An MDL Approach for Multiple Low Observable Track Initiation*

H. Chen[a], T. Kirubarajan[b], Y. Bar-Shalom[a], and K. R. Pattipati[a]

[a]University of Connecticut
Dept. of Electrical & Computer Engineering
Storrs, CT 06269-2157, USA

[b]McMaster University
Dept. of Electrical & Computer Engineering
Hamilton, Ontario, Canada

ABSTRACT

In this paper the track initiation problem is formulated as multiple composite hypothesis testing using maximum likelihood estimation with probabilistic data association (ML-PDA), an algorithm known to work under very low SNR conditions. This algorithm does not have to make a decision as to which measurement is target originated. The hypothesis testing is based on the minimum description length (MDL) criterion. We first review some well-known approaches for statistical model selection and the advantage of the MDL criterion. Then we present an approximate penalty in accounting for the model complexity to simplify the calculation of MDL. Finally, we apply the MDL approach for the detection and initiation of tracks of incoming tactical ballistic missiles in the exo-atmospheric phase using a surface based electronically scanned array (ESA) radar. The targets are characterized by low SNR, which leads to low detection probability and high false alarm rate. The target acquisition problem is formulated using a batch of radar scans to detect the presence of up to two targets. The ML-PDA estimator is used to initiate the tracks assuming the target trajectories follow a deterministic state propagation. The approximate MDL criterion is used to determine the number of valid tracks in a surveillance region. The detector and estimator are shown to be effective even at 4.4 dB average SNR in a resolution cell (at the end of the signal processing chain).

Keywords: Track Initiation, Minimum Description Length, Maximum Likelihood Estimator, Probabilistic Data Association, Model Selection

1. INTRODUCTION

A substantial threat comes from theater ballistic missiles (TBM). This has stimulated extensive research work in the area of ballistic missile defense (BMD). The first step is to detect the incoming missiles, then track their trajectories and eventually direct interceptors or other defense apparatus to counter them. The acquisition of a missile consists of missile track detection and its initial state (position and velocity) estimation at a certain time instant. During its exo-atmospheric flight, the predorminant force acting on the missile is earth's gravity, which yields a determinstic target trajectory given by the laws of Kepler [3]. The estimation of the exo-atmospheric trajectory is equivalent to the estimation of the target state at some reference time t_0, i.e., it amounts to the estimation of a fixed dimensional parameter vector. The measurements are obtained by a ground-based radar searching for the missile in a predetermined "cueing region". The radar is a phased array monopulse radar and the cueing region is informed by early warning satellites. The radar uses amplitude comparison monopulse processing to obtain the angle measurements. The technical challenges in this estimation problem consist of short time period available to acquire the missile tracks, presence of spurious measurements due to the low detection threshold made necessary by the desire to extend the acquisition range of the targets, and most of all, unknown number of targets. The problem can be formally written as multiple composite hypothesis testing where the number of candidate hypotheses is yet to be determined. From a statistical standpoint, it is a model selection problem where each model represents a different number of tracks to explain the observed data.

Previous work [17] proposed the track initiation algorithm of a ballistic missile using line of sight (LOS) measurements. The monopulse radar model developed in [4] has been widely used in the benchmark tracking problems in [8], [16] and [13] with SNR of at least 10 dB. Few results are available for the estimation of missile trajectories under lower SNR. We use a monopulse radar to measure the full position of the incoming target and want to extend the missile acquisition range, which requires acquisition capability of low SNR targets. A more recent paper [15] developed a track initiation algorithm for monopulse radar to work under average SNR as low as 4 dB. The algorithm

*Research sponsored by ONR/BMDO Grant N00014-00-1-0740, AFOSR Grant F49620-00-1-0052 and ONR Grant N00014-97-1-0502.

used in [15] was the maximum likelihood estimator with probabilistic data association (ML-PDA), which does not have to make a decision as to which measurement is target originated. However, the above works are applicable only for acquisition of a single target. The acquisition of multiple targets in the presence of high false alarm rate and missed detections is the topic of this paper.

The need for model selection arises when we want to determine the number of tracks using the ML-PDA based on the available data. There exist several criteria for model selection, namely, the Akaike information criterion (AIC) [1], the Bayes information criterion (BIC) [12], stochastic complexity (SC) [9], generalized ML rule [7], the information-theoretic measure of complexity (ICOMP) [5] and minimum description length (MDL) [10]. AIC and BIC are the most commonly used methods which penalize a complicated model by some terms depending on the number of unknown parameters being determined from the data. However, they are not invariant under reparameterization. The only approach invariant under reparameterization is the MDL criterion, which compares different model classes by the likelihood function augmented with the model complexity [11]. The idea behind MDL is the notion that knowledge (some regularity property in the data) and data redundancy are interrelated. There must be some kind of redundancy in the data because without it, every point in the data will be unique. In such a case, there is no regularity to be learned or extracted from the data. In other words, the more we compress the data by extracting redundancy from it, the more we learn about the regularity underlying the data. The model with minimum redundancy would be the best choice. With the MDL criterion, we can gain some insights on the performance limit of the ML-PDA for initiating multiple tracks.

The problem of testing multiple composite hypotheses is formulated in Section 2. The application of the MDL criterion to the missile track acquisition is presented in Section 3 with the simulation results based on scenarios with up to two targets under various SNRs. Section 4 summarizes the results.

