = 1$. As an example, consider our scenario of Section III. Each scan yields a two-dimensional measurement of position but the target state is four-dimensional, so we need $I = 2$ "initialising" scans before the state information matrix is full rank. It is only from the third scan that likelihoods can be calculated from position measurements.

Taking the logarithm of the statistic and discarding further constant terms, we have

$$t_{\mathbf{c}} = \sum_{n=1}^{N} \left\{ \|\mathbf{S}_n^{-1/2}(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})\|^2 - a_n \right\}. \tag{4}$$

The summation here does not need the same care as was the case in the denominator of (3) since the problematic part of the p.d.f., the determinant of $\mathbf{S}_n^{-1}$, is discarded amongst the other constant terms.

Under the null hypothesis, the a_n are χ^2_2 distributed and the $\mathbf{y}_n$ are uniformly distributed. Given a threshold θ on $t_{\mathbf{c}}$, we confirm the track, which under H_0 means we declare a false track, if

$$\ell(\mathbf{y}_1, \dots \mathbf{y}_N) \triangleq \sum_{n=1}^{N} \|\mathbf{S}_n^{-1/2}(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n|n-1})\|^2 < \theta + \sum_{n=1}^{N} a_n. \tag{5}$$

Consider the probability that the L.H.S. is less than the R.H.S. when the latter is held constant at a value α. To analyse this probability, we stack measurement and residual vectors to create

$$\mathbf{y} \triangleq \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{pmatrix} \quad \text{and} \quad \mathbf{r}(\mathbf{y}) \triangleq \begin{pmatrix} \mathbf{S}_1^{-1/2}[\mathbf{y}_1 - \mathbf{H}\hat{\mathbf{x}}_{1|0}] \\ \vdots \\ \mathbf{S}_N^{-1/2}[\mathbf{y}_N - \mathbf{H}\hat{\mathbf{x}}_{N|N-1}] \end{pmatrix} \tag{6}$$

where, as discussed above, the first I subvectors of $\mathbf{r}(\mathbf{y})$ are identically zero. We can then write $\ell(\mathbf{y}) = \|\mathbf{r}(\mathbf{y})\|^2$. Let

$$\mathcal{C}(\alpha) \triangleq \left\{ \mathbf{y} \in \mathbb{R}^{Nk_{\mathrm{m}}} \mid \ell(\mathbf{y}) < \alpha \right\}. \tag{7}$$

We see that, because $\mathbf{y}$ is uniformly distributed on $\mathcal{S}^N$,

$$\Pr\{\ell(\mathbf{y}) < \alpha\} = \frac{\mathrm{vol}[\mathcal{C}(\alpha) \cap \mathcal{S}^N]}{\mathrm{vol}\, \mathcal{S}^N}. \tag{8}$$

The nullspace of $\ell(\cdot)$ is non-trivial. The residuals are zero if the measurements conform without error to a target state that evolves strictly according to its model dynamics. That is, $\mathbf{r}(\mathbf{y}) = \mathbf{0}$ if and only if $\mathbf{y} \in \langle \mathbf{V} \rangle$ where $\langle \mathbf{V} \rangle$ represents the column space of the observability matrix, $\mathbf{V}$, a matrix of blocks of the form $\mathbf{H}\mathbf{F}^{n-1}$, $n = 1, \dots, N$, stacked vertically . Define the matrix $\mathbf{W}$ such that $\langle \mathbf{W} \rangle$ is the orthogonal complement to $\langle \mathbf{V} \rangle$ in $\mathbb{R}^{Nk_{\mathrm{m}}}$. Clearly, $\ell(\mathbf{y}) > 0$ if and only if $\mathbf{y}$ has a component in $\langle \mathbf{W} \rangle$. We define $\mathcal{A} \triangleq \langle \mathbf{V} \rangle$ and $\mathcal{B}(\alpha) \triangleq \{\mathbf{y} \in \langle \mathbf{W} \rangle \mid \ell(\mathbf{y}) < \alpha\}$. The set $\mathcal{B}(\alpha)$ has a well-defined volume with respect to $\langle \mathbf{W} \rangle$. This is the subject of the following lemma. The proof of the lemma is omitted for brevity.

Lemma 1. *Let* $\mathbf{U}$ *be an orthonormal basis for* $\mathcal{A}$ *and let* $\mathbf{U}$ *be partitioned vertically so that* $\mathbf{U}_1$ *is the square, upper submatrix. Then,*

$$\mathrm{vol}_{\langle \mathbf{W} \rangle}\, \mathcal{B}(\alpha) = \frac{(\pi\alpha)^{d/2} |\det \mathbf{U}_1|}{\Gamma(\frac{1}{2}d + 1)} \prod_{n=I+1}^{N} \sqrt{\det \mathbf{S}_n}$$

when $\alpha \geqslant 0$*, where* $d = Nk_m - k_s$.

