

NUWC-NPT Technical Report 12,312
11 September 2019

Robust Approaches to Long-Term Maneuvering-Target Tracking

Michael J. Walsh
Sensors and Sonar Systems Department

Marcus L. Graham
Undersea Warfare Combat Systems Department



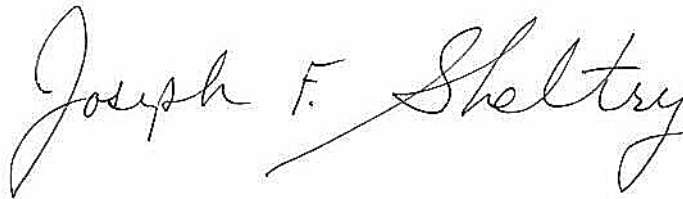
**Naval Undersea Warfare Center Division
Newport, Rhode Island**

PREFACE

This report was prepared under Naval Undersea Warfare Center Division Newport (NUWCDIVNPT) Network Activity No. 300000133944/0010, "Robust Methods for Long-Term Tracking of Maneuvering Targets," principal investigator Michael J. Walsh (Code 1511). The report was funded by a NUWCDIVNPT Section 219 Basic Research project.

The technical reviewer for this report was Phillip L. Ainsleigh (Code 1511).

Reviewed and Approved: 11 September 2019

A handwritten signature in black ink that reads "Joseph F. Sheltry". The signature is written in a cursive style with a long horizontal stroke at the end.

**Joseph F. Sheltry
Head, Sensors and Sonar Systems Department**



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OPM control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 11-09-2019	2. REPORT TYPE Technical Report	3. DATES COVERED (From - To)
--	---	-------------------------------------

4. TITLE AND SUBTITLE Robust Approaches to Long-Term Maneuvering-Target Tracking	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S) Michael J. Walsh Marcus L. Graham	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Undersea Warfare Mission Engineering and Analysis Department Naval Undersea Warfare Center Division Newport 1176 Howell Street Newport, RI 02841-1708	8. PERFORMING ORGANIZATION REPORT NUMBER TR 12,312
---	--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Undersea Warfare Mission Engineering and Analysis Department Naval Undersea Warfare Center Division Newport 1176 Howell Street Newport, RI 02841-1708	10. SPONSORING/MONITOR'S ACRONYM NUWC DIVNPT
	11. SPONSORING/MONITORING REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT Two complimentary approaches to long-term tracking of a maneuvering target, which are robust to changes in both target maneuver and measurement noise characteristics, are derived in this report. The first approach uses a hidden Gauss-Markov model (HGMM) to estimate the target maneuver and measurement noise characteristics at each update, and produces the track estimate as an ancillary output. The second approach uses continuous Gaussian mixture models (CGMMs) to accommodate heavy-tailed, non-Gaussian behavior in these characteristics at each update, and produces the track estimate as the primary output. Both approaches reduce to iterative Kalman smoothing problems in the linear case.

15. SUBJECT TERMS Maneuvering Target Tracking, Target Motion Analysis, Hidden Markov Model, Hidden Gauss-Markov Model, Continuous Gaussian Mixture Model, Interacting Multiple Model, Singer Model, Expectation-Maximization, Probabilistic Multi-Hypothesis Tracking, Kalman Filter, Kalman Smoother, Robust Estimation
--

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 43	19a. NAME OF RESPONSIBLE PERSON Michael J. Walsh
a. REPORT (U)	b. ABSTRACT (U)	c. THIS PAGE (U)			19b. TELEPHONE NUMBER (Include area code) (401) 832-4155

TABLE OF CONTENTS

Section	Page
LIST OF ABBREVIATIONS AND ACRONYMS	ii
1 INTRODUCTION	1
1.1 Background and Preview of Results	1
1.2 Motion Models.....	4
1.2.1 Constant Velocity.....	4
1.2.2 Singer	5
1.3 Measurement Models.....	6
2 HGMM APPROACH	7
2.1 Estimation of Process Noise Scales Q ,	8
2.2 Estimation of Measurement Noise Scales, \mathcal{R}	9
2.3 A Bayesian HGMM for Q and \mathcal{R}	10
2.4 Algorithm Convergence.....	12
3 CGMM APPROACH.....	15
3.1 Process and Measurement Noise Models	15
3.2 Complete-, Incomplete-, and Missing-Data PDFS	16
3.3 State Estimation	19
3.4 Algorithm Convergence.....	21
4 EXAMPLE.....	23
5 CONCLUDING REMARKS AND FUTURE WORK	33
APPENDIX—CLOSED-FORM EXPRESSIONS FOR CGMM INTEGRALS	A-1
REFERENCES	R-1

LIST OF ILLUSTRATIONS

Figure	Page
1 Uniform and Inverse-Gamma Priors for q for Three Values of the Lower/Upper Bounds $[a, b]$ and Shape/Scale Parameters $\{\alpha, \beta\}$, Respectively	12
2 CGMs for the Measurement Noise PDF pVn for Three Different Priors for the Covariance Scale, r , in Equation (65) with $\bar{R}_n = I_2$	17
3 Measurement and Process Noise Covariance Scale Priors Used in Example	23
4 True Track and Measurements (Left), Measurement Noise Scale Initializations (Top Right), and Process Noise Scale Initializations (Bottom Right) at Iteration $\ell = 0$ for HGMM Approach—Measurement Outliers are Circled.....	25

LIST OF ILLUSTRATIONS (Cont'd)

Figure	Page
5	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 1$26
6	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 2$27
7	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 4$28
8	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 8$29
9	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 16$30
10	True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 256$31
11	True Track and Measurements (Left, Red), CGMM Track Estimate (Left, Black), Effective Measurement Noise Scales (Top Right), and Effective Process Noise Scales (Bottom Right) at Iteration $\ell = 16$32

LIST OF ABBREVIATIONS AND ACRONYMS

CGM	Continuous Gaussian Mixture
CGMM	CGM Model
E-Step	Expectation Step
EM	Expectation-Maximization
HGMM	Hidden Gauss-Markov Model
HMM	Hidden Markov Model
IMM	Interacting Multiple Model
M-Step	Maximization Step
NUWCDIVNPT	Naval Undersea Warfare Center Division Newport
PDA	Probabilistic Data Association
PDF	Probability Density Function
PMHT	Probabilistic Multi-Hypothesis Tracking

ROBUST APPROACHES TO LONG-TERM MANEUVERING-TARGET TRACKING

1. INTRODUCTION

1.1 BACKGROUND AND PREVIEW OF RESULTS

This report is concerned with long-term tracking of a maneuvering targets, where long-term is meant to imply tens of minutes or more. In this time frame, a contact of interest (or target) is expected to occasionally maneuver with varying degrees, depending on the target's objective and the environment. Tracking a maneuvering target from noisy data in the possible presence of spurious measurements (or clutter) is challenging because it is difficult to distinguish between clutter and measurements from a target that has maneuvered or deviated from its assumed motion model. The difficulty is compounded by often low signal-to-noise ratios and missed target detections, high false-alarm rates and dense clutter, and nonlinear mappings between the measurements (e.g., bearings and ranges) and the target kinematic states.

There is a large body of literature on maneuvering target tracking. The most popular approach to date is the interacting multiple model (IMM) filter (reference 1), which augments the standard (i.e., fixed motion model) Kalman filter with multiple motion models. In this approach, the target is assumed to switch (or jump) between models according to a discrete-time Markov chain with known transition probabilities. Several modifications and extensions of the IMM filter have been proposed, the most prominent of which is the IMM probabilistic data association (IMM/PDA) filter (reference 2), which replaces the Kalman filter with the PDA filter (reference 3) to handle clutter.

Long-term track estimates can be significantly improved by processing a batch of data scans collected over time to produce smoothed track estimates over the batch, rather than processing the data sequentially as it arrives from one scan to the next (yielding filtered but coarser track estimates over the same batch). Though batch implementations of the IMM/PDA filter have been proposed (reference 4), these have a computational cost that grows exponentially with batch length and are only practical as fixed-lag smoothers for relatively few scans of data; hence the batch IMM/PDA approach is not suitable for long-term tracking applications where the best possible track accuracy is required over the entire length of the track using all available data.

Two complementary approaches to long-term maneuvering-target tracking are derived in this report, both of which have batch implementations that grow linearly with batch length. Both approaches adopt the same system model as the standard Kalman filter with a state equation that describes the evolution of the target kinematic state, and a measurement equation that relates the target state to a measurement of the target at each scan. In particular, let n denote the scan index, let x_n denote the target state at time t_n , where x is an $L \times 1$ vector of kinematic parameters (such as position and velocity), and let z_n denote the measurement of target state at t_n , where z is an

$M \times 1$ vector. Then, for the standard Kalman filter the system model is given by

$$x_n = A_n x_{n-1} + w_n, \quad (1)$$

$$z_n = B_n x_n + v_n, \quad (2)$$

where A_n is a known $L \times L$ state transition matrix, B_n is a known $M \times L$ measurement matrix, and w_n and v_n are independent, zero-mean, Gaussian distributed process and measurement noise terms with known $L \times L$ and $M \times M$ covariance matrices Q_n and R_n , respectively. The probability density functions (PDFs) for w_n and v_n are written as

$$p(w_n) = \mathcal{N}(w_n|0, Q_n), \quad (3)$$

$$p(v_n) = \mathcal{N}(v_n|0, R_n), \quad (4)$$

where $\mathcal{N}(\xi|\mu, P)$ denotes the multivariate Gaussian PDF with respect to the realization, ξ , of a Gaussian random variable with mean μ and covariance matrix P , given by

$$\mathcal{N}(\xi|\mu, P) = \frac{1}{|2\pi P|^{1/2}} \exp\left[-\frac{1}{2}(\xi - \mu)^T P^{-1}(\xi - \mu)\right]. \quad (5)$$

The additive process noise terms, w_n , in the state equation of the standard Kalman filter model are the only mechanisms for modeling target maneuvers outside of the deterministic motion model specified by the state transition matrices, A_n , and the magnitudes of these maneuvers are fixed by the values of the process noise covariance matrices, Q_n . Without knowing the target maneuver sequence *a priori*, it is impossible to choose fixed values for Q_n to adequately model the maneuver behavior of the target. To address this problem, IMM approaches assume a fixed, finite set of K known motion models, $\{A_n(k), Q_n(k) : k = 1, 2, \dots, K\}$, and a discrete-time Markov chain with known transition probabilities to allow switching (or jumping) between models based on evidence for such maneuvers in the data. However, due to the need to enumerate all possible model transition sequences for optimal state estimation, these approaches are only computationally feasible for fixed-lag smoothing of relatively few scans of data.