2. MULTIPLE COMPOSITE HYPOTHESIS TESTING AND MDL CRITERION

The multiple composite hypothesis testing can be formulated as follows. Assume we have K different models for hypothesis testing, namely, H_1, H_2, ..., H_K. Denote $\mathbf{Z} = \{\mathbf{z}_1, ..., \mathbf{z}_N\}$ as the data sequence of length N. Under hypothesis H_i, we have the conditional probability density function (pdf) $p(\mathbf{Z}|\mathbf{x}_i, H_i)$ with unknown parameter vector $\mathbf{x}_i$ of dimension n_i. Hence we may equate a model with a parametric pdf $p(\mathbf{Z}|\mathbf{x}_i, H_i)$. Here the conditioning on H_i can not be omitted since different models may have different functional forms of the conditional pdf. Assuming the priors $P(H_i)$ $(i = 1, ..., K)$ are equal, we want to choose model j for which $p(\mathbf{Z}|H_j)$ is maximum. Since the parameter $\mathbf{x}_i$ is unknown, we assign a prior pdf $p(\mathbf{x}_i)$ for each hypothesis H_i. The likelihood function of H_i is

$$p(\mathbf{Z}|H_i) = \int p(\mathbf{Z}|\mathbf{x}_i, H_i) p(\mathbf{x}_i) d\mathbf{x}_i \tag{1}$$

Due to the difficulty in performing the integration and the lack of prior knowledge, this approach is not very practical. An alternative is to use the maximum likelihood (ML) estimate $\hat{\mathbf{x}}_i$ for the unknown $\mathbf{x}_i$ and choose H_j for which $p(\mathbf{Z}|\hat{\mathbf{x}}_j, H_j)$ is maximum, i.e.,

$$j = \arg\max_i p(\mathbf{Z}|\hat{\mathbf{x}}_i, H_i) \tag{2}$$

where

$$\hat{\mathbf{x}}_i = \arg\max_{\mathbf{x}_i} p(\mathbf{Z}|\mathbf{x}_i, H_i), \quad i = 1, ..., K \tag{3}$$

This is the generalized likelihood ratio test (GLRT) for multiple composite hypothesis testing. However, the above decision rule tends to choose a more sophisticated model as the number of data points increases because the modeling error decreases as we add more unknown parameters to the model. To compare models with different parameters (as well as different functional forms), we have to penalize a sophisticated model which overfits the limited amount of data.

From an information theoretic point of view, we prefer to choose a model that has minimum description length (MDL) in explaining the data. In [10], the model complexity is measured by the normalized maximum likelihood (NML), i.e.,

$$\hat{p}(\mathbf{Z}|H_i) = \frac{p(\mathbf{Z}|\hat{\mathbf{x}}_i, H_i)}{\int_{\hat{\mathbf{x}}_i(\mathbf{Z}) \in \Omega_i} p(\mathbf{Z}|\hat{\mathbf{x}}_i, H_i) d\mathbf{Z}} \tag{4}$$

where $\mathbf{Z}$ is a data sequence of length N that yields the MLE $\hat{\mathbf{x}}_i$ and Ω_i is a subset of the parameter space for H_i that makes the integral finite.[1] It is shown in [10] that under some mild conditions, the ideal code length of the NML is

$$-\ln\hat{p}(\mathbf{Z}|H_i) = -\ln p(\mathbf{Z}|\hat{\mathbf{x}}_i, H_i) + \frac{n_i}{2}\ln\left(\frac{N}{2\pi}\right) + \ln\int_{\Omega_i} |I(\mathbf{x}_i)|^{\frac{1}{2}} d\mathbf{x}_i + o(1) \tag{5}$$

Here $I(\mathbf{x}_i)$ is the Fisher information matrix (FIM) per data point given by

$$I(\mathbf{x}_i) = \lim_{N\to\infty} \frac{1}{N} E\left(\left[\nabla_{\mathbf{x}_i} \ln p(\mathbf{Z}|\mathbf{x}_i, H_i)\right] \left[\nabla_{\mathbf{x}_i} \ln p(\mathbf{Z}|\mathbf{x}_i, H_i)\right]' \right) \tag{6}$$

where

$$\nabla_{\mathbf{x}_i} \ln p(\mathbf{Z}|\mathbf{x}_i, H_i) = \left[\frac{\partial \ln p(\mathbf{Z}|\mathbf{x}_i, H_i)}{\partial (\mathbf{x}_i)_1} \cdots \frac{\partial \ln p(\mathbf{Z}|\mathbf{x}_i, H_i)}{\partial (\mathbf{x}_i)_{n_i}} \right]' \tag{7}$$

and $o(1)$ converges to zero as $N \to \infty$.

The MDL model selection criterion can be equivalently written as choosing model j as

$$j = \arg\max_i \left\{ \ln p(\mathbf{Z}|\hat{\mathbf{x}}_i, H_i) - \frac{n_i}{2}\ln\left(\frac{N}{2\pi}\right) - \ln\int_{\Omega_i} |I(\mathbf{x}_i)|^{\frac{1}{2}} d\mathbf{x}_i \right\}, \quad i = 1, ..., K \tag{8}$$

Note that the first term in the right hand side (r.h.s.) of (8) is the log-likelihood function. The second term in the r.h.s. of (8) is a penalty for the number of unknown parameters used to fit the data. The third term in the r.h.s. of (8) is a penalty for the functional complexity measured over the subset Ω_i. In general, it is difficult to compute the third penalty term. In the AIC [1] the penalty term n_i is used while in the BIC [12] the penalty term $\frac{n_i}{2}\log N$ is used. They all ignore the differences in model complexity in different functional forms. In [7], the penalty as a function of the FIM is used rather than as a function of the number of parameters, which again can only choose the correct model in an asymptotic sense. The MDL criterion is an invariant measure for model selection [10]. It has the advantage of not favoring a particularly complex model, and is suitabe for the multiple composite hypothesis testing problem considered here.

3. APPLICATION: ACQUISITION OF BALLISTIC MISSILES

The main focus of the ballistic missile acquisition is to initialize the missile trajectories using the measurements obtained by a surface-based radar for a short period of time (typically less than 10 s) before the missile reentry phase. The radar beam pointing directions are assumed to cover a specific "cueing region" (taken here as $(30\,\text{km})^3$) provided by some early warning. The measurements obtained by the radar require a monopulse processing technique to obtain the full position measurements with associated variances. Due to low SNR and thus the low detection threshold to obtain a reasonable detection probability (around 0.6 for a dwell), the set of measurements contains a large number of false alarms (detection threshold exceedances due to noise). In addition, the number of targets within the cueing region is unknown. A track initiation algorithm has to determine how many targets are in the cueing region (model selection problem). If there are targets, the initial state of each target has to be estimated. In the following subsections, we present the target and radar model and the ML-PDA algorithm for missile trajectory parameter estimation. The detailed missile trajectory, radar model and the monopulse processing technique can be found in [6].