Returning to the calculation of $\mathrm{vol}[\mathcal{C}(\alpha) \cap \mathcal{S}^N]$, we write

$$\begin{aligned} \mathrm{vol}[\mathcal{C}(\alpha) \cap \mathcal{S}^N] &= \mathrm{vol}[\{\mathcal{A} + \mathcal{B}(\alpha)\} \cap \mathcal{S}^N] \\ &\approx \mathrm{vol}_{\langle \mathbf{V} \rangle}(\mathcal{A} \cap \mathcal{S}^N)\, \mathrm{vol}_{\langle \mathbf{W} \rangle}\, \mathcal{B}(\alpha) \end{aligned} \tag{9}$$

which is increasingly accurate as α becomes smaller. The volume of $\mathcal{A} \cap \mathcal{S}^N$ is dependent on the particular geometry of the scenario, its target dynamics and measurements.

Example. For our scenario of Section III, we find that

$$\mathrm{vol}_{\langle \mathbf{V} \rangle}(\mathcal{A} \cap \mathcal{S}^N) = \frac{h^2 w^2 N^2 (N+1)}{12(N-1)} \tag{10}$$

and, from Lemma 1, that

$$\mathrm{vol}_{\langle \mathbf{W} \rangle}[\mathcal{B}(\alpha)] = \frac{12(\pi\alpha)^{N-2}}{N^2(N^2-1)} \prod_{n=3}^{N} \sqrt{\det \mathbf{S}_n}$$

so we have $\Pr\{\ell(\mathbf{y}) < \alpha\} = 0$ when $\alpha < 0$ or, when $\alpha \geqslant 0$,

$$\Pr\{\ell(\mathbf{y}) < \alpha\} \approx \frac{1}{(N-1)!(N-1)} \left(\frac{\pi\alpha}{hw}\right)^{N-2} \prod_{n=3}^{N} \sqrt{\det \mathbf{S}_n}.$$

We have found, then, that $\Pr\{\ell(\mathbf{y}) < \alpha\} \approx \beta_N \max\{\alpha^{d/2}, 0\}$, for some constant β_N where d is defined in Lemma 1. We are now able to evaluate the probability of declaring a false track, P_{FT}. For threshold θ, $P_{\mathrm{FT}} = \Pr\{t_{\mathbf{c}} < \theta\} = \Pr\{\ell(\mathbf{y}) < \theta + A_N\}$ where $A_N = \sum_{n=1}^{N} a_n$, the sum of the received energy along the hypothesised target trajectory, is a χ^2_{2N} r.v. It follows that $P_{\mathrm{FT}} \approx \beta_N E[\max\{(\theta + A_N)^{d/2}, 0\}]$. We omit the details of the derivation but the form of the expectation depends on whether θ is positive, negative or zero. In the case of practical interest, where $\theta > 0$,

$$P_{\mathrm{FT}} \approx \frac{\beta_N}{2} \left(\frac{\theta}{\sqrt{2}}\right)^{d+2} U(\tfrac{1}{2}d + 2, d + 3, \tfrac{1}{2}\theta) \tag{11}$$

where $U(a, b, z)$ is Tricomi's confluent hypergeometric function.

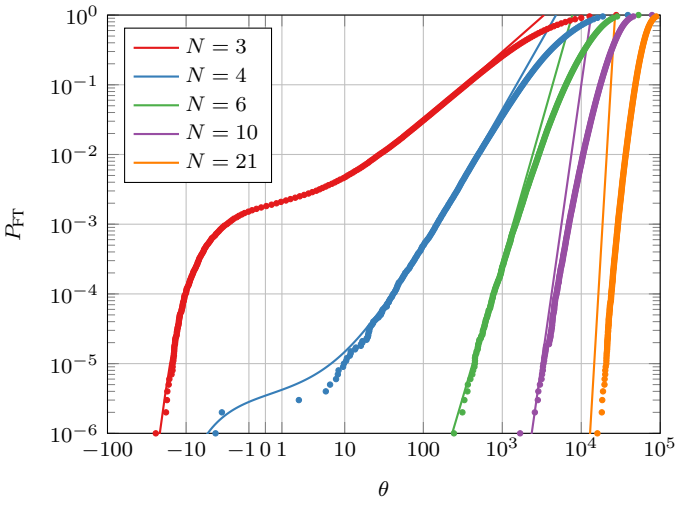

Fig. 1. Comparison of theoretical approximation and Monte Carlo simulation of false-track probabilities in track-before-detect.

Example. In Figure 1, we compare the theoretical approximation (11) relating the threshold value to the false-track probability against a Monte Carlo simulation involving a million trials for each value of N. Excellent agreement is observed when P_{FT} is sufficiently small. Because the log likelihood ratio (LLR) can take negative values, but also tends to scale exponentially with N, we plot the horizontal axis on an arcsinh scale.

Let's turn now to the alternate hypothesis and the probability of detecting a true track. In this case, the contribution to the LLR detection statistic $t_{\mathbf{c}}$ arising from Kalman-filter residues, *i.e.*, $\ell(\mathbf{y})$, is chi-square distributed with $Nk_{\mathrm{m}} - k_{\mathrm{s}}$ degrees of freedom. However, the "penalty" term, subtracted from these residues, arises from the returned energies along the postulated target positions. Under H_1, the signal amplitudes are unknown parameters to which noise is added. Hence, $A_N = \sum_{n=1}^{N} a_n$ has a non-central chi-squared distribution with $2N$ degrees of freedom.