In the maneuvering-target approaches derived here, and as in reference 5, the process noise covariance matrices, Q_n , are assumed to have known structures but unknown magnitudes, the latter of which are estimated—either explicitly or implicitly—from all available data. In particular, it is assumed the process noise covariance matrix Q_n has the form

$$Q_n = q_n \bar{Q}_n, \quad (6)$$

where \bar{Q}_n is a positive-definite matrix with known structure, and q_n is the process noise scale associated with the target state update for scan n . Then, given a batch of N data scans, the sequence of process noise covariance scales,

$$\mathcal{Q} = \{q_n : n = 1, 2, \dots, N\}, \quad (7)$$

describes the target maneuver behavior over the entire length of the batch. The first approach to maneuvering-target tracking, derived in section 2, is based on the theory of hidden Gauss-Markov

models (HGMMs)* developed in references 8 and 9. In this approach, the target states $X = \{x_n : n = 0, 1, \dots, N\}$ are treated as missing data, and the process noise covariance scales, Q , are estimated from the batch of data $Z = \{z_n : n = 1, 2, \dots, N\}$ using the expectation-maximization (EM) method (references 10–12). Interestingly, though estimates of the target states, X , are not the primary outputs of this approach, meaningful estimates of X are nevertheless available as ancillary outputs of the expectation step (E-step) of the EM procedure for estimating the process noise covariance scales, Q . The same is true of a similar approach derived in reference 13, where the target states, X , are treated as missing data, and the EM method is used to estimate a sequence of target maneuver inputs (accelerations).

The second approach to maneuvering-target tracking, derived in section 3, is based on the use of continuous Gaussian mixture models (CGMMs) for the PDFs of the process noise terms, w_n , and is similar in approach and derivation to the probabilistic multi-hypothesis tracking (PMHT) method (references 14 and 15) originally developed for multitarget tracking, and later applied to maneuvering-target tracking (references 5, 16, 17, and 18). In this approach, the process noise covariance scales, Q , are treated as missing data, and the target states, X , are estimated from the batch of data, Z , using the EM method. Though estimates of the target states are the primary outputs of this approach, estimates of the process noise covariance scales can be calculated from the missing-data PDF for Q required by the estimation procedure for X .

Both the HGMM and CGMM approaches to long-term maneuvering-target tracking can be extended to include a clutter model. Even in the case where only one measurement is observed at each scan, clutter modeling can make the resulting algorithm more robust to outliers. As for the process noise covariance matrices, Q_n , the measurement noise covariance matrices, R_n , in both approaches are assumed to have known structures but unknown magnitudes, the latter of which are estimated—either explicitly or implicitly—from all available data. In particular, it is assumed the measurement noise covariance matrix, R_n , has the form

$$R_n = r_n \bar{R}_n, \tag{8}$$

where \bar{R}_n is a positive-definite matrix with known structure, and r_n is the measurement noise scale associated with the measurement for scan n . Then, given a batch of N data scans, the sequence of measurement noise covariance scales,

$$\mathcal{R} = \{r_n : n = 1, 2, \dots, N\}, \tag{9}$$

describes the measurement noise characteristics over the entire length of the batch. In the HGMM approach, the measurement noise covariance scales, \mathcal{R} , are estimated from the batch of data along with the process noise scales, Q , in an iterative procedure after the missing target states, X , are integrated out of the problem via the E-step of the EM method. In the CGMM approach, the measurement noise covariance scales, \mathcal{R} , are treated as missing data along with the process noise covariance scales, Q , and the target states, X , are estimated from the batch of data, Z , using EM; as for estimates of Q , estimates of the measurement noise covariance scales can be calculated from the missing-data PDF for \mathcal{R} required by the estimation procedure for X in this approach.

*HGMMs are a special case of continuous-state hidden Markov models (HMMs), which extend discrete-state HMMs (references 6 and 7) to include the continuous-state model in equations (1) and (2).

The remainder of this section discusses the additional assumptions underlying the derivations of the HGMM and CGMM approaches presented in sections 2 and 3, respectively. Specifically, it is assumed the target kinematic states at times $\{t_n : n = 0, 1, \dots, N\}$ are to be estimated from the measurements $Z = \{z_n : n = 1, 2, \dots, N\}$ of the target taken at times $\{t_n : n = 1, 2, \dots, N\}$. While no measurement of the target is assumed at time t_0 , the target state at this time, x_0 , is assumed to be Gaussian distributed with known mean, μ_0 , and covariance matrix, P_0 , so that its prior PDF is given by

$$p(x_0) = \mathcal{N}(x_0 | \mu_0, P_0). \quad (10)$$

Also, both the target state and measurement equations are assumed to be linear, as in the system model given by equations (1) and (2). A linear state equation is a practical assumption as it includes constant-velocity motion—the normal steady-state mode for many targets of interest—in a Cartesian coordinate system. Forms of the state transition matrices, A_n , and unscaled process noise covariance matrices, \bar{Q}_n , for both the constant-velocity motion and commonly used Singer motion models (reference 19) are given in the following subsection. A linear measurement equation is a somewhat more limiting assumption, as many sensors of practical interest estimate target kinematic parameters in non-Cartesian coordinates (e.g., range, bearing, and their rates, etc.). However, measurement conversion procedures such as those described in reference 20, which eliminate biases in measurements converted to Cartesian coordinates, are available for effectively circumventing nonlinear measurement equations in Kalman filtering. Extension of these measurement conversion procedures to the smoothing problem would appear to be straightforward but are not pursued here.

1.2 MOTION MODELS

The state transition matrices, A_n , and unscaled process noise covariance matrices, \bar{Q}_n , for the constant-velocity motion model and commonly used Singer motion model are provided in this subsection. To simplify notation, the discussion will be restricted to two dimensions, though extension to three dimensions is straightforward. In each case, the time difference between t_n and t_{n-1} will be denoted by $\tau_n = t_n - t_{n-1}$.

1.2.1 Constant Velocity

Let the target state, x , be defined by its position, (u_1, u_2) , and velocity, (\dot{u}_1, \dot{u}_2) , in a two-dimensional Cartesian coordinate system as

$$x = \begin{bmatrix} u_1 \\ \dot{u}_1 \\ u_2 \\ \dot{u}_2 \end{bmatrix}. \quad (11)$$

For constant-velocity target motion, the state transition matrices, A_n , are given by

$$A_n = I_2 \otimes \begin{bmatrix} 1 & \tau_n \\ 0 & 1 \end{bmatrix}, \quad (12)$$

where I_2 denotes the 2×2 identity matrix, and \otimes denotes the Kronecker product. The unscaled process noise covariance matrices, \bar{Q}_n , for this motion model are obtained via discretization of the continuous-time version of the state equation given by equation (1) (see reference 21, section 6.2.2), and are given by

$$\bar{Q}_n = I_2 \otimes \begin{bmatrix} \tau_n^3/3 & \tau_n^2/2 \\ \tau_n^2/2 & \tau_n \end{bmatrix}. \quad (13)$$

1.2.2 Singer

For the constant-velocity motion model, the additive process noise terms, w_n , in the state equation of the standard Kalman filter model are the only mechanisms for modeling target maneuvers. The Singer motion model presented in reference 19 provides additional accommodations for maneuvers by effectively adding acceleration to the target state while maintaining constant-velocity motion as the normal mode.

Let the target state, x , be defined by its position, (u_1, u_2) , velocity, (\dot{u}_1, \dot{u}_2) , and acceleration, (\ddot{u}_1, \ddot{u}_2) , in a two-dimensional Cartesian coordinate system as

$$x = \begin{bmatrix} u_1 \\ \dot{u}_1 \\ \ddot{u}_1 \\ u_2 \\ \dot{u}_2 \\ \ddot{u}_2 \end{bmatrix}. \quad (14)$$

The state transition matrices, A_n , and process noise covariance matrices, Q_n , for the Singer model are parameterized by two parameters: the magnitude (or variance) and time constant (or duration) of the target maneuver, denoted $\sigma^2 > 0$ and $1/\alpha \geq 0$, respectively. In particular, the state transition matrices are given by equation (13) in reference 19 as

$$A_n = I_2 \otimes \begin{bmatrix} 1 & \tau_n & (-1 + \alpha\tau_n + e^{-\alpha\tau_n})/\alpha^2 \\ 0 & 1 & (1 - e^{-\alpha\tau_n})/\alpha \\ 0 & 0 & e^{-\alpha\tau_n} \end{bmatrix}. \quad (15)$$

The unscaled process noise covariance matrices for the Singer model are given by

$$\bar{Q}_n = I_2 \otimes \bar{Q}_n^*, \quad (16)$$

where \bar{Q}_n^* is a symmetric matrix with upper-triangular elements given by equations (19) and (20) in reference 19; if $\bar{Q}_n^*(i, j)$ denotes the element in the i th row and j th column of \bar{Q}_n^* , then

$$\begin{aligned} \bar{Q}_n^*(1, 1) &= (1 - e^{-2\alpha\tau_n} + 2\alpha\tau_n + 2\alpha^3\tau_n^3/3 - 2\alpha^2\tau_n^2 - 4\alpha\tau_n e^{-\alpha\tau_n})/\alpha^4, \\ \bar{Q}_n^*(1, 2) &= (e^{-2\alpha\tau_n} + 1 - 2e^{-\alpha\tau_n} + 2\alpha\tau_n e^{-\alpha\tau_n} - 2\alpha\tau_n + \alpha^2\tau_n^2)/\alpha^3, \\ \bar{Q}_n^*(1, 3) &= (1 - e^{-2\alpha\tau_n} - 2\alpha\tau_n e^{-\alpha\tau_n})/\alpha^2, \\ \bar{Q}_n^*(2, 2) &= (4e^{-\alpha\tau_n} - 3 - e^{-2\alpha\tau_n} + 2\alpha\tau_n)/\alpha^2, \\ \bar{Q}_n^*(2, 3) &= (e^{-2\alpha\tau_n} + 1 - 2e^{-\alpha\tau_n})/\alpha, \\ \bar{Q}_n^*(3, 3) &= 1 - e^{-2\alpha\tau_n}. \end{aligned} \quad (17)$$

The target maneuver magnitude, σ^2 , in the Singer model corresponds to the process noise covariance scales, q_n , in equation (6); hence, the process noise covariance matrices for the Singer model are given by equation (6), with \bar{Q}_n given by equation (16), and

$$q_n = \sigma^2 \quad (18)$$

for $n = 1, 2, \dots, N$. Thus, when the Singer motion model is used in the HGMM and CGMM approaches developed in sections 2 and 3, the resulting maneuvering-target tracking algorithms may be viewed as generalizations of the Kalman-Singer smoother derived in reference 19, in which the target maneuver magnitude parameter σ^2 is automatically adapted over the batch.