3.1. Coordinate Systems and Target Trajectory

The motion of a missile above the atmosphere is governed by Kepler's laws. The earth centered inertial (ECI) coordinate system is used for the state propagation of the targets. This coordinate system has its origin at the center of the earth, with the positive x axis pointing along the vernal equinox and the positive z axis to the north pole. The radar and target locations are better interpreted in the earth centered earth fixed (ECEF) coordinate system. This coordinate system has its origin at the center of the earth, with the positive x axis passing through the prime meridian at the equator and the positive z axis passing through the north pole. Let $\mathbf{x}_k = [\mathbf{r}_k \;\; \dot{\mathbf{r}}_k]'$ be the

[1] The integral is finite if Ω_i is an open bounded set and its closure contains no singular point of the Fisher information matrix.

6-dimensional target state vector at time t_k, where the vectors $\mathbf{r}_k = [\xi(t_k)\ \ \eta(t_k)\ \ \zeta(t_k)]'$ and $\dot{\mathbf{r}}_k = [\dot{\xi}(t_k)\ \ \dot{\eta}(t_k)\ \ \dot{\zeta}(t_k)]'$ are the position and velocity of the target in ECI coordinates, respectively. Given the initial state $\mathbf{x}_0$ of the target at time t_0, the state at t_k is

$$\mathbf{x}_k = [\mathbf{r}_k(\mathbf{x}_0, t_0, t_k)'\ \ \dot{\mathbf{r}}_k(\mathbf{x}_0, t_0, t_k)']' \tag{9}$$

where $\mathbf{r}_i(\cdot)$ and $\dot{\mathbf{r}}_i(\cdot)$ are determinstic functions given in [6]. Another coordinate system is located at the radar. This coordinate has its origin at the radar, with positive x pointing toward north and positive z axis pointing up. It is often called the local Cartesian frame of radar or ENU frame indicating the axes pointing toward East, North and Up. Given the radar position $\mathbf{O}^{radar}$ in ECEF coordinates, the state of the target in radar coordinates is given by

$$\mathbf{x}_k^{radar} = \text{AXIS}(\mathbf{x}_k, \mathbf{O}^{radar}) \tag{10}$$

where AXIS$(\cdot)$ is the coordinate conversion given in [6]. Figure 1 illustrates the missile trajectory and the cueing region. The state estimate of the target is obtained in radar coordinates and then converted into ECEF coordinates.

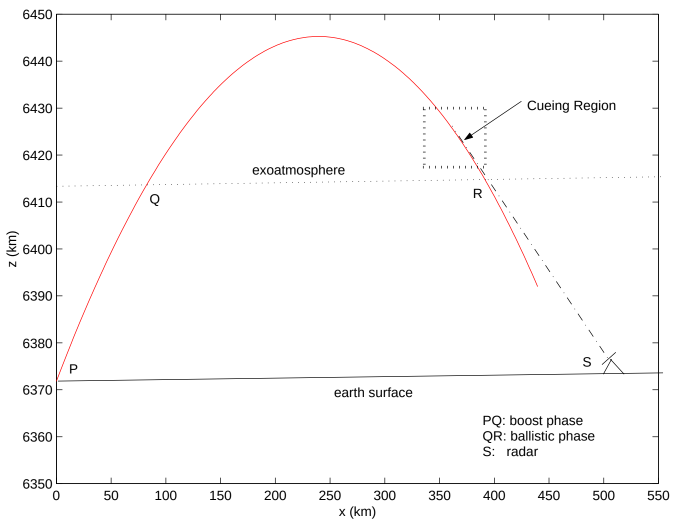

Figure 1. Illustration of the stages of a ballistic missile trajectory

3.2. The Log-Likelihood Ratio with Probabilistic Data Association

We now present the maximum likelihood (ML) estimator combined with the probabilistic data association (PDA) technique to tackle the uncertainty of measurement origin. We assume that the radar receiver waits for N scans before a decision about track existence is made. Thus we have N sets of measurements made at times t_i, $i = 1, 2, ..., N$ where the index i denotes the scan number. The k-th measurement in the i-th scan is defined by the vector

$$z_{ik} = [\Re_{ik}\ \ r_{ik}\ \ \beta_{ik}\ \ \epsilon_{ik}] \tag{11}$$

where $\Re_{ik}$, r_{ik}, β_{ik}, ϵ_{ik} are the signal plus noise to noise ratio (SNNR), slant range, bearing and elevation, respectively. The SNNR is a target feature and it has been shown to reduce the data association uncertainty [15]. Its model is presented in sequel. The radar performs a volumetric search over the cueing region based on prior information. It is assumed that all the dwells in a scan are made at (practically) the same time. Denote by m_i the number of measurements in the i-th scan obtained at time t_i. The set is written as

$$Z(i) = \{z_{ik}\}_{k=1}^{m_i} \tag{12}$$

The cumulative set of measurements during the entire period is

$$\mathbf{Z} = \{Z(i)\}_{i=1}^{N} \tag{13}$$

The following assumptions about the statistical characteristics of the measurements are made:

1. The measurements at two different scans are conditionally independent, i.e.,

$$p[Z(i), Z(j)|\mathbf{x}] = p[Z(i)|\mathbf{x}]p[Z(j)|\mathbf{x}] \tag{14}$$

2. A measurement that originated from a target at a particular scan is given by

$$r_{ik} = r_0 + (n + \frac{1}{2})\Delta_r, \quad \text{if } r_0 + n\Delta_r < r_i(\mathbf{x}_i) < r_0 + (n+1)\Delta_r \tag{15}$$

$$\beta_{ik} = \beta_i(\mathbf{x}_i) + w_{ik}^{\beta} \tag{16}$$

$$\epsilon_{ik} = \epsilon_i(\mathbf{x}_i) + w_{ik}^{\epsilon} \tag{17}$$

where r_0 is the reference range and n is the index of the range bins (range cells) of size Δ_r;

$$w_{ik}^{\beta} \sim \mathcal{N}(0, (\sigma_{ik}^{\beta})^2) \tag{18}$$

$$w_{ik}^{\epsilon} \sim \mathcal{N}(0, (\sigma_{ik}^{\epsilon})^2) \tag{19}$$

and the bearing and elevation measurement noise standard deviations σ_{ik}^{β} and σ_{ik}^{ϵ} are obtained as in [6].