Example. Suppose we are presented with a Swerling Type I target, as in our example scenario of Section III. If we knew in advance that all targets conformed to this model then the likelihood ratio test should account for it. Rather, let us leave the test unchanged and merely analyse its behaviour. Now, A_N/s has a chi-squared distribution with $2N$ degrees of freedom and s is a scaling factor such that $s = 1 + \text{SNR} = 1 + (\sigma^2_{\text{signal}}/\sigma^2_{\text{noise}})$. Again omitting the derivation, we find that, for $\theta \geqslant 0$,

$$P_{\text{DT}} = 1 - \frac{se^{-\theta/2}}{N!\,2^{N-3}(s+1)^{N+1}} \sum_{n=0}^{N-3} \frac{(n+N)!}{n!\,(N-3-n)!} \cdot \left(\frac{2s}{s+1}\right)^n \theta^{N-3-n}\, {}_2F_1\left(1, n+N+1; N+1; \frac{1}{s+1}\right)$$

where ${}_2F_1(a,b;c;z)$ is the Gaussian hypergeometric function.

The approximation for P_{DT} has similar or better accuracy to that for P_{FT}. However, we do not present comparisons with Monte Carlo simulations here, for this or for the other detection statistics presented in this paper, due to lack of space.

B. Detect Then Track

When we test the hypotheses of (2) for the DTTk model, the likelihood inverse-ratio, discarding constant terms, can be written

$$R_{\mathbf{m}} = \prod_{n=1}^{N} \left[\frac{1-P_{\text{D}}}{\text{vol}\,\mathcal{S}} + P_{\text{D}}\mathcal{N}(\mathbf{y}_n; \mathbf{H}\mathbf{x}_n, \mathbf{R})\right] \prod_{n=2}^{N} \mathcal{N}(\mathbf{x}_n; \mathbf{F}\mathbf{x}_{n-1}, \mathbf{Q})$$

The likelihood explicitly includes terms relating to the evolution of the state $\mathbf{x}_n$, which is missing in the TkBD likelihood ratio in (3), because the standard Kalman filter doesn't maximise their likelihood in a mixture distribution. Therefore, we distribute the product so that

$$R_{\mathbf{m}} = \sum_{\mathcal{D}\in\mathbb{P}_N} P_{\text{D}}^{|\mathcal{D}|}(1-P_{\text{D}})^{|\mathcal{M}|} \cdot \left[\prod_{n\in\mathcal{D}} \mathcal{N}(\mathbf{y}_n; \mathbf{H}\mathbf{x}_n, \mathbf{R}) \prod_{n=2}^{N} \mathcal{N}(\mathbf{x}_n; \mathbf{F}\mathbf{x}_{n-1}, \mathbf{Q})\right] \quad (12)$$

where $\mathbb{P}_N$ is the power set of $\{1,\ldots,N\}$, $\mathcal{D}$ is an element of the power set, itself a set, $|\mathcal{D}|$ is its cardinality, $\mathcal{M}$ is the complement of $\mathcal{D}$ in $\{1,\ldots,N\}$ and $|\mathcal{M}| = N - |\mathcal{D}|$.

The term inside the square brackets in (12), taken in isolation, is maximised with respect to the $\mathbf{x}_n$ using the Kalman filter, assuming that the only measurements available are at those scans indicated in $\mathcal{D}$. This implies a different set of state estimates for each element in the sum over $\mathcal{D}$. It yields an upper bound on the likelihood, namely

$$R_{\mathbf{m}} \leqslant \sum_{\mathcal{D}\in\mathbb{P}_N} L_{\mathcal{D}}(\mathbf{y}) \qquad \text{where} \quad (13)$$

$$L_{\mathcal{D}}(\mathbf{y}) \triangleq P_{\text{D}}^{|\mathcal{D}|}\left(\frac{1-P_{\text{D}}}{\text{vol}\,\mathcal{S}}\right)^{|\mathcal{M}|} \prod_{n\in\mathcal{D}\setminus\mathcal{I}} \mathcal{N}(\mathbf{y}_n, \mathbf{H}\hat{\mathbf{x}}_{n,\mathcal{D}}, \mathbf{S}_{n,\mathcal{D}}). \quad (14)$$

We index the measurement covariance, $\mathbf{S}_n$, according to the element of $\mathbb{P}_N$ to which it belongs, *i.e.*, $\mathbf{S}_{n,\mathcal{D}}$. Likewise, the corresponding value of $\hat{\mathbf{x}}_{n|n-1}$ is labelled $\hat{\mathbf{x}}_{n,\mathcal{D}}$. We define a set $\mathcal{I}$ which consists of the first I elements of $\mathcal{D}$. (The assumption that I is independent of $\mathcal{D}$ is not completely general but holds for our example.) We see that $|\mathbf{S}^{-1}_{n,d}| > 0$ when $d \in \mathcal{T} \triangleq \mathcal{D} \setminus \mathcal{I}$. Although (13) is an upper bound, we will use it as an approximation.

When $|\mathcal{D}| \leqslant I$, $L_{\mathcal{D}}(\mathbf{y})$ is constant. Therefore, we choose to omit those terms from the sum in (13). We write the sum as instead being over $\mathcal{D} \in \mathbb{P}_{N,I}$, denoting the power set of $\{1,\ldots,N\}$ restricted to sets of cardinality greater than I.