1.3 MEASUREMENT MODELS

The forms of the measurement matrices, B_n , and unscaled measurement covariance matrices, \bar{R}_n , in equation (2) depend on the state space and measurement space as defined by the sensor, as well as the sensor measurement characteristics. For example, suppose the sensor is able to measure the individual components of target position in Cartesian coordinates, and the measurement components have the same variance but are correlated with correlation coefficient ρ . Then, for the constant-velocity motion model with target state defined by equation (11), the measurement matrices, B_n , are given by

$$B_n = I_2 \otimes [1 \ 0] . \quad (19)$$

Likewise, for the Singer motion model with target state defined by equation (14), the measurement matrices are given by

$$B_n = I_2 \otimes [1 \ 0 \ 0] . \quad (20)$$

In both cases, the unscaled measurement covariance matrices, \bar{R}_n , are interpreted as a correlation matrices, and take the form

$$\bar{R}_n = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} . \quad (21)$$

For nonlinear measurement models such as those for sensors that estimate range and bearing to the target, the measurement conversion procedures discussed in reference 20 produce converted measurements which are, in general, correlated, and hence corresponding converted measurement covariance matrices which are, in general, non-diagonal. For use in the HGMM and CGMM approaches to maneuvering-target tracking derived here, these converted measurement covariance matrices may be substituted directly for the unscaled measurement covariance matrices, \bar{R}_n , in equation (8).

2. HGMM APPROACH

HGMM theory is fully developed in reference 8 for the general case where the entire collection of system parameters in the linear-Gaussian state-space model given by equations (1) through (4) and equation (10), namely

$$\mathcal{S} = \{\mu_0, P_0\} \cup \{A_n, Q_n, B_n, R_n : n = 1, 2, \dots, N\}, \quad (22)$$

are to be estimated from K sequences of measurements $\mathcal{Z} = \{Z^k : k = 1, 2, \dots, K\}$ of possibly unequal lengths, with the k th measurement sequence Z^k denoted by

$$Z^k = \{z_n^k : n \in \mathcal{N}^k \subseteq \{1, 2, \dots, N\}\}. \quad (23)$$

In this application of the theory to the long-term maneuvering-target tracking problem, the state transition and measurement matrices, as well as the mean and covariance of the target state at t_0 , are assumed known, and the process and measurement noise covariance matrices, Q_n and R_n , are assumed to take the forms given by equations (6) and (8), respectively, with matrices \bar{Q}_n and \bar{R}_n of known form as discussed in sections 1.2 and 1.3. The system parameters to be estimated here are the process and measurement noise covariance scales, \mathcal{Q} and \mathcal{R} , respectively; moreover, only one measurement sequence Z is assumed.

The HGMM-based algorithm for estimating the system parameters of interest is derived via the EM method, which leads to an iterative procedure for the parameter estimates that guarantees a monotonic increase in the incomplete-data PDF (likelihood function), $p(Z|\mathcal{S})$, at each iteration and, under mild regularity conditions, convergence to at least a local maximum of this function (references 10, 12, and 22). In the sequel, let $\mathcal{S} \equiv \mathcal{S}(\mathcal{Q}, \mathcal{R})$ denote the collection of system parameters in the maneuvering-target tracking problem formulated here, with

$$\mathcal{S}(\mathcal{Q}, \mathcal{R}) = \{\mathcal{Q}, \mathcal{R}\} \cup \{\mu_0, P_0\} \cup \{A_n, \bar{Q}_n, B_n, \bar{R}_n : n = 1, 2, \dots, N\}, \quad (24)$$

where all parameters in the collection other than \mathcal{Q} and \mathcal{R} are assumed known. Let $\ell = 1, 2, \dots$ denote the EM iteration index, and let $\mathcal{Q}^{(\ell)}$ and $\mathcal{R}^{(\ell)}$ denote the process and measurement noise covariance scale estimates at the ℓ th EM iteration. The EM method consists of two steps at each iteration: an E-step and a maximization step (M-step). The output of the E-step for the ℓ th iteration is an auxiliary function of the system parameters, \mathcal{S} , denoted $\Upsilon(\mathcal{S}|\mathcal{S}^{(\ell)})$, defined as the expectation of the logarithm of the complete-data PDF, $p(X, Z|\mathcal{S})$, conditioned on the observed data, Z , and the estimate of the system parameters at the beginning of the iteration, $\mathcal{S}^{(\ell)} \equiv \mathcal{S}(\mathcal{Q}^{(\ell)}, \mathcal{R}^{(\ell)})$:

$$\begin{aligned} \Upsilon(\mathcal{S}|\mathcal{S}^{(\ell)}) &= \mathbb{E}[\log p(X, Z|\mathcal{S})|Z, \mathcal{S}^{(\ell)}] \\ &= \int [\log p(X, Z|\mathcal{S})|Z, \mathcal{S}^{(\ell)}] p(X|Z, \mathcal{S}^{(\ell)}) dX. \end{aligned} \quad (25)$$

The M-step for the ℓ th iteration yields an updated estimate for the unknown system parameters as the value of \mathcal{S} that maximizes the auxiliary function:

$$\mathcal{S}^{(\ell+1)} = \arg \max_{\mathcal{Q}, \mathcal{R}} \Upsilon(\mathcal{S}(\mathcal{Q}, \mathcal{R})|\mathcal{S}^{(\ell)}). \quad (26)$$

This updated estimate is used, in turn, as the input to the next (i.e., $(\ell + 1)$ th) EM iteration. As shown for the general case in reference 8, this auxiliary function separates into two auxiliary functions, one whose unknown parameters include only the process noise covariance scales, \mathcal{Q} , and one whose unknowns include only the measurement noise covariance scales, \mathcal{R} :

$$\Upsilon(\mathcal{S}(\mathcal{Q}, \mathcal{R})|\mathcal{S}^{(\ell)}) = \Upsilon_{\mathcal{Q}}(\mathcal{Q}|\mathcal{S}^{(\ell)}) + \Upsilon_{\mathcal{R}}(\mathcal{R}|\mathcal{S}^{(\ell)}). \quad (27)$$

Hence, the M-step in equation (26) separates into two independent maximization problems, one for the updated estimate of \mathcal{Q} , and one for the updated estimate of \mathcal{R} :

$$\mathcal{Q}^{(\ell+1)} = \arg \max_{\mathcal{Q}} \Upsilon_{\mathcal{Q}}(\mathcal{Q}|\mathcal{S}^{(\ell)}), \quad (28)$$

$$\mathcal{R}^{(\ell+1)} = \arg \max_{\mathcal{R}} \Upsilon_{\mathcal{R}}(\mathcal{R}|\mathcal{S}^{(\ell)}). \quad (29)$$

The details of these maximization sub-problems are provided in the following two subsections.

While estimates for the process and measurement noise scales, \mathcal{Q} and \mathcal{R} , are the primary outputs of the HGMM algorithm for this problem, meaningful estimates of the target states, X , are available as ancillary outputs from the E-step of the final EM iteration. Indeed, as shown in reference 8, evaluation of the auxiliary function in equation (25) requires evaluation of the conditional state PDFs $\gamma(x_n) \equiv p(x_n|Z, \mathcal{S})$. For the linear-Gaussian case, these PDFs are given by

$$\gamma(x_n) = \mathcal{N}(x_n|\mu_{n|N}, P_{n|N}), \quad (30)$$

where $\mu_{n|N}$ and $P_{n|N}$ are the target states estimates and associated covariance matrices, respectively, obtained from a Kalman smoother with system parameters, \mathcal{S} , applied to the data, Z . In the following two subsections, the notation $\mu_{n|N}^{(\ell)}$ and $P_{n|N}^{(\ell)}$ will be used to denote the state estimates and associated covariance matrices obtained from the E-step of the ℓ th EM iteration, conditioned on the system parameters $\mathcal{S}(\mathcal{Q}^{(\ell)}, \mathcal{R}^{(\ell)})$.