Next we present the assumptions made for the statistical characteristics of false alarms.

1. In the case of a false alarm, the bearing and elevation measurements are Gaussian distributed around the commanded bearing and elevation (beam boresight), respectively. That is, if a measurement z_{ik} is a false alarm, we have (if SNR in range bin n exceeds the detection threshold)

$$r_{ik} = r_0 + (n + \frac{1}{2})\Delta_r \tag{20}$$

$$\beta_{ik} \sim \mathcal{N}(b_{ik}, (\sigma_f^{\beta})^2) \tag{21}$$

$$\epsilon_{ik} \sim \mathcal{N}(e_{ik}, (\sigma_f^{\epsilon})^2) \tag{22}$$

where b_{ik} and e_{ik} are the bearing and elevation of the beam boresight for the k-th measurement in the i-th radar scan; σ_f^{β} and σ_f^{ϵ} are the standard deviations which can be obtained for various detection thresholds.[2]

2. The number of false measurements in a volume V_g (measurement validation gate around the hypothesized trajectory) at each scan follows a Poisson probability mass function (pmf) with known expected number of false measurements λ per unit volume. That is, the probability that there are m_i false measurements in a volume V_g is given by

$$\mu(m_i) = \frac{e^{-\lambda V_g}(\lambda V_g)^{m_i}}{m_i!} \tag{23}$$

where the parameter λ is the ratio of the false alarm probability in a resolution cell to the cell volume.

If there are m_i measurements in the i-th scan, we assume at most one measurement originates from the target. Hence the mutually exclusive and exhaustive events giving rise to these measurements are

$$\chi_k(i) = \begin{cases} \text{all measurements are false} & k = 0 \\ \text{measurement } z_{ik} \text{ is from the target} & k = 1, ..., m_i \end{cases} \tag{24}$$

This assumption may not hold exactly since a target may have multiple detections from several adjacent beams in one scan. In this case, only the detection from the main beam is considered as target originated measurement and the rest of them are taken as false alarms. Denote by $p_0^{\tau}(\Re_{ij})$, $p_0(r_{ij})$, $p_0(\beta_{ij})$ and $p_0(\epsilon_{ij})$ the probability density functions of SNNR, slant range, bearing and elevation of the false measurement z_{ij} with the detection threshold τ, respectively. We have [6]

$$p_0^{\tau}(\Re_{ij}) = \frac{1}{P_{FA}} p_0(\Re_{ij}), \quad \Re_{ij} > \tau \tag{25}$$

[2] This is the consequence of the "centralilty tedency" of the false measurements, discussed in [15].

Denote by $p_1^\tau(\Re_{ij}|\mathbf{x}_i)$, $p_1(r_{ij}|\mathbf{x}_i)$, $p_1(\beta_{ij}|\mathbf{x}_i)$ and $p_1(\epsilon_{ij}|\mathbf{x}_i)$ the probability density functions of SNNR, slant range, bearing and elevation from the target originated measurement z_{ij} at state $\mathbf{x}_i$ (the full position at reference time t_i) with the detection threshold τ, respectively. We have [6]

$$p_1^\tau(\Re_{ij}|\mathbf{x}_i) = \frac{1}{P_D} p(\Re_{ij}|\mathbf{x}_i), \quad \Re_{ij} > \tau \tag{26}$$

Under H_0, the pdf of the measurements $Z(i)$ corresponding to the hypothesis that all measurements are false is given by

$$p_0(Z(i)|\delta_0(i)) = \prod_{j=1}^{m_i} p_0^\tau(\Re_{ij}) p_0(r_{ij}) p_0(\beta_{ij}) p_0(\epsilon_{ij}) \tag{27}$$

Under H_1, the pdf of the measurements $Z(i)$ if the k-th measurement originates from the target is given by

$$p_1(Z(i)|\mathbf{x}_i, \delta_k(i)) = p_1^\tau(\Re_{ik}|\mathbf{x}_i) p_1(r_{ik}|\mathbf{x}_i) p_1(\beta_{ik}|\mathbf{x}_i) p_1(\epsilon_{ik}|\mathbf{x}_i) \prod_{j=1, j\neq k}^{m_i} p_0^\tau(\Re_{ij}) p_0(r_{ij}) p_0(\beta_{ij}) p_0(\epsilon_{ij}) \tag{28}$$

where $\mathbf{x}_i$ is the target position at time t_i. From (25), (26) the amplitude likelihood ratio is [6]

$$\rho_{ik} = \frac{p_1^\tau(\Re_{ik}|\mathbf{x}_i)}{p_0^\tau(\Re_{ik})} = \zeta_i(1 + \nu_i \Re_{ik}) e^{\nu_i \Re_{ik}} \tag{29}$$

and the likelihood ratio of the position component between a target-originated and a false measurement is

$$\frac{p_1(r_{ik}|\mathbf{x}_i) p_1(\beta_{ik}|\mathbf{x}_i) p_1(\epsilon_{ik}|\mathbf{x}_i)}{p_0(r_{ik}) p_0(\beta_{ik}) p_0(\epsilon_{ik})} = \frac{\sigma_f^\beta \sigma_f^\epsilon}{\sigma_{ik}^\beta \sigma_{ik}^\epsilon} \exp\{\psi(z_{ik}, \mathbf{x}_i)\} \tag{30}$$

where, with $\alpha_i \Sigma_i^4$ denoting the average SNR (multiplied by 2)[3],

$$\zeta_i = \frac{16 P_{FA}}{P_D (4 + \alpha_i \Sigma_i^4)^2} \tag{31}$$

$$\nu_i = \frac{\alpha_i \Sigma_i^4}{4 + \alpha_i \Sigma_i^4} \tag{32}$$

$$\psi(z_{ik}, \mathbf{x}_i) = -\frac{[\beta_{ik} - \beta_i(\mathbf{x}_i)]^2}{2(\sigma_{ik}^\beta)^2} - \frac{[\epsilon_{ik} - \epsilon_i(\mathbf{x}_i)]^2}{2(\sigma_{ik}^\epsilon)^2} + \frac{(\beta_{ik} - b_{ik})^2}{2(\sigma_f^\beta)^2} + \frac{(\epsilon_{ik} - e_{ik})^2}{2(\sigma_f^\epsilon)^2} \tag{33}$$