As in (5), we define

$$\ell_{\mathcal{D}}(\mathbf{y}) \triangleq \sum_{n\in\mathcal{D}} \|\mathbf{S}^{-1/2}_{n,\mathcal{D}}(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n,\mathcal{D}})\|^2. \quad (15)$$

By taking the logarithm of $L_{\mathcal{D}}(\mathbf{y})$ in (14) but without inverting its sign or discarding constant terms, we see that $L_{\mathcal{D}}(\mathbf{y}) > \Theta$ if and only if $\ell_{\mathcal{D}}(\mathbf{y}) < 2(\theta - \delta_{\mathcal{D}})$ where $\theta = -\log\Theta$ and

$$\delta_{\mathcal{D}} = \tfrac{1}{2}\sum_{n\in\mathcal{T}} \log(|2\pi\mathbf{S}_{n,\mathcal{D}}|) - |\mathcal{D}|\log P_{\text{D}} - |\mathcal{M}|\log\left(\frac{1-P_{\text{D}}}{\text{vol}\,\mathcal{S}}\right). \quad (16)$$

We can stack those measurement and residual vectors indexed in $\mathcal{D}$ to form $\mathbf{y}_{\mathcal{D}}$ and $\mathbf{r}_{\mathcal{D}}(\mathbf{y}_{\mathcal{D}})$ analogously to $\mathbf{y}$ and $\mathbf{r}(\mathbf{y})$ in (6), define $\mathcal{C}_{\mathcal{D}}(\alpha)$ analogously to $\mathcal{C}(\alpha)$ in (7), *etc.*, so that, analogously to (9),

$$\text{vol}[\mathcal{C}_{\mathcal{D}}(\alpha) \cap \mathcal{S}^{|\mathcal{D}|}] \approx \text{vol}_{\langle\mathbf{V}_{\mathcal{D}}\rangle}(\mathcal{A}_{\mathcal{D}} \cap \mathcal{S}^{|\mathcal{D}|})\,\text{vol}_{\langle\mathbf{W}_{\mathcal{D}}\rangle}\,\mathcal{B}_{\mathcal{D}}(\alpha).$$

Moreover, the obvious corollary to Lemma 1 states that, when $\alpha \geqslant 0$,

$$\text{vol}_{\langle\mathbf{W}_{\mathcal{D}}\rangle}\,\mathcal{B}_{\mathcal{D}}(\alpha) = \frac{(\pi\alpha)^{d/2}|\det\mathbf{U}_{\mathcal{D},1}|}{\Gamma(\frac{1}{2}d+1)} \prod_{n\in\mathcal{T}} \sqrt{\det\mathbf{S}_{n,\mathcal{D}}}$$

where $d = |\mathcal{D}|k_{\text{m}} - k_{\text{s}}$.

To compute the false-track probability P_{FT}, we then have

$$\begin{aligned} P_{\text{FT}} &\leqslant \Pr\left\{\sum_{\mathcal{D}\in\mathbb{P}_{N,I}} L_{\mathcal{D}}(\mathbf{y}) > \Theta\right\} \\ &\approx \sum_{\mathcal{D}\in\mathbb{P}_{N,I}} \Pr\left\{L_{\mathcal{D}}(\mathbf{y}) > \Theta\right\} \\ &= \sum_{\mathcal{D}\in\mathbb{P}_{N,I}} \Pr\left\{\ell_{\mathcal{D}}(\mathbf{y}) < 2(\theta-\delta_{\mathcal{D}})\right\} \\ &= \sum_{\mathcal{D}\in\mathbb{P}_{N,I}} \frac{\text{vol}_{\langle\mathbf{V}_{\mathcal{D}}\rangle}(\mathcal{A}_{\mathcal{D}} \cap \mathcal{S}^{|\mathcal{D}|})\,\text{vol}_{\langle\mathbf{W}_{\mathcal{D}}\rangle}\,\mathcal{B}_{\mathcal{D}}[2(\theta-\delta_{\mathcal{D}})]}{\text{vol}\,\mathcal{S}^{|\mathcal{D}|}}. \end{aligned} \quad (17)$$

Example. For the example scenario of Section III, we can show that

$$P_{\text{FT}} \approx \sum_{\mathcal{D}\in\mathbb{P}_{N,2}} \frac{(d_2-d_1)^2}{(|\mathcal{D}|-2)!\,(d_{|\mathcal{D}|}-d_1)^2}\left(\frac{2\pi\max\{\theta-\delta_{\mathcal{D}},0\}}{hw}\right)^{|\mathcal{D}|-2} \cdot \sqrt{\det\mathbf{S}_{d_3,\mathcal{D}}}\cdots\sqrt{\det\mathbf{S}_{d_{|\mathcal{D}|},\mathcal{D}}}. \quad (18)$$

Figure 2 plots our approximation of the false-track probability in the example scenario. Excellent agreement is observed in all cases between the theory and the Monte Carlo simulation results, generated from one million trials for each value of N.