2.1 ESTIMATION OF PROCESS NOISE SCALES, \mathcal{Q}

Following the developments in section 3.4 of reference 8, the auxiliary function for \mathcal{Q} conditioned on $\mathcal{S}^{(\ell)}$ is given by

$$\Upsilon_{\mathcal{Q}}(\mathcal{Q}|\mathcal{S}^{(\ell)}) = \sum_{n=1}^N \iint \gamma^{(\ell)}(x_n, x_{n-1}) \log p(x_n|x_{n-1}, q_n) dx_{n-1} dx_n, \quad (31)$$

where

$$\gamma^{(\ell)}(x_n, x_{n-1}) = \mathcal{N}(x_n|\mu_{n|N}^{(\ell)}, P_{n|N}^{(\ell)}) \mathcal{N}(x_{n-1}|\lambda_n^{(\ell)}, \Lambda_n^{(\ell)}), \quad (32)$$

$$p(x_n|x_{n-1}, q_n) = \mathcal{N}(x_n|A_n x_{n-1}, q_n \bar{Q}_n), \quad (33)$$

with

$$\lambda_n^{(\ell)} = (I - H_n^{(\ell)} A_n) \mu_{n-1|n-1}^{(\ell)} + H_n^{(\ell)} x_n, \quad (34)$$

$$\Lambda_n^{(\ell)} = (I - H_n^{(\ell)} A_n) P_{n-1|n-1}^{(\ell)}, \quad (35)$$

and

$$H_n^{(\ell)} = P_{n-1|n-1}^{(\ell)} A_n^T (P_{n|n-1}^{(\ell)})^{-1}. \quad (36)$$

The filtered state estimates, $\mu_{n-1|n-1}^{(\ell)}$, and associated covariance matrices, $P_{n-1|n-1}^{(\ell)}$, as well as the predicted state covariance matrices, $P_{n|n-1}^{(\ell)}$, in equations (34) through (36) are obtained from the forward pass of the Kalman smoother for $\mu_{n|N}^{(\ell)}$ and $P_{n|N}^{(\ell)}$. Substituting equations (32) and (33) into equation (31), evaluating the logarithm and dropping terms not dependent on q_n yields

$$\Upsilon_{\mathcal{Q}}(\mathcal{Q}|\mathcal{S}^{(\ell)}) = \sum_{n=1}^N \left[-\frac{L}{2} \log q_n - \frac{1}{2} q_n^{-1} \Psi_n(\mathcal{S}^{(\ell)}) \right], \quad (37)$$

where $\Psi_n(\mathcal{S}^{(\ell)})$ is a function independent of \mathcal{Q} , given by

$$\begin{aligned} \Psi_n(\mathcal{S}^{(\ell)}) = & \iint \mathcal{N}(x_n | \mu_{n|N}^{(\ell)}, P_{n|N}^{(\ell)}) \mathcal{N}(x_{n-1} | \lambda_n^{(\ell)}, \Lambda_n^{(\ell)}) \\ & \times (x_n - A_n x_{n-1})^T \bar{Q}_n^{-1} (x_n - A_n x_{n-1}) dx_{n-1} dx_n. \end{aligned} \quad (38)$$

The EM update for q_n is then easily obtained by setting the derivative of equation (37) with respect to q_n equal to zero, and solving the resulting expression for q_n evaluated at $q_n^{(\ell+1)}$, yielding

$$q_n^{(\ell+1)} = \frac{1}{L} \Psi_n(\mathcal{S}^{(\ell)}). \quad (39)$$

Evaluating the integrals in the expression for $\Psi_n(\mathcal{S}^{(\ell)})$ in equation (38) is simplified by repeated application of the following identity from reference 8:

$$\int \mathcal{N}(x | \mu, P) (x^T G x + x^T g + c) dx = \text{tr} [G(P + \mu\mu^T)] + \mu^T g + c, \quad (40)$$

where G is a matrix and g is vector, each sized according to the dimension of x , and c is a scalar. Using this identity, equation (38) simplifies to

$$\Psi_n(\mathcal{S}^{(\ell)}) = \text{tr} (\bar{Q}_n^{-1} C_{x_n x_n}^{(\ell)}) - 2 \text{tr} (\bar{Q}_n^{-1} A_n (C_{x_n x_{n-1}}^{(\ell)})^T) + \text{tr} (A_n^T \bar{Q}_n^{-1} A_n C_{x_{n-1} x_{n-1}}^{(\ell)}), \quad (41)$$

with

$$C_{x_n x_n}^{(\ell)} = P_{n|N}^{(\ell)} + \mu_{n|N}^{(\ell)} (\mu_{n|N}^{(\ell)})^T, \quad (42)$$

$$C_{x_n x_{n-1}}^{(\ell)} = P_{n,n-1|N}^{(\ell)} + \mu_{n|N}^{(\ell)} (\mu_{n-1|N}^{(\ell)})^T, \quad (43)$$

$$P_{n,n-1|N}^{(\ell)} = P_{n|N}^{(\ell)} (H_n^{(\ell)})^T. \quad (44)$$

2.2 ESTIMATION OF MEASUREMENT NOISE SCALES, \mathcal{R}

Following the developments in section 3.4 of reference 8, the auxiliary function for \mathcal{R} conditioned on $\mathcal{S}^{(\ell)}$ is given by

$$\Upsilon_{\mathcal{R}}(\mathcal{R}|\mathcal{S}^{(\ell)}) = \sum_{n=1}^N \int \gamma^{(\ell)}(x_n) \log p(z_n | x_n, r_n) dx_n, \quad (45)$$

where

$$\gamma^{(\ell)}(x_n) = \mathcal{N}(x_n | \mu_{n|N}^{(\ell)}, P_{n|N}^{(\ell)}), \quad (46)$$

$$p(z_n | x_n, r_n) = \mathcal{N}(z_n | B_n x_n, r_n \bar{R}_n). \quad (47)$$

Substituting equations (46) and (47) into equation (45), evaluating the logarithm and dropping terms not dependent on r_n yields

$$\Upsilon_{\mathcal{R}}(\mathcal{R} | \mathcal{S}^{(\ell)}) = \sum_{n=1}^N \left[-\frac{M}{2} \log r_n - \frac{1}{2} r_n^{-1} \Phi_n(z_n | \mathcal{S}^{(\ell)}) \right], \quad (48)$$

where $\Phi_n(z_n | \mathcal{S}^{(\ell)})$ is a data-dependent function independent of \mathcal{R} , given by

$$\Phi_n(z_n | \mathcal{S}^{(\ell)}) = \int \mathcal{N}(x_n | \mu_{n|N}^{(\ell)}, P_{n|N}^{(\ell)}) (z_n - B_n x_n)^T \bar{R}_n^{-1} (z_n - B_n x_n) dx_n. \quad (49)$$

The EM update for r_n is then easily obtained by setting the derivative of equation (48) with respect to r_n equal to zero, and solving the resulting expression for r_n evaluated at $r_n^{(\ell+1)}$, yielding

$$r_n^{(\ell+1)} = \frac{1}{M} \Phi_n(z_n | \mathcal{S}^{(\ell)}). \quad (50)$$

Evaluating the integral in the expression for $\Phi_n(z_n | \mathcal{S}^{(\ell)})$ in equation (49) is simplified by repeated application of the identity in equation (40); using this identity, equation (49) simplifies to

$$\Phi_n(z_n | \mathcal{S}^{(\ell)}) = \text{tr} (B_n^T \bar{R}_n^{-1} B_n C_{z_n x_n}^{(\ell)}) - 2 \text{tr} (B_n^T \bar{R}_n^{-1} C_{z_n x_n}^{(\ell)}) + \text{tr} (\bar{R}_n^{-1} C_{z_n z_n}^{(\ell)}), \quad (51)$$

with

$$C_{z_n x_n}^{(\ell)} = z_n (\mu_{n|N}^{(\ell)})^T, \quad (52)$$

$$C_{z_n z_n}^{(\ell)} = z_n z_n^T. \quad (53)$$

2.3 A BAYESIAN HGMM FOR \mathcal{Q} AND \mathcal{R}

The unknown system parameters in the HGMM theory developed in reference 8 are treated as deterministic parameters. For the maneuvering-target tracking application, treating the unknown process and measurement noise covariance scales, q_n and r_n , as random variables with prior PDFs given by $f(q_n)$ and $f(r_n)$, respectively, is considered. This modeling change leads to a Bayesian HGMM for \mathcal{Q} and \mathcal{R} . In the following, the impact of a prior on the estimation of the process noise scales, q_n , is examined; the impact on the measurement noise scales, r_n , is similar.

Assuming the process noise covariance scales, q_n , are independent and identically distributed random variables with prior PDF $f(q_n)$, it is straightforward to modify the derivations in sections 2.4 and 3.4 of reference 8 to show the auxiliary function for \mathcal{Q} given by equation (37) becomes

$$\Upsilon_{\mathcal{Q}}(\mathcal{Q} | \mathcal{S}^{(\ell)}) = \sum_{n=1}^N \left[-\frac{L}{2} \log q_n - \frac{1}{2} q_n^{-1} \Psi_n(\mathcal{S}^{(\ell)}) + \log f(q_n) \right]. \quad (54)$$

As in the deterministic case, the EM update for q_n in the Bayesian case is obtained by setting the derivative of equation (54) with respect to q_n equal to zero, and solving the resulting expression for q_n evaluated at $q_n^{(\ell+1)}$.

Choice of the prior PDF for q_n can reflect actual prior knowledge of the range of values for q_n , or can be used to effectively emphasize certain values of q_n over others; for example, a prior PDF, $f(q_n)$, with more mass near zero can be used to emphasize constant-velocity motion over maneuverability. A uniform prior for q_n on the interval $[a, b] \subset (0, \infty)$, given by

$$f(q_n) = \begin{cases} \frac{1}{b-a}, & q_n \in [a, b], \\ 0, & \text{otherwise,} \end{cases} \quad (55)$$

provides no prior information on q_n beyond its bounding values, and only serves to constrain the update $q_n^{(\ell+1)}$ given by equation (39) between these bounds. A useful and more informative alternative to the uniform prior for q_n is the inverse-gamma prior (reference 23), with PDF on the interval $(0, \infty)$ given by

$$f(q_n) = \frac{\beta^\alpha}{\Gamma(\alpha)} q_n^{-\alpha-1} \exp\left(-\frac{\beta}{q_n}\right), \quad q_n \in (0, \infty), \quad (56)$$

where $\alpha > 0$ and $\beta > 0$ are real-valued shape and scale parameters, respectively, and $\Gamma(\cdot)$ denotes the gamma function. Substituting equation (56) into equation (54) and simplifying the resulting expression, the auxiliary function for \mathcal{Q} reduces to

$$\Upsilon_{\mathcal{Q}}(\mathcal{Q}|\mathcal{S}^{(\ell)}) = \sum_{n=1}^N \left\{ -\frac{L+2(\alpha+1)}{2} \log q_n - \frac{1}{2} q_n^{-1} [\Psi_n(\mathcal{S}^{(\ell)}) + 2\beta] \right\}. \quad (57)$$

Hence, in the case of an inverse-gamma prior for the process noise covariance scales, q_n , the EM update for q_n becomes

$$q_n^{(\ell+1)} = \frac{1}{L+2(\alpha+1)} [\Psi_n(\mathcal{S}^{(\ell)}) + 2\beta], \quad (58)$$

a simple modification of the form of the update given by equation (39) for the deterministic case.