Denote by $p(Z(i)|\mathbf{x}_i)$ the likelihood function that one of the measurements in $Z(i)$ is target originated with target position $\mathbf{x}_i$. Similarly, let $p(Z(i)|\delta_0(i))$ be the likelihood function that all measurements are false. The likelihood ratio is

$$\phi(Z(i), \mathbf{x}_i) = \frac{p(Z(i)|\mathbf{x}_i)}{p(Z(i)|\delta_0(i))} \tag{34}$$

Assuming the prior probabilities of each measurement to be target originated are equal, the likelihood ratio of one target present versus no target is, using the PDA approach, given by [15]

$$\phi(Z(i), \mathbf{x}_i) = \frac{1}{C_i} \left\{ \frac{P_D}{\lambda V_g} \sum_{k=1}^{m_i} \frac{\rho_{ik} \sigma_f^\beta \sigma_f^\epsilon}{\sigma_{ik}^\beta \sigma_{ik}^\epsilon} \exp(\psi(z_{ik}, \mathbf{x}_i)) + (1 - P_D) \right\} \tag{35}$$

where C_i is a normalizing constant. Using the conditional independence of the measurements among different scans, the likelihood ratio of the entire measurement set can be written as

$$\Lambda(\mathbf{Z}, \mathbf{x}) = \prod_{i=1}^{N} \phi(Z(i), \mathbf{x}_i) \tag{36}$$

[3] The SNR is the ratio of the total target power to the total noise power in the I and Q radar receiver channel, with each of them normalized to unity. Thus SNR=$\alpha_i \Sigma_i^4/2$ and $\alpha_i \Sigma_i^4$ is the normalized signal power (average target return) as it passes through the antenna pattern. For more details, see [4,15].

The log-likelihood ratio is given by

$$\ln(\Lambda(\mathbf{Z},\mathbf{x})) = \sum_{i=1}^{N} \ln\left\{ \frac{P_D}{\lambda V_g} \sum_{k=1}^{m_i} \frac{\rho_{ik}\sigma_f^\beta \sigma_f^\epsilon}{\sigma_{ik}^\beta \sigma_{ik}^\epsilon} \exp(\psi(z_{ik},\mathbf{x}_i)) + (1-P_D) \right\} - \sum_{i=1}^{N} \ln c_i \tag{37}$$

where c_i is a constant given by

$$c_i = \begin{cases} 1 & m_i = 0 \\ (1-P_D) + \frac{m_i}{\lambda V_g} P_D & m_i > 0 \end{cases} \tag{38}$$

Note that (37) includes all the measurements without any decision as which are target originated. This is the essence of the ML-PDA approach. In (37) the second summation involving c_i can be omitted when finding the maximum likelihood estimate. The maximization of the log-likelihood ratio was accomplished in [15] using the quasi-Newton method following a preliminary grid search. Here, instead of evaluating the log-likelihood ratio with a grid search, we use a less expensive grid search based on the number of nonempty gates. This will speed up finding the initial point before using the quasi-Newton method.

3.3. Acceptance of the Estimate

We first consider the hypothesis test of one target vs. no target:

- H_1: There is one track corresponding to an existing target and $\hat{\mathbf{x}}$ is the global maximum of the likelihood ratio
- H_0: $\hat{\mathbf{x}}$ does not correspond to a valid track

We use the log-likelihood ratio $\log(\Lambda(\mathbf{Z},\mathbf{x}))$ with the penalty terms given by the approximate MDL criterion as the test statistic for the hypothesis test of H_1 against H_0. Denote $T(\mathbf{Z})$ as the test statistic. We have

$$T(\mathbf{Z}) = \ln\Lambda(\mathbf{Z},\hat{\mathbf{x}}) - \frac{k}{2}\ln\left(\frac{S_N}{2\pi}\right) - \ln\int_\Omega |I(\mathbf{x})|^{\frac{1}{2}} d\mathbf{x} \tag{39}$$

where S_N is the number of measurements given by

$$S_N = \sum_{i=1}^{N} m_i \tag{40}$$

and k is the the dimension of the MLE $\hat{\mathbf{x}}$ (in this case $k=6$) which is

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} \ln\Lambda(\mathbf{Z},\mathbf{x}) \tag{41}$$

The parameter space Ω covers the cueing region and all possible velocity vectors of the missile. Without considering the information reduction factor, the integral $\int_\Omega |I(\mathbf{x})|^{\frac{1}{2}} d\mathbf{x}$ is obtained numerically and can be computed off-line. We choose H_1 when $T(\mathbf{Z}) > 0$. Alternatively, we can use $\log\Lambda(\mathbf{Z},\hat{\mathbf{x}})$ as the test statistic and compare it with a prespecified threshold for certain power based on the Gaussian approximation under H_1. The numerical approach for obtaining the mean and variance of the test statistic under H_1 can be found in [15]. We will compare the MDL with the GLRT in simulation.

If a track is accepted, the MLE $\hat{\mathbf{x}}$ is used as the state estimate with the corresponding covariance given by approximate Cramer-Rao lower bound (CRLB) which will be presented in the next subsection. The estimator uses all the measurements $\mathbf{Z}$ to extract one track at a time. The measurements having the association probability greater than 0.5 to the presumed track at each scan are deleted after the track has been initiated. If a measurement at a specific scan is deleted, all the measurements from its adjacent beams (in the same range bin) are also deleted (if there is any) to compensate for the effect of multiple detections. It is very important to remove all measurements from the main beam and the adjacent beams after a track is declared. The number of false tracks decreases when using this procedure especially when the SNR is low. The rest of the measurements are used to extract another target using the same ML-PDA estimator with MDL model selection criterion. The procedure continues until the test statistic $T(\mathbf{Z})$ is less than 0.