To compute the probability of track detection, we assume that the term involving the true set of scans in which the target was detected, $\mathcal{D}_0$, dominates in (13), so that $R_{\mathbf{m}} \approx L_{\mathcal{D}_0}(\mathbf{y})$. It follows that

$$P_{\text{DT}|\mathcal{D}_0} \approx \Pr\{\ell_{\mathcal{D}_0}(\mathbf{y}) < 2(\theta-\delta_{\mathcal{D}_0})\} \quad (19)$$

and therefore

$$P_{\text{DT}} = \sum_{\mathcal{D}_0\in\mathbb{P}_{N,I}} P_{\text{D}}^{|\mathcal{D}_0|}(1-P_{\text{D}})^{|\mathcal{M}_0|} P_{\text{DT}|\mathcal{D}_0} = E_{\mathcal{D}_0}[P_{\text{DT}|\mathcal{D}_0}]. \quad (20)$$

Expanding (19), we have

$$P_{\text{DT}|\mathcal{D}_0} \approx \Pr\left\{\sum_{n\in\mathcal{D}_0} \|\mathbf{S}^{-1/2}_{n,\mathcal{D}_0}(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n,\mathcal{D}_0})\|^2 < 2\theta - 2\delta_{\mathcal{D}_0}\right\}. \quad (21)$$

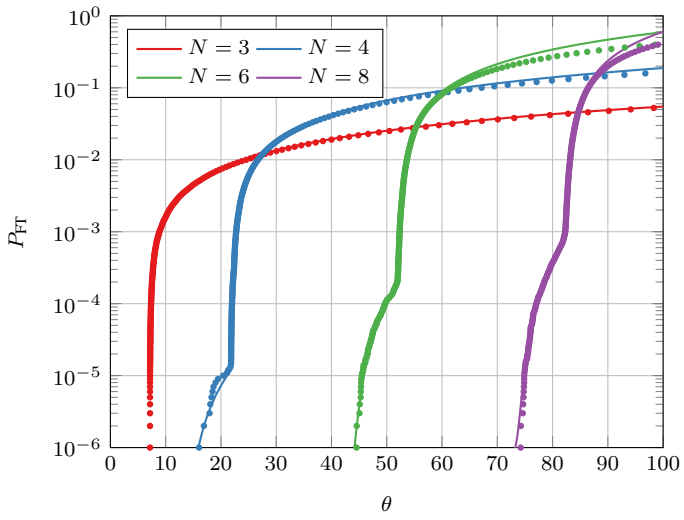

Fig. 2. Comparison of theoretical approximation and Monte-Carlo simulation of false-track probabilities in detect-then-track.

Under H_1, $\sum_{n\in\mathcal{D}_0} \|\mathbf{S}_{n,\mathcal{D}_0}^{-1/2}(\mathbf{y}_n - \mathbf{H}\hat{\mathbf{x}}_{n,\mathcal{D}_0})\|^2$ is a χ^2 r.v. with $|\mathcal{D}_0|k_{\text{m}} - k_{\text{s}}$ degrees of freedom. Therefore,

$$P_{\text{DT}|\mathcal{D}_0} \approx \frac{\gamma([|\mathcal{D}_0|k_{\text{m}} - k_{\text{s}}]/2, \max\{2\theta - 2\delta_{\mathcal{D}_0}, 0\})}{\Gamma(|\mathcal{D}_0|k_{\text{m}} - k_{\text{s}}]/2)}. \tag{22}$$

Example. For our recurring example of Section III, we combine (20) and (22) to obtain

$$P_{\text{DT}} \approx \sum_{\mathbb{P}_{N,2}} P_{\text{D}}^{|\mathcal{D}_0|}(1-P_{\text{D}})^{|\mathcal{M}_0|}\frac{\gamma(|\mathcal{D}_0| - 2, \max\{2[\theta - \delta_{\mathcal{D}_0}], 0\})}{(|\mathcal{D}_0| - 3)!} \tag{23}$$

where $\gamma(s,x)$ is the lower incomplete gamma function.

VI. Time Between False Tracks

Consider the GLR statistic proposed for TkBD in Section V-A. Each scan yields amplitude measurements in each of C different resolution cells. Therefore there are C^N arrangements of bins that might correspond to a track. If we assume that the thresholded statistics are i.i.d. Bernoulli r.v.s and that P_{FT} is small then the probability of at least one false track being declared in a *batch* of N scans is $P_{\text{FTIB}} = 1 - (1 - P_{\text{FT}})^{C^N} \approx C^N P_{\text{FT}}$. The aggregation of Bernoulli r.v.s in this way itself constitutes a new Bernoulli r.v. A new instance of the r.v. is available after every batch of N scans. Each new r.v. may be assumed to be independent of the last. The expected number of batches between false track declarations is therefore $P_{\text{FTIB}}^{-1} \approx (C^N P_{\text{FT}})^{-1}$. It follows that the expected time between false track declarations, t_{FT}, is obtained by multiplying by Nt_{scan} and hence $t_{\text{FT}} \approx (Nt_{\text{scan}})/(C^N P_{\text{FT}})$.