The inverse-gamma distribution is a natural choice for the prior distribution on q_n in a Bayesian model for \mathcal{Q} , as the inverse-gamma is the conjugate prior distribution for the variance in a Gaussian likelihood function with known mean—meaning the posterior distribution for the variance is also inverse-gamma (reference 23). Examples of the inverse-gamma PDF for several values of shape and scale parameters $\{\alpha, \beta\}$ are shown in figure 1. For comparison, examples of the uniform PDF for several values of the support interval $[a, b]$ are also shown.

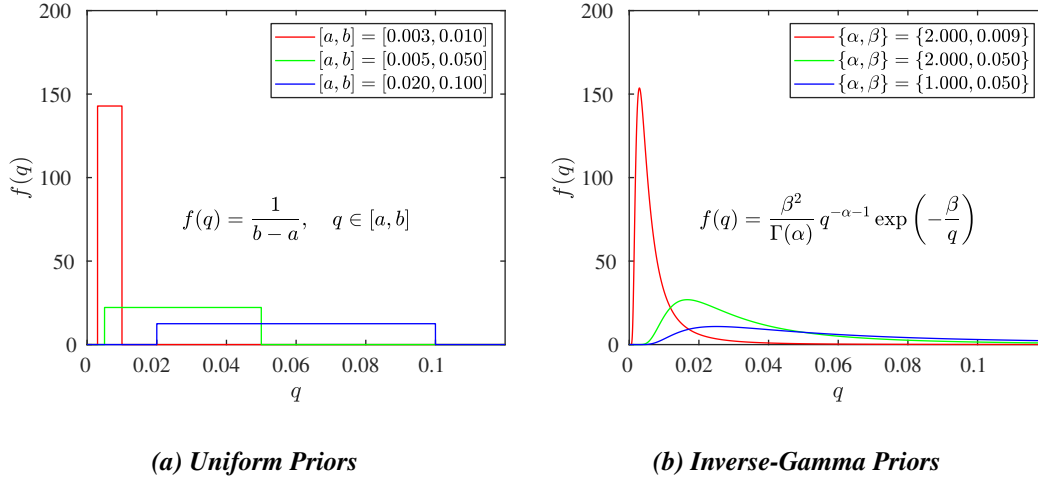


Figure 1. Uniform and Inverse-Gamma Priors for q for Three Values of the Lower/Upper Bounds $[a, b]$ and Shape/Scale Parameters $\{\alpha, \beta\}$, Respectively

2.4 ALGORITHM CONVERGENCE

The iterative estimation procedure for the process and measurement noise scales, \mathcal{Q} and \mathcal{R} , may be terminated when the differences between successive estimates fall below prescribed thresholds, ϵ_q and ϵ_r , respectively. With some abuse of notation, these differences can be written in terms of an arbitrary norm $\|\cdot\|$, and the sets \mathcal{Q} and \mathcal{R} can be interpreted as vectors. Then, the estimation procedure may be terminated at iteration ℓ^* , where

$$\ell^* = \min \ell : \|\mathcal{Q}^{(\ell+1)} - \mathcal{Q}^{(\ell)}\| < \epsilon_q \wedge \|\mathcal{R}^{(\ell+1)} - \mathcal{R}^{(\ell)}\| < \epsilon_r. \quad (59)$$

Alternatively, the estimation procedure may be terminated when the difference between successive values of the likelihood function, $p(Z|\mathcal{S})$, (or, equivalently, the log-likelihood function) falls below a prescribed threshold, $\epsilon_{\mathcal{L}}$, that is, when

$$\ell^* = \min \ell : \log p(Z|\mathcal{S}^{(\ell+1)}) - \log p(Z|\mathcal{S}^{(\ell)}) < \epsilon_{\mathcal{L}}. \quad (60)$$

The likelihood function $p(Z|\mathcal{S})$ is derived in reference 8 and is given in terms of the iterate, $\mathcal{S}^{(\ell)}$, by

$$p(Z|\mathcal{S}^{(\ell)}) = \prod_{n=1}^N \mathcal{N}(\nu_n^{(\ell)} | 0, \Sigma_n^{(\ell)}), \quad (61)$$

where the predicted measurement error (or innovation) vectors, $\nu_n^{(\ell)}$, are given by

$$\nu_n^{(\ell)} = z_n - B_n \mu_{n|n-1}^{(\ell)}, \quad (62)$$

and the predicted measurement error covariance matrices, $\Sigma_n^{(\ell)}$, by

$$\Sigma_n = r_n^{(\ell)} \bar{R}_n + B_n P_{n|n-1}^{(\ell)} B_n^T. \quad (63)$$

The predicted state estimates, $\mu_{n|n-1}^{(\ell)}$, and associated covariance matrices, $P_{n|n-1}^{(\ell)}$, in equations (62) and (63) are obtained from the forward pass of the Kalman smoother for the E-step of the ℓ th iteration of the algorithm. In practice, both equations (59) and (60) may be used in combination as termination criteria. In any event, because the incomplete-data log-PDF should increase monotonically, evaluation of the difference in equation (60) at each iteration is a valuable check against algorithm implementation.

3. CGMM APPROACH

The CGMM approach to maneuvering-target tracking is dual to the HGMM approach in that the roles of the target states and process and measurement noise covariance scales are swapped. In the CGMM approach, the covariance scales, \mathcal{Q} and \mathcal{R} , are treated as the missing data, and estimates of the target states, X , are the primary outputs. The CGMM approach and resulting estimation procedure for the target states are derived in the following subsections.

3.1 PROCESS AND MEASUREMENT NOISE MODELS

Whereas the HGMM approach assumes the process and measurement noise terms, w_n and v_n , in equations (1) and (2) are Gaussian distributed with PDFs given by equations (3) and (4) with unknown covariance scales, q_n and r_n , respectively, the CGMM approach assumes these noise terms are distributed according to the marginal PDFs

$$p(w_n) = \int_0^\infty \pi(q) \mathcal{N}(w_n|0, q\bar{Q}_n) dq, \quad (64)$$

$$p(v_n) = \int_0^\infty \chi(r) \mathcal{N}(v_n|0, r\bar{R}_n) dr, \quad (65)$$

where $\pi(q)$ and $\chi(r)$ are the prior PDFs for the process and measurement noise scales, respectively. The PDFs in equations (64) and (65) are continuous analogs of discrete Gaussian mixtures formed by weighted sums of the Gaussian PDFs given by equations (3) and (4) for finite sets of values for q_n and r_n , respectively. Hence, the PDFs in equations (64) and (65) are referred to as continuous Gaussian mixtures (CGMs), and the maneuvering-target tracking approach based on these models for process and measurement noise is referred to as the CGM model (CGMM) approach.* Plots of the CGMM for the measurement noise PDF given by equation (65) for several choices of the measurement noise scale prior PDF, $\chi(r)$, are shown in figure 2 for the case $M = 2$ and $\bar{R}_n = I_2$. Figures 2(a) through 2(b) show results for the Dirac delta prior,

$$\chi(r) = \delta(r - r^*), \quad r \in (0, \infty), \quad (66)$$

where $\delta(\cdot)$ denotes the Dirac delta function and r^* denotes the value for r at which all of its probability mass is fixed. It follows that for $\chi(r) = \delta(r - r_n)$, the CGMM for $p(v_n)$ reduces to

$$p(v_n) = \int_0^\infty \delta(r - r_n) \mathcal{N}(v_n|0, r\bar{R}_n) dr = \mathcal{N}(v_n|0, r_n\bar{R}_n), \quad (67)$$

the standard Gaussian PDF for $p(v_n)$ given by equation (4), with measurement covariance matrix, R_n , of the form given by equation (8). Figures 2(c) through 2(d) and figures 2(e) through 2(f)

*A related approach based on PMHT, called mixed-model PMHT, is developed in reference 18. In this approach, a finite set of $J > 1$ target motion models is assumed, and the conditional state densities, $p(x_n|x_{n-1})$, are modeled as discrete mixtures of these models. Loosely speaking, the CGMM approach is a limiting form of the mixed-model PMHT approach in the limit $J \rightarrow \infty$ for the linear-Gaussian case where the motion models differ only in the process noise covariance scales, $q_n(j)$, $j = 1, 2, \dots, J$, with $q_n(j) \in \{q(j) : q(j) \in (0, \infty), j = 1, 2, \dots, J\}$.

show results for the uniform and inverse-gamma priors given by equations (55) and (56), respectively. These priors for r result in measurement noise distributions (figures 2(d) and 2(f)) with heavier tails than the Gaussian distribution (figure 2(b)) associated with the standard Kalman filter model, making the CGMM approach robust to measurement outliers. The CGMM for the process noise PDF given by equation (64) shares the same characteristics as equation (65), though the dimension of w_n is typically greater than that of v_n (i.e., typically $L > M$). In this case, a uniform or inverse-gamma prior for the process noise scale, q , leads to a CGMM for $p(w_n)$ with heavier tails than the standard Gaussian distribution, making the CGMM approach robust to outliers in the target maneuver behavior.