3.4. Cramer-Rao Lower Bound of the Estimator

The Cramer-Rao lower bound (CRLB) is a lower limit on the variance that can be achieved by an unbiased estimator. The MLE can be said to be efficient if its sample variance is statistically commensurate with the CRLB. If an MLE exists, its variance achieves the CRLB asymptotically. Thus for the MLE $\hat{\mathbf{x}}$ of the state $\mathbf{x}$, we have

$$E[(\mathbf{x}-\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})'] \geq J^{-1} \tag{42}$$

where J is the Fisher information matrix (FIM). If the noises of target originated measurements are Gaussian, the FIM can be written as

$$\begin{aligned} J \;=\; & \textstyle\sum_{i=1}^{N} q_i(P_D, \lambda, V_g, \alpha_i \Sigma_i^4) \Big(\frac{1}{(\sigma^r)^2} [\nabla_{\mathbf{x}} r_i(\mathbf{x})][\nabla_{\mathbf{x}} r_i(\mathbf{x})]' \\ & + \frac{1}{(\sigma^\beta)^2} [\nabla_{\mathbf{x}} \beta_i(\mathbf{x})][\nabla_{\mathbf{x}} \beta_i(\mathbf{x})]' + \frac{1}{(\sigma^\epsilon)^2} [\nabla_{\mathbf{x}} \epsilon_i(\mathbf{x})][\nabla_{\mathbf{x}} \epsilon_i(\mathbf{x})]' \Big) \end{aligned} \tag{43}$$

where $q_i(\cdot)$ is the information reduction factor for the i^{th} scan that accounts for the loss of information due to the presence of false alarms and the less-than-unity probability of detection [15]. However, the noise of the range measurement is not Gaussian. We use a moment matching technique and let

$$\sigma^r = \left(\frac{(2\Delta_r)^2}{12} \right)^{\frac{1}{2}} = 51.96 \tag{44}$$

To simplify the calculation, we also assume that the information reduction factor is a constant across different scans. Only those scans with at least one measurement in the gate are considered in obtaining the FIM. Thus we can write the FIM as

$$\begin{aligned} J \;=\; & q(P_D, \lambda V_g, g, \alpha \Sigma^4) \textstyle\sum_{i=1}^{N} \Big(\frac{1}{(\sigma^r)^2} [\nabla_{\mathbf{x}} r_i(\mathbf{x})][\nabla_{\mathbf{x}} r_i(\mathbf{x})]' \\ & + \frac{1}{(\sigma^\beta)^2} [\nabla_{\mathbf{x}} \beta_i(\mathbf{x})][\nabla_{\mathbf{x}} \beta_i(\mathbf{x})]' + \frac{1}{(\sigma^\epsilon)^2} [\nabla_{\mathbf{x}} \epsilon_i(\mathbf{x})][\nabla_{\mathbf{x}} \epsilon_i(\mathbf{x})]' \Big) \end{aligned} \tag{45}$$

The standard deviations of the bearing and elevation measurements depend on the target position within the radar beam. Assuming the target is uniformly distributed within the rectangular region covered by one radar beam, the average measurement noise for a target is $\sigma^\beta = \sigma^\epsilon = 0.0067$rad when the detection threshold is 3.25. If the detection probability per scan is $P_D = P_{D\text{ave}} = 0.464$, we have the average target return $\alpha\Sigma^4 = 5.76$. The false alarm probability is obtained by averaging the false detections in eight adjacent beams when the target is uniformly distributed within the rectangle area of the main beam (the one pointing to the target). Using the integral approach developed in [15], the information reduction factor is $q(\cdot) = 0.346$ for a single target with Poisson distributed false alarms. This information reduction factor $q(\cdot)$ is averaged over 100 randomly generated target positions uniformly distributed within the rectangular region of the main beam. It approximately quantifies the estimation accuracy. For multiple target scenarios, we assume the approximate CRLB for a single target is still valid for quantifying the efficiency of the estimator. For Swerling III targets with SNR=6 dB at boresight and various detection thresholds, we list the detection probability, false alarm probability and the information reduction factor averaged over the above rectangular region in Table 1.

3.5. Simulation Results

In this section we consider a two-target scenario to illustrate the operation of the ML-PDA estimator under various SNRs. The scenario contains two Swerling III targets with the same average RCS. We assume that the two targets enter the radar acquisition region at $t = 0$ s with the initial state $\mathbf{r}_0^{1true} = [720.0 \;\; 6050.0 \;\; 2180.0]'$ km, $\dot{\mathbf{r}}_0^{1true} = [-0.045 \;\; -1.531 \;\; -0.694]'$ km/s, and $\mathbf{r}_0^{2true} = [718.0 \;\; 6052.0 \;\; 2179.5]'$ km, $\dot{\mathbf{r}}_0^{2true} = [-0.045 \;\; -1.541 \;\; -0.695]'$ km/s; the radar position is $\mathbf{O}^{radar} = [800 \;\; 6000 \;\; 2000]'$ km, in ECEF coordinates. The radar parameters are given in Table 2. At this point the slant range of the radar to target 1 is 204.1 km and the slant range of radar to target 2 is 203.2 km.

It is assumed that the radar cueing region for initial radar beam pointing is available from prior information. The radar cueing volume is approximately $(30\,\text{km})^3$ and the radar beam packing is 7 by 7 rectangular with each dwell consisting 330 bins of length 90 m. The average output SNR at boresight for each target is 6 dB to 8 dB for the cases considered. Notice that the *average SNR in a resolution cell* is 4.4 dB when the *target SNR at boresight* is 6 dB. The detection threshold for acquisition of 6 dB targets is $\tau = 2.87$, yielding an average detection probability 0.52. The spatial density of false alarms in the validation gate is dorminated by the target-due extraneous detections in the eight adjacent beams when the target is uniformly distributed within the rectangular region centered at the main beam.[4] With the parameter settings specified as above we have λV_g=1.2. The radar scan rate is 10 Hz. The total number of scans N varies from 30 to 60. The scenarios are denoted as H_0: no target is present; H_1: only target 1 is present; H_2: both target 1 and target 2 are present. We are interested in the probability of choosing the correct model under each case. The target SNR at boresight varies from 6 dB to 8 dB. In each case, the detection threshold is chosen to maintain $\lambda V_g \leq 1.2$.