On the other hand, the GLR statistic for DTTk that we proposed in Section V-B relies on just the detected measurements, not those in every cell. Sensor false-alarms in each scan are governed by a spatial false-alarm rate λ. The number of false-alarms is approximately a Poisson r.v. with parameter $F = \lambda \operatorname{vol}\mathcal{S}$. Let the r.v. representing the number of detections in the nth scan be F_n. The F_n are i.i.d. with $E[F_n] = F$. We have, again for P_{FT} sufficiently small,

$$P_{\text{FTIB}} = 1 - E_{F_1,\ldots,F_N}[(1 - P_{\text{FT}})^{F_1\cdots F_N}] \approx F^N P_{\text{FT}}. \tag{24}$$

It follows that, for the DTTk track detection statistic of Section V-B, $t_{\text{FT}} \approx (Nt_{\text{scan}})/(F^N P_{\text{FT}})$.

VII. Analysis of IPDA

IPDA maintains an estimate of target state, covariance and probability of existence. The state and covariance estimates are updated in a Bayesian manner which extends the classical Kalman filter in a natural way, although the details do not concern us here.

IPDA is usually implemented with reference to a "gate", a hyperellipse surrounding the predicted target position outside of which measurements are rejected automatically as being unassociated with the target. The gate serves primarily to reduce the computational burden so, to reduce instead the burden on the statistical analysis, we will discuss a "gateless" version of IPDA here. Further, IPDA, as originally proposed, models the evolution of target existence between scans using a two- or three-state hidden Markov model. Practically, this means that a decaying or "forgetting" factor is also applied to the PoE from scan to scan. We neglect this aspect of IPDA.

At scan n, the PoE, r_n, is updated according to the formula

$$\frac{r_n}{1-r_n} = \left[1 - P_{\text{D}} + \frac{P_{\text{D}}}{\lambda}\sum_{m=1}^{f_n}\mathcal{N}(\mathbf{y}_{m,n}; \mathbf{H}\hat{\mathbf{x}}_{n|n-1}, \mathbf{S}_n)\right]\frac{r_{n-1}}{1-r_{n-1}}. \tag{25}$$

We introduce ϕ_n as a shorthand for the probability ratio $r_n/(1-r_n)$. We further adopt the suggestion of Musicki *et al.* that the spatial false-alarm rate λ be estimated directly from the number of detections in the scan. That is, at scan n, $\lambda_n = f_n/\operatorname{vol}\mathcal{S}$. This will be helpful in showing more clearly the relationship with the GLR of Section V-B.

We develop (25) so that

$$\begin{aligned}\frac{\phi_N}{\phi_0} &= \Big(\prod_{n\in\mathcal{J}}\lambda_n^{-1}\Big)\prod_{n\in\mathcal{J}}\sum_{m=1}^{f_n}\frac{(1-P_{\text{D}})}{\operatorname{vol}\mathcal{S}} + P_{\text{D}}\mathcal{N}(\mathbf{y}_{m,n}; \mathbf{H}\hat{\mathbf{x}}_{n|n-1}, \mathbf{S}_n)\\ &= \Big(\prod_{n\in\mathcal{J}}\lambda_n^{-1}\Big)\sum_{m_{n_1}=1}^{f_{n_1}}\cdots\sum_{m_{n_{|\mathcal{J}|}}=1}^{f_{n_{|\mathcal{J}|}}}\sum_{\mathcal{D}\in\mathbb{P}_{N,\mathcal{I}}}\\ &\qquad P_{\text{D}}^{|\mathcal{T}|}\left(\frac{1-P_{\text{D}}}{\operatorname{vol}\mathcal{S}}\right)^{|\mathcal{M}|}\prod_{n\in\mathcal{T}}\mathcal{N}(\mathbf{y}_{m_n,n}; \mathbf{H}\hat{\mathbf{x}}_{n|n-1}, \mathbf{S}_n)\end{aligned}$$

where $\mathcal{J}$ is the complement to $\mathcal{I}$ in $\{1,\ldots,N\}$ and $\mathbb{P}_{N,\mathcal{I}}$ is the restriction of $\mathbb{P}_N$ to elements that contain $\mathcal{I}$ as a subset. We define

$$\psi_{\mathcal{I}}(\mathbf{y}_{\mathcal{I}}) \triangleq P_{\text{D}}^{|\mathcal{I}|}\Big(\prod_{n\in\mathcal{J}}\lambda_n\Big)\left[\frac{\phi_N}{\phi_0} - (1-P_{\text{D}})^{|\mathcal{J}|}\right] \tag{26}$$

where we explicitly parametrise ψ by the set $\mathcal{I}$ of scans used for initialisation and by $\mathbf{y}_{\mathcal{I}}$, the initialising detections within those scans. Observe that we also subtract off the constant term that corresponds to a sequence consisting entirely of missed detections apart from the initialising set. This is in keeping with our convention from Section V-B. Then, we can show that

$$\psi_{\mathcal{I}}(\mathbf{y}_{\mathcal{I}}) = \sum_{\mathcal{D}\in\mathbb{P}_{N,\mathcal{I}}}\sum_{m_{t_1}=1}^{f_{t_1}}\cdots\sum_{m_{t_{|\mathcal{T}|}}=1}^{f_{t_{|\mathcal{T}|}}} G_{\mathcal{D}}(\mathbf{y}_{\mathcal{T}}) \tag{27}$$