3.2 COMPLETE-, INCOMPLETE-, AND MISSING-DATA PDFS

Estimation of the target states, X , from the observed data, Z , for the linear-CGMM model given by equations (1), (2), (10), (64), and (65) would be straightforward if observations of the process and measurement noise covariance scales, \mathcal{Q} and \mathcal{R} , were also available. In this case, maximizing the complete-data PDF, given by

$$\begin{aligned} p(X, Z, \mathcal{Q}, \mathcal{R}) &= p(X, Z | \mathcal{Q}, \mathcal{R}) p(\mathcal{Q}, \mathcal{R}) \\ &= p(Z | X, \mathcal{R}) p(X | \mathcal{Q}) p(\mathcal{Q}) p(\mathcal{R}), \end{aligned} \quad (68)$$

with respect to X would amount to a Kalman smoothing problem conditioned on Z , \mathcal{Q} , and \mathcal{R} , with the likelihood function, $p(Z | X, \mathcal{R})$, given by

$$p(Z | X, \mathcal{R}) = \prod_{n=1}^N p(z_n | x_n, r_n), \quad (69)$$

and the target state prior PDF, $p(X | \mathcal{Q})$, given by

$$p(X | \mathcal{Q}) = p(x_0) \prod_{n=1}^N p(x_n | x_{n-1}, q_n), \quad (70)$$

with the densities $p(x_0)$, $p(x_n | x_{n-1}, q_n)$, and $p(z_n | x_n, r_n)$ given by equations (10), (33), and (47), respectively. In reality, observations of \mathcal{Q} and \mathcal{R} are unavailable (or missing), and X must be estimated from the incomplete-data PDF, given by

$$p(X, Z) = \iint p(X, Z, \mathcal{Q}, \mathcal{R}) d\mathcal{Q} d\mathcal{R} \quad (71)$$

$$= \left[\int p(\mathcal{R}) p(Z | X, \mathcal{R}) d\mathcal{R} \right] \left[\int p(\mathcal{Q}) p(X | \mathcal{Q}) d\mathcal{Q} \right]. \quad (72)$$

The CGMM approach estimates the target states, X , from the observed data, Z , by finding the value of X that maximizes the incomplete-data PDF $p(X, Z)$; however, it does so by using the EM method to find values of X that maximize a sequence of simpler estimation problems involving the complete-data PDF, $p(X, Z, \mathcal{Q}, \mathcal{R})$, as described in the following subsection.

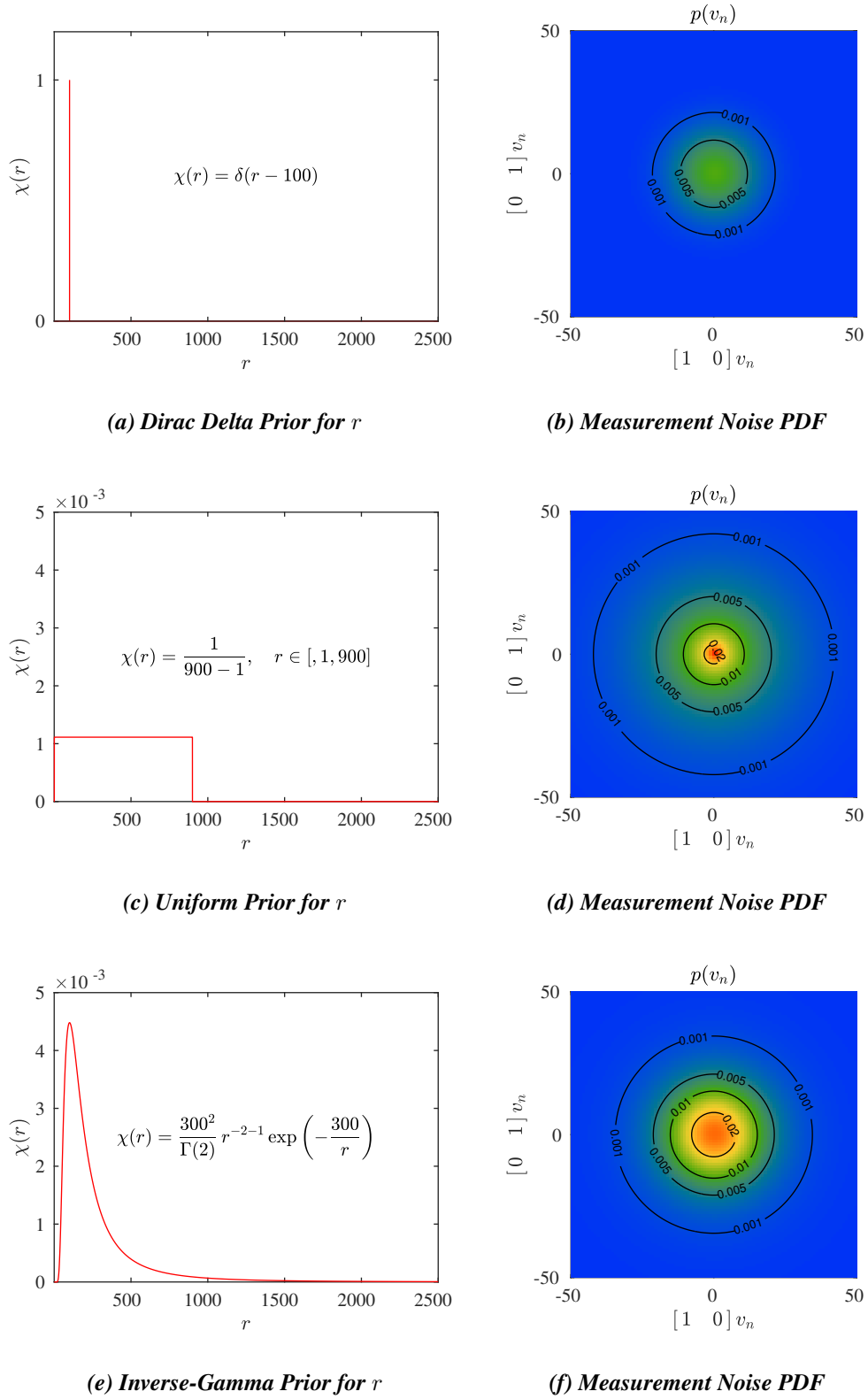


Figure 2. CGMs for the Measurement Noise PDF $p(v_n)$ for Three Different Priors for the Covariance Scale, r , in Equation (65), with $\bar{R}_n = I_2$

Before proceeding, it will be useful to derive the distribution of the missing data, \mathcal{Q} and \mathcal{R} , conditioned on X and Z , as this distribution is necessary for the E-step of the EM method. The conditional missing-data PDF is defined in terms of the complete-data and incomplete-data PDFs by

$$p(\mathcal{Q}, \mathcal{R}|X, Z) = \frac{p(X, Z, \mathcal{Q}, \mathcal{R})}{p(X, Z)}. \quad (73)$$

Assuming the missing process and measurement noise scales are independent, the prior PDFs for \mathcal{Q} and \mathcal{R} are given by the products

$$p(\mathcal{Q}) = \prod_{n=1}^N \pi(q_n), \quad (74)$$

$$p(\mathcal{R}) = \prod_{n=1}^N \chi(r_n). \quad (75)$$

Substituting equations (69), (70), (74), and (75) into equation (68) and rearranging terms yields the following expression for the complete-data PDF:

$$p(X, Z, \mathcal{Q}, \mathcal{R}) = p(x_0) \left[\prod_{n=1}^N \pi(q_n) p(x_n|x_{n-1}, q_n) \right] \left[\prod_{n=1}^N \chi(r_n) p(z_n|x_n, r_n) \right]. \quad (76)$$

Likewise, substituting equations (69), (70), (74), and (75) into equation (72) and rearranging terms yields the following expression for the incomplete-data PDF:

$$p(X, Z) = p(x_0) \left[\prod_{n=1}^N \int_0^\infty \pi(q) p(x_n|x_{n-1}, q) dq \right] \left[\prod_{n=1}^N \int_0^\infty \chi(r) p(z_n|x_n, r) dr \right]. \quad (77)$$

Finally, substituting equations (76) and (77) into equation (73) and collecting terms in gives the conditional missing-data PDF as

$$\begin{aligned} p(\mathcal{Q}, \mathcal{R}|X, Z) &= p(\mathcal{Q}|X)p(\mathcal{R}|X, Z) \\ &= \left[\prod_{n=1}^N \phi_n(q_n) \right] \left[\prod_{n=1}^N \psi_n(r_n) \right], \end{aligned} \quad (78)$$

where $\phi_n(q_n) \equiv p(q_n|x_n, x_{n-1})$ denotes the conditional PDF of the missing processing noise covariance scale, q_n , given by

$$\phi_n(q_n) = \frac{\pi(q_n) p(x_n|x_{n-1}, q_n)}{\int_0^\infty \pi(q) p(x_n|x_{n-1}, q) dq}, \quad (79)$$

and $\psi_n(r_n) \equiv p(r_n|z_n, x_n)$ denotes the conditional PDF of the missing measurement noise covariance scale, r_n , given by

$$\psi_n(r_n) = \frac{\chi(r_n) p(z_n|x_n, r_n)}{\int_0^\infty \chi(r) p(z_n|x_n, r) dr}. \quad (80)$$

3.3 STATE ESTIMATION

As for the HGMM approach, the EM method leads to an iterative procedure for the estimates of interest—in this case, those of the target states, X —that guarantees a monotonic increase in the incomplete-data PDF, $p(X, Z)$, and, under mild regularity conditions, convergence to at least a local maximum of $p(X, Z)$ at each iteration. As in section 2, let $\ell = 1, 2, \dots$ denote the EM iteration index, and let $X^{(\ell)}$ denote the target state estimate at the ℓ th EM iteration. Then the auxiliary function of the target states, X , at the beginning of the ℓ th iteration, denoted $\Theta(X|X^{(\ell)})$, is given by

$$\Theta(X|X^{(\ell)}) = \text{E} [\log p(X, Z, \mathcal{Q}, \mathcal{R})|X^{(\ell)}, Z] \quad (81)$$

$$= \iint [\log p(X, Z, \mathcal{Q}, \mathcal{R})] p(\mathcal{Q}, \mathcal{R}|X^{(\ell)}, Z) d\mathcal{Q}d\mathcal{R}, \quad (82)$$

and the updated estimate for X at the ℓ th iteration by

$$X^{(\ell+1)} = \arg \max_X \Theta(X|X^{(\ell)}). \quad (83)$$

Substituting equations (76) and (78) into equation (82), evaluating the logarithm and simplifying the resulting expression yields the following general form for the auxiliary function:

$$\begin{aligned} \Theta(X|X^{(\ell)}) = & \log p(x_0) + \left[\sum_{n=1}^N \int_0^\infty \phi_n^{(\ell)}(q) \log p(x_n|x_{n-1}, q) dq \right] \\ & + \left[\sum_{n=1}^N \int_0^\infty \psi_n^{(\ell)}(r) \log p(z_n|x_n, r) dr \right], \end{aligned} \quad (84)$$

with $\phi_n^{(\ell)}(q) = p(q|x_n^{(\ell)}, x_{n-1}^{(\ell)})$ and $\psi_n^{(\ell)}(r) = p(r|z_n, x_n^{(\ell)})$. For the complete-data, linear-Gaussian model, the conditional state and measurement densities in equation (84) are given by equations (33) and (47), repeated here for convenience:

$$p(x_n|x_{n-1}, q) = \mathcal{N}(x_n|A_n x_{n-1}, q\bar{Q}_n), \quad (85)$$

$$p(z_n|x_n, r) = \mathcal{N}(z_n|B_n x_n, r\bar{R}_n). \quad (86)$$

Substituting the Gaussian PDFs given by equations (85) and (86) into equation (84), evaluating the logarithms, simplifying the resulting expression, and dropping terms not dependent on the target states, X , gives the auxiliary function at the ℓ th iteration in simplified form:

$$\begin{aligned} \Theta(X|X^{(\ell)}) = & -\frac{1}{2}(x_0 - \mu_0)^T P_0^{-1}(x_0 - \mu_0) \\ & -\frac{1}{2} \left[\sum_{n=1}^N (x_n - A_n x_{n-1})^T (\tilde{Q}_n^{(\ell)})^{-1} (x_n - A_n x_{n-1}) \right] \\ & -\frac{1}{2} \left[\sum_{n=1}^N (z_n - B_n x_n)^T (\tilde{R}_n^{(\ell)})^{-1} (z_n - B_n x_n) \right], \end{aligned} \quad (87)$$

where $\tilde{Q}_n^{(\ell)}$ is the effective process noise covariance matrix for the target state estimate at the ℓ th iteration, given by

$$\tilde{Q}_n^{(\ell)} = \tilde{q}_n^{(\ell)} \bar{Q}_n, \quad (88)$$

with

$$\tilde{q}_n^{(\ell)} = \left[\int_0^\infty \frac{\phi_n^{(\ell)}(q)}{q} dq \right]^{-1}, \quad (89)$$

and $\tilde{R}_n^{(\ell)}$ is the effective measurement noise covariance matrix at the ℓ th iteration, given by

$$\tilde{R}_n^{(\ell)} = \tilde{r}_n^{(\ell)} \bar{R}_n, \quad (90)$$

with

$$\tilde{r}_n^{(\ell)} = \left[\int_0^\infty \frac{\psi_n^{(\ell)}(r)}{r} dr \right]^{-1}. \quad (91)$$

The conditional missing-data PDFs $\phi_n^{(\ell)}(q)$ and $\psi_n^{(\ell)}$ required for evaluating the effective process and measurement noise covariance scales given by equations (89) and (91) for the linear-Gaussian model are obtained by substituting equations (85) and (86) into equations (79) and (80), yielding

$$\phi_n^{(\ell)}(q) = \frac{\pi(q) \mathcal{N}(x_n^{(\ell)} | A_n x_{n-1}^{(\ell)}, q \bar{Q}_n)}{\int_0^\infty \pi(\tau) \mathcal{N}(x_n^{(\ell)} | A_n x_{n-1}^{(\ell)}, \tau \bar{Q}_n) d\tau}, \quad (92)$$

and

$$\psi_n^{(\ell)}(r) = \frac{\chi(r) \mathcal{N}(z_n | B_n x_n^{(\ell)}, r \bar{R}_n)}{\int_0^\infty \chi(\tau) \mathcal{N}(z_n | B_n x_n^{(\ell)}, \tau \bar{R}_n) d\tau}. \quad (93)$$

The simplified form of the auxiliary function, $\Theta(X|X^{(\ell)})$, in equation (87) is functionally equivalent to the posterior log-PDF for the standard Kalman smoothing problem for the linear-Gaussian system with states X , measurements Z , and system matrices $\{A_n, \tilde{Q}_n^{(\ell)}, B_n, \tilde{R}_n^{(\ell)} : n = 1, 2, \dots, N\}$. Hence, the target state estimate at the ℓ th EM iteration, $X^{(\ell)}$, may be obtained by standard Kalman smoothing techniques.

Though the integrals required by the calculations of the effective process and measurement noise covariance scales in equations (89) and (91) look formidable, they are one-dimensional integrals that may be efficiently evaluated by standard numerical techniques (reference 24). Moreover in at least two cases of practical interest shown in the appendix, the integrals required by the CGMM approach may be evaluated in closed-form in terms of well-known special functions.

3.4 ALGORITHM CONVERGENCE

As for the HGMM approach, the iterative estimation procedure for the target states may be terminated when the difference between successive estimates falls below a prescribed threshold, ϵ_X . Specifically, interpreting the set of states, X , as a concatenated vector, the estimation procedure may be terminated at iteration ℓ^* , where

$$\ell^* = \min \ell : \|X^{(\ell+1)} - X^{(\ell)}\| < \epsilon_X. \quad (94)$$

Alternatively, the estimation procedure may be terminated when the difference between successive values of the incomplete-data PDF, $p(X, Z)$, or, equivalently, the incomplete-data log-PDF, falls below a prescribed threshold, $\epsilon_{\mathcal{L}}$, that is, when

$$\ell^* = \min \ell : \log p(X^{(\ell+1)}, Z) - \log p(X^{(\ell)}, Z) < \epsilon_{\mathcal{L}}. \quad (95)$$

The incomplete-data log-PDF at the ℓ th iteration is obtained by substituting equations (85) and (86) into equation (77) and evaluating the logarithm of the resulting expression at $X = X^{(\ell)}$:

$$\begin{aligned} \log p(X^{(\ell)}, Z) = & \log \mathcal{N}(x_0^{(\ell)} | \mu_0, P_0) + \left[\sum_{n=1}^N \log \int_0^\infty \pi(q) \mathcal{N}(x_n^{(\ell)} | A_n x_{n-1}^{(\ell)}, q \bar{Q}_n) dq \right] \\ & + \left[\sum_{n=1}^N \log \int_0^\infty \chi(r) \mathcal{N}(z_n | B_n x_n^{(\ell)}, r \bar{R}_n) dr \right]. \end{aligned} \quad (96)$$

In practice, both equations (94) and (95) may be used in combination as termination criteria. Again, as for the HGMM approach, evaluation of the difference in equation (95) at each iteration is a valuable check against algorithm implementation because the incomplete-data log-PDF should increase monotonically with each iteration.

4. EXAMPLE

This section compares the HGMM and CGMM approaches to maneuvering-target tracking via a simple simulated example of a target that executes a 90-degree maneuver over 2 minutes in between two constant-velocity legs, each 10 minutes in length. For both approaches, an inverse-gamma prior and a uniform prior are chosen for the measurement and process noise covariance scales, respectively. These priors are shown in figure 3.

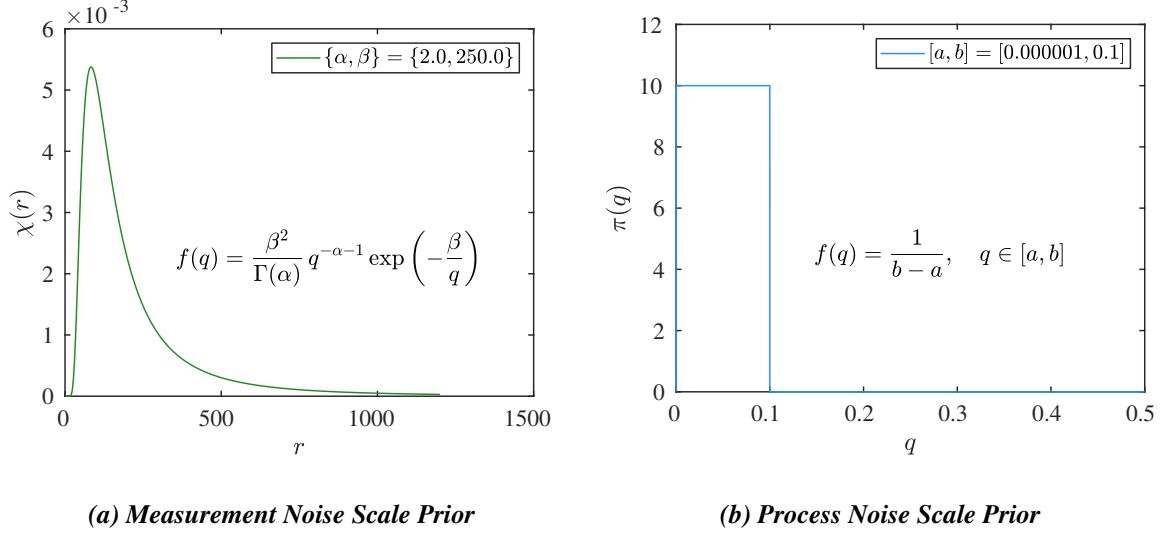


Figure 3. Measurement and Process Noise Covariance Scale Priors Used in Example

Figures 4 through 11 are shown at the end of the section to facilitate comparisons. The true target track is shown by the solid red line in the left-hand plot of figure 4. The target initially heads due north at 5 knots until maneuvering to starboard at 45-degrees/minute and then continues due east at the same speed. Independent measurements of target position, shown as red dots in figure 4, are obtained every 10 seconds, where the error in each component of each measurement is normally distributed with zero mean and a standard deviation of 10 yards. Furthermore, 10 measurements are chosen at random as outliers, each uniformly distributed in a square box with side lengths 200 yards, centered at the true target position; the outlier measurements are identified with red circles in figure 4.

The outputs of the HGMM approach for several EM iterations are shown in figures 4 through 10. The right-hand plots of figure 4 show the initial values used for the measurement and process noise scales, $r_n^{(0)}$ and $q_n^{(0)}$, respectively, to start the iterations. Values of $r_n^{(0)}$ associated with outlier measurements are identified with circles in the top right-hand plot. In the bottom right-hand plot, two red triangles are placed on the measurement index axis to denote the beginning and end of the target maneuver. Figure 5 shows the output of the HGMM approach after the first iteration. While the estimated target track, shown as the solid black line the

left-hand plot, does a reasonable job of fitting the data during the constant-velocity legs, it does a poor job of fitting the data during the maneuver. This result is due to the small initial, constant values of the process noise scales used to estimate this track, shown in the lower right-hand plot in figure 4. In fact, the track estimate for the first iteration of the HGMM approach is the output of the standard Kalman smoother for this problem for the fixed measurement and process noise covariance scales shown in the right-hand plots of figure 4. However, the right-hand plots of figure 5 show the estimates of the measurement and process noise scales to be used for the track estimate at the next iteration, shown in the left-hand plot of figure 6. As shown by the top right-hand plot in figure 5, the HGMM approach is able to identify the outliers via inflated measurement noise scale estimates after one iteration. Likewise, as shown by the lower right-hand plot, the approach is able to identify the target maneuver via inflated process noise scale estimates. Both sets of noise scale estimates lead to a better track estimate at the next iteration as shown in the left-hand plot of figure 6. Indeed, the inflated measurement noise scale estimates for the outliers lead to better constant-velocity leg estimates, while the inflated process noise scale estimates around the maneuver lead to a better estimate of the constant turn-rate trajectory.