The results from 100 Monte Carlo runs for three scenarios (H_0, H_1, H_2) are listed in Table 3. When two tracks are initiated, the track to truth association is based on the geometric distance between the true and estimated state. If the distance is greater than three times the measurement standard deviation, the track is declared as false. If we want the probability of choosing the correct model to be greater than 0.95 for $N = 60$, the target SNR has to be

[4] These unavoidable extra detections cannot be combined via clustering with the main beam detection because this leads to a significant degradation of the resulting single measurement. This is due to the fact that the measurements in the adjacent beams (if a target detection occurs there) have a strong tendency of being attracted toward the center of the corresponding adjacent beam, i.e., they are biased. Thus such detections/measurements are treated as false, i.e., lumped with the false alarms.

threshold τ	P_{FA}	P_D at boresight	$P_{D\text{ave}}$	information reduction q
3.15	0.112	0.586	0.483	0.3792
3.10	0.115	0.593	0.489	0.3813
3.05	0.119	0.599	0.496	0.3826
3.00	0.122	0.605	0.503	0.3848
2.95	0.126	0.611	0.509	0.3861
2.90	0.130	0.617	0.516	0.3876
2.85	0.134	0.624	0.523	0.3891
2.80	0.138	0.630	0.530	0.3906
2.75	0.143	0.636	0.537	0.3913
2.70	0.147	0.643	0.544	0.3931
2.65	0.151	0.649	0.551	0.3949
2.60	0.156	0.656	0.559	0.3958
2.55	0.161	0.662	0.566	0.3969

Table 1. Detection thresholds used for the target with SNR=6 dB at boresight and the corresponding information reduction factors.

Parameter	Description	Value
N	radiating elements per side	55
P_t	transmitter power (MW)	1
$G_t = G_r = G$	antenna gain	$\frac{\pi}{2}N^2$
λ	nominal wavelength (m)	0.075
F_t	transmitter propagation factor	1
F_r	receiver propagation factor	1
L_{tot}	total system losses	144.5
k_0	Boltzmann constant (J/K)	1.38×10^{-23}
T_0	reference temperature (K)	290
F_n	receiver noise figure	2
R_{stc}	range for STC (km)	25
sq_0	broadside squint angle (deg)	0.9
b_0	bearing broadside angle (deg)	0
e_0	elevation broadside angle (deg)	15
τ_e	pulse length (μs)	675

Table 2. Radar Parameters

greater than 6 dB at boresight (or equivalently, the average SNR has to be greater than 4.4 dB). When the target SNR increases, the performance of the MDL model selection also improves.

For the same scenarios, we also use the GLRT with a threshold to obtain 95% acquisition probability under H_1 for a single target (based on the Gaussian approximation as in [15]). The results are listed in Table 4. We can see that the GLRT yields lower acquisition probabilities under both one-target and two-target scenarios compared with the MDL approach. This is because the threshold chosen by obtaining 95% target acquisition probability is not optimal in comparison with the penalty terms used in MDL.

N		40	50	60
SNR=6dB $\tau = 2.87$	$\hat{\mathrm{P}}(H_0\|H_0)$	0.96	0.98	0.99
	$\hat{\mathrm{P}}(H_1\|H_1)$	0.92	0.95	0.97
	$\hat{\mathrm{P}}(H_2\|H_2)$	0.93	0.96	0.98
SNR=7dB $\tau = 3.15$	$\hat{\mathrm{P}}(H_0\|H_0)$	0.98	1	1
	$\hat{\mathrm{P}}(H_1\|H_1)$	0.96	0.99	1
	$\hat{\mathrm{P}}(H_2\|H_2)$	0.98	1	1
SNR=8dB $\tau = 3.63$	$\hat{\mathrm{P}}(H_0\|H_0)$	1	1	1
	$\hat{\mathrm{P}}(H_1\|H_1)$	1	1	1
	$\hat{\mathrm{P}}(H_2\|H_2)$	1	1	1

Table 3. Acquisition results of three scenarios using MDL with various SNRs (at beam boresight), 100 runs.

N		40	50	60
SNR=6dB $\tau = 2.87$	$\hat{\mathrm{P}}(H_0\|H_0)$	0.93	0.98	1
	$\hat{\mathrm{P}}(H_1\|H_1)$	0.92	0.94	0.95
	$\hat{\mathrm{P}}(H_2\|H_2)$	0.85	0.89	0.88
SNR=7dB $\tau = 3.15$	$\hat{\mathrm{P}}(H_0\|H_0)$	1	1	1
	$\hat{\mathrm{P}}(H_1\|H_1)$	0.93	0.96	0.94
	$\hat{\mathrm{P}}(H_2\|H_2)$	0.86	0.87	0.89
SNR=8dB $\tau = 3.63$	$\hat{\mathrm{P}}(H_0\|H_0)$	1	1	1
	$\hat{\mathrm{P}}(H_1\|H_1)$	0.94	0.92	0.95
	$\hat{\mathrm{P}}(H_2\|H_2)$	0.88	0.89	0.87

Table 4. Acquisition results of three scenarios using GLRT with various SNRs (at beam boresight), 100 runs.

The CRLB and the sample standard deviation for the two-target scenario are listed in Table 5 with $N = 50$ and $\tau = 2.87$ for the 6 dB target and $\tau = 3.63$ for the 8 dB target. In most cases, the sample standard deviation is slightly larger than the approximate CRLB. This can be explained by several factors: a) the false alarms are not uniformly distributed in the cueing region; b) the range measurement noise is not Gaussian; c) the target trajectories used in simulation yield different off-boresight angles from scan to scan and could be better than the average CRLB. The results are acceptable in view of the fact that they are based on a 100 run average, which allow $\pm 14\%$ error within the 95% confidence region.