where the $\mathbf{y}_{\mathcal{T}}$ are the sets of measurements implied by the summations over the $m_{t_1},\ldots,m_{t_{|\mathcal{T}|}}$ and, like (14),

$$G_{\mathcal{D}}(\mathbf{y}_{\mathcal{T}}) = P_{\text{D}}^{|\mathcal{D}|}\left(\frac{1-P_{\text{D}}}{\operatorname{vol}\mathcal{S}}\right)^{|\mathcal{M}|}\prod_{n\in\mathcal{M}} f_n \prod_{n\in\mathcal{T}}\mathcal{N}(\mathbf{y}_{n,\mathcal{T}}; \mathbf{H}\hat{\mathbf{x}}_{n|n-1}, \mathbf{S}_n).$$

Consider applying $\psi_{\mathcal{I}}$ to a threshold Θ under the null hypothesis in which the $\mathbf{y}_{m,n}$ are assumed to be uniformly distributed on the surveillance space. Using the same assumptions that we used to

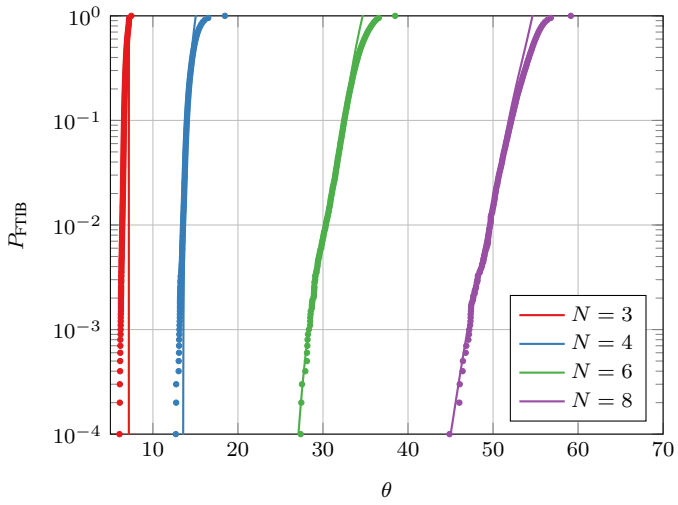

Fig. 3. Comparison of theoretical approximations and Monte-Carlo simulation of batch false-track probabilities in IPDA.

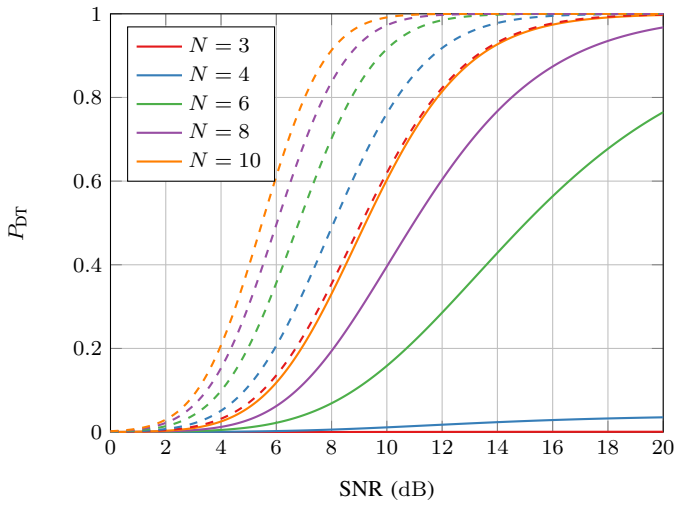

Fig. 4. TOC as a function of SNR when $t_{\text{FT}} = 90\,\text{s}$.

develop the approximation for P_{FT} in (18), namely that the sum can be approximated by the probability that any one of the terms exceeds the threshold and then applying the union bound, we then have that

$$P_{\text{FT};\mathcal{I}}(\mathbf{y}_{\mathcal{I}}) \triangleq \Pr\{\psi_{\mathcal{I}}(\mathbf{y}_{\mathcal{I}}) > \Theta\} \approx \sum_{m_{n_1}=1}^{f_{n_1}} \cdots \sum_{m_{n_{|\mathcal{T}|}}=1}^{f_{n_{|\mathcal{T}|}}} \sum_{\mathcal{D}\in\mathbb{P}_{N,\mathcal{I}}} \Pr\left\{G_{\mathcal{D}}(\mathbf{y}_{\mathcal{T}}) > \Theta\right\}.$$

With reference to (17), we have

$$\Pr\left\{G_{\mathcal{D}}(\mathbf{y}_{\mathcal{T}}) > \Theta\right\} \approx \frac{\text{vol}\{\mathcal{C}_{\mathcal{D}}[2(\theta-\epsilon_{\mathcal{D}})] \cap \mathcal{S}^{|\mathcal{D}|}\}}{\text{vol}\,\mathcal{S}^{|\mathcal{D}|}} \quad \text{where}$$

$$\epsilon_{\mathcal{D}} = \tfrac{1}{2}\sum_{n\in\mathcal{T}} \log(|2\pi\mathbf{S}_{n,\mathcal{D}}|) - |\mathcal{D}|\log P_{\text{D}} - |\mathcal{M}|\log[\lambda(1-P_{\text{D}})].$$

Within a batch of N consecutive scans, we compute $\psi_{\mathcal{I}}$ for all possible sets $\mathcal{I}$ of initiating scans and for all possible sensor-level detections within those scans. It follows that

$$\begin{aligned} P_{\text{FTIB}} &= 1 - E_{F_1,\ldots,F_N}\left[\prod_{\mathcal{I}\in\mathbb{I}_N}\prod_{m_{n_1}=1}^{F_{n_1}}\cdots\prod_{m_{n_I}}^{F_{n_I}}\left\{1-P_{\text{FT};\mathcal{I}}(\mathbf{y}_{\mathcal{I}})\right\}\right] \\ &\approx E_{F_1,\ldots,F_N}\left[\sum_{\mathcal{I}\in\mathbb{I}_N}\sum_{m_{n_1}=1}^{F_{n_1}}\cdots\sum_{m_{n_I}}^{F_{n_I}} P_{\text{FT};\mathcal{I}}(\mathbf{y}_{\mathcal{I}})\right] \\ &\approx \sum_{\mathcal{D}\in\mathbb{P}_{N,|\mathcal{I}|}} F^{|\mathcal{D}|}\frac{\text{vol}\{\mathcal{C}_{\mathcal{D}}[2(\theta-\epsilon_{\mathcal{D}})]\cap\mathcal{S}^{|\mathcal{D}|}\}}{\text{vol}\,\mathcal{S}^{|\mathcal{D}|}} \end{aligned} \tag{28}$$

where $\mathbb{I}_N$ denotes all subsets of $\{1,\ldots,N\}$ of length I.

Example. We plot, in Figure 3, the probability of false track within a batch for IPDA according to our approximation (28) in solid lines. The parameters are again those of the example scenario in Section III. Monte Carlo simulations, each involving 10,000 trials, are plotted for comparison using circular markers. Apart from $N = 3$ and except where P_{FTIB} is close to unity, the agreement is seen to be excellent.

The track detection probability in IPDA requires but little explanation. The same approach as was used to arrive at (23) may be used again for IPDA, substituting only $\epsilon_{\mathcal{D}_0}$ for $\delta_{\mathcal{D}_0}$.

VIII. Tracker Operating Characteristic

To illustrate the relative performance of the TkBD GLR statistic derived in Section V-A and the IPDA PoE, we conclude the paper by plotting the TOC for the example scenario of Section III.

In Figure 4, we plot the TOC for IPDA using solid lines. The theoretical approximations only are plotted. Using dashed lines, the TOC is plotted for the TkBD statistic. In all cases, we maintain $t_{\text{FT}} = 90\,\text{sec}$. Performance for $N = 3$ with the TkBD statistic is almost identical to that for $N = 10$ with IPDA. In terms of reflected energy, the gap between TkBD and IPDA narrows as N increases. Yet even at $N = 10$, the gap is $5.5\,\text{dB}$ when $P_{\text{DT}} = 0.9$. As has been widely reported, *e.g.*, [2] and references therein, there is a substantial advantage in using available amplitude or energy measurements in track confirmation. Using the contributions of this paper, we can now quantify the advantages using accurate closed-form approximations.

References

[1] D. Musicki, R. Evans, and S. Stankovic, "Integrated probabilistic data association," *IEEE Trans. Automat. Control*, vol. 39, no. 6, pp. 1237–1241, Jun. 1994.

[2] S. J. Davey, "SNR limits on Kalman filter detect-then-track," *IEEE Signal Process. Letters*, vol. 20, no. 8, pp. 767–770, Aug. 2013.

[3] D. Mušicki and R. Evans, "Data association and track initialisation," *IEEE Trans. Aerospace Electron. Syst.*, vol. 40, no. 2, pp. 387–398, Apr. 2004.

[4] T. L. Song and D. Mušicki, "Target existence based resource allocation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4496–4506, Sep. 2010.

[5] S. Schoenecker, T. Luginbuhl, P. Willett, and Y. Bar-Shalom, "Extreme-value analysis for ML-PMHT, Part 2: Target trackability," *IEEE Trans. Aerospace Electron. Syst.*, vol. 50, no. 4, pp. 2515–2527, Oct. 2014.

[6] Y. Bar-Shalom, L. J. Campo, and P. B. Luh, "From receiver operating characteristic to system operating characteristic: Evaluation of a track formation system," *IEEE Trans. Automat. Control*, vol. 35, no. 2, pp. 172–179, Feb. 1990.

[7] C. Jauffret and Y. Bar-Shalom, "Track formation with bearing and frequency measurements in clutter," *IEEE Trans. Aerospace Electron. Syst.*, vol. 26, no. 6, pp. 999–1010, Nov. 1990.

[8] T. Kirubarajan and Y. Bar-Shalom, "Low observable target motion analysis using amplitude information," *IEEE Trans. Aerospace Electron. Syst.*, vol. 32, no. 4, pp. 1367–1384, Oct. 1996.

[9] W. R. Blanding, P. K. Willett, and Y. Bar-Shalom, "Offline and real-time methods for ML-PDA track validation," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1994–2006, May 2007.