4. CONCLUSION

In this paper the track initiation problem was formulated as multiple composite hypothesis testing using maximum likelihood estimation with probabilistic data association (ML-PDA). The number of tracks was determined based on the minimum description length (MDL) criterion. We applied the MDL approach for the detection and initiation of tracks of incoming tactical ballistic missiles in the exo-atmospheric phase using a surface based electronically scanned array (ESA) radar. The targets were characterized by low SNR, which lead to low detection probability and high false alarm rate. A batch of radar scans were processed to detect the presence of up to two targets. The ML-PDA estimator was used to initiate the tracks assuming the target trajectories follow a deterministic state propagation.

The approximate MDL criterion was used to determine the number of valid tracks in a surveillance region. The detector/estimator was shown to be efficient even at 4.4 dB average SNR (within the beam, i.e., in a resolution cell).

REFERENCES

[1] Akaike, H., "A New Look at the Statistical Model Identification", *IEEE Trans. on Automatic Control*, vol.19, pp. 716-723, 1974.

[2] Bar-Shalom, Y., and Li, X. R., "Multisensor, Multitarget Tracking: Principles and Techniques", YBS Publishing, 1995.

[3] Bate, R., et al., *Fundamentals of Astrodynamics*. New York, NY: Dover, 1971.

[4] Blair, W. D., Watson, G. A., Kirubarajan T., and Bar-Shalom Y., "Benchmark for Radar Allocation and Tracking Targets in ECM", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 4, pp. 1097–1114, October 1998.

[5] Bozdogan, H., "On the Information-Based Measure of Covariance Complexity and Its Application to the Evaluation of Multivariate Linear Models", *Commun. Statist-Theory Meth.*, vol.19, pp. 221-278, 1990.

[6] Chen, H., Kirubarajan, T., and Bar-Shalom, Y., "Multiple Low Observable Ballistic Missile Track Initiation Using an ESA Radar", ESP Lab Tech Report, University of Connecticut, March 2001.

[7] Kay, S., *Fundamentals of Statistical Signal Processing, Vol. II–Detection Theory*, Prentice Hall, 1998.

[8] Kirubarajan, T., Bar-Shalom, Y., Blair, W. D., and Watson, G. A., "IMMPDAF for Radar Management and Tracking Benchmark with ECM", *IEEE Transactions on Aerospace and Electronic Systems*, vol.34, no. 4, pp. 1115-1132, October 1998.

[9] Rissanen, J., "Stochastic Complexity and the MDL Principle", *Econometric Reviews*, vol.6, pp. 85-102, 1987.

[10] Rissanen, J., "Fisher Information and Stochastic Complexity", *IEEE Trans. on Information Theory*, vol.42, pp. 40-47, January 1996.

[11] Rissanen, J., "Hypothesis Selection and Testing by the MDL Principle", *The Computer Journal*, vol.42, pp. 260-269, 1999.

[12] Schwartz, G., "Estimating the Dimension of a Model", *Annals of Statistics*, vol.6, pp. 461-464, 1978.

SNR	parameter of target 1	$\hat{\sigma}$	σ_{CRLB}	parameter of target 2	$\hat{\sigma}$	σ_{CRLB}
6 dB	ξ (m)	698.8	755.6	ξ (m)	720.2	754.4
6 dB	η (m)	1012.2	860.9	η (m)	965.3	862.2
6 dB	ζ (m)	398.2	371.7	ζ (m)	408.4	381.6
6 dB	$\dot{\xi}$ (m/s)	214.3	233.6	$\dot{\xi}$ (m/s)	219.6	233.3
6 dB	$\dot{\eta}$ (m/s)	297.8	265.5	$\dot{\eta}$ (m/s)	317.1	265.9
6 dB	$\dot{\zeta}$ (m/s)	138.4	114.3	$\dot{\zeta}$ (m/s)	130.6	117.4
8 dB	ξ (m)	628.3	648.9	ξ (m)	650.6	647.9
8 dB	η (m)	864.6	739.3	η (m)	901.5	740.4
8 dB	ζ (m)	310.4	319.2	ζ (m)	346.3	327.7
8 dB	$\dot{\xi}$ (m/s)	209.7	200.6	$\dot{\xi}$ (m/s)	210.5	200.3
8 dB	$\dot{\eta}$ (m/s)	267.8	228.0	$\dot{\eta}$ (m/s)	268.1	228.3
8 dB	$\dot{\zeta}$ (m/s)	88.7	98.1	$\dot{\zeta}$ (m/s)	96.9	100.8

Table 5. CRLB and the parameter estimates' sample standard deviations for the two-target scenario with various SNRs (at beam boresight).

[13] Sinha, A., Kirubarajan, T., and Bar-Shalom, Y., "Maximum Likelihood Angle Extractor for Two Closely-Spaced Targets", Proc. IEEE 2001 Radar Conf., Atlanta, GA, May 2001; submitted to *IEEE Transactions on Aerospace and Electronic Systems*, September 2000.

[14] Sinha, A., Kirubarajan, T., and Bar-Shalom, Y., "Maximum Likelihood Angle Extractor for a Single Swerling III Target", ESP Lab Tech Report, University of Connecticut, November 2000.

[15] Sivananthan, S., Bar-Shalom, Y., and Kirubarajan, T. "A Radar Power Multiplier Algorithm for Acquisition of Low Observable Ballistic Missiles Using ESA Radar", *IEEE Transactions on Aerospace and Electronic Systems*, vol.37, no. 2. pp. 401-418, April 2001.

[16] Watson, G. A., and McCabe, D. H., "Benchmark Problem with a Multisensor Framework for Radar Resource Allocation and Tracking of Highly Maneuvering Targets, Closely-Spaced Targets, and Targets in the Presence of Sea-Surface Induced Multipath", *Technical Report NSWCDD/TR-99/32*, Dahlgren VA, May 1999.

[17] Yeddanapudi, M., Bar-Shalom, Y., Pattipati, K. R., and Deb, S., "Ballistic missile track initiation from satellite observations", *IEEE Transactions on Aerospace and Electronic Systems*, vol.31, no. 3, pp. 1054-1071, July 1995.