Figures 7 through 10 show the improvements in the HGMM track estimate and the evolution of the measurement and process noise scale estimates for several additional iterations. Interestingly, while the track estimate appears to converge after relatively few iterations (around 8), estimates of the noise scales—particularly the process noise scales—continue to noticeably change over many iterations, though the magnitudes of the changes decrease with each iteration. This behavior is typical of EM algorithms, which tend to take large steps toward the nearest local solution after a few iterations, only to then experience increasingly slower convergence with each iteration. Methods for accelerating EM convergence are surveyed reference 12.

For comparison, outputs for the final iteration of the CGMM approach are shown in figure 11. The right-hand plots show the effective measurement and process noise scales used to establish the track estimate shown by the solid black line in the left-hand plot, given by equations (89) and (91), respectively. While the CGMM effective measurement noise scales for the outlier measurements in the top right-hand plot of figure 11 are noticeably larger than the corresponding HGMM measurement noise scale estimates in the top right-hand plot of figure 10, the trends in these quantities are nearly identical. However, the changes in the CGMM effective process noise scales around the maneuver, shown in the bottom right-hand plot of figure 11, are much more abrupt than the more gradual changes in the corresponding HGMM process noise scale estimates shown in the bottom right-hand plot of figure 10. Indeed, the CGMM effective process noise scales are effectively flat for the majority of the scenario, except for two spikes near the endpoints of the maneuver interval—a behavior reminiscent of bang-bang control. In contrast, the HGMM process noise scale factors exhibit a more gradual ramp-up/ramp-down around these two points. A more detailed comparison of the behavioral characteristics of the two approaches, along with thorough performance evaluations, are left for future work.

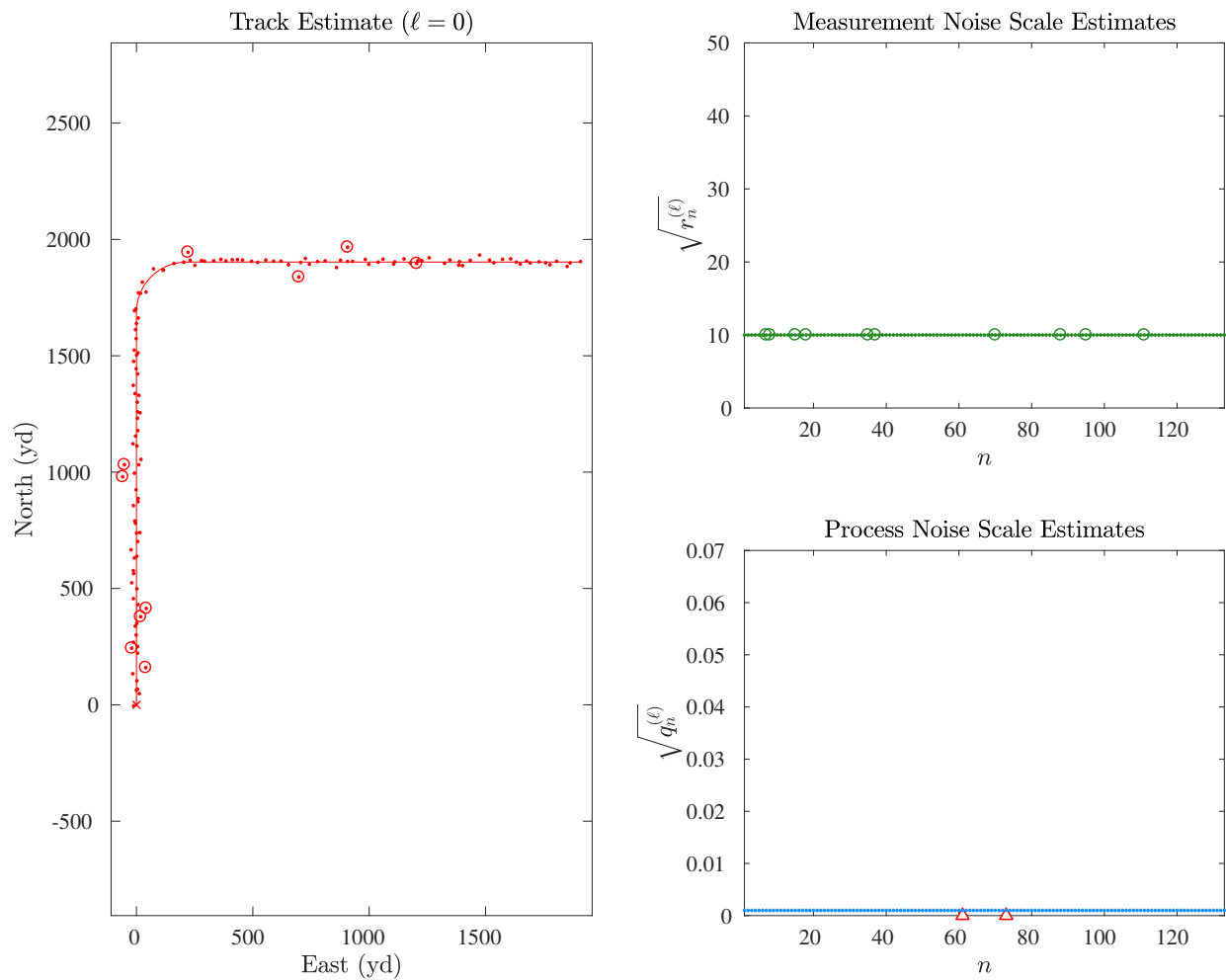


Figure 4. True Track and Measurements (Left), Measurement Noise Scale Initializations (Top Right), and Process Noise Scale Initializations (Bottom Right) at Iteration $\ell = 0$ for HGMM Approach—Measurement Outliers are Circled

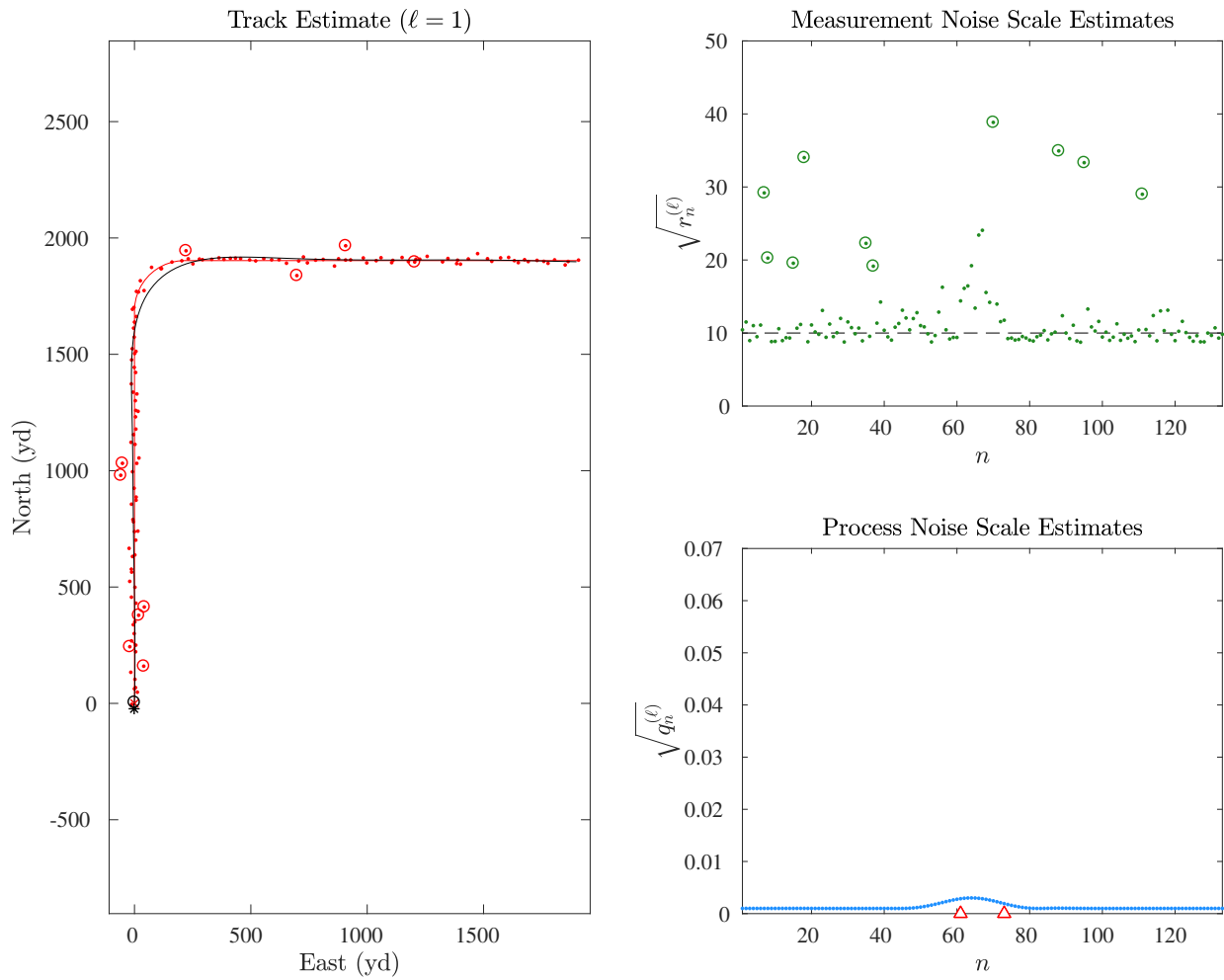


Figure 5. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 1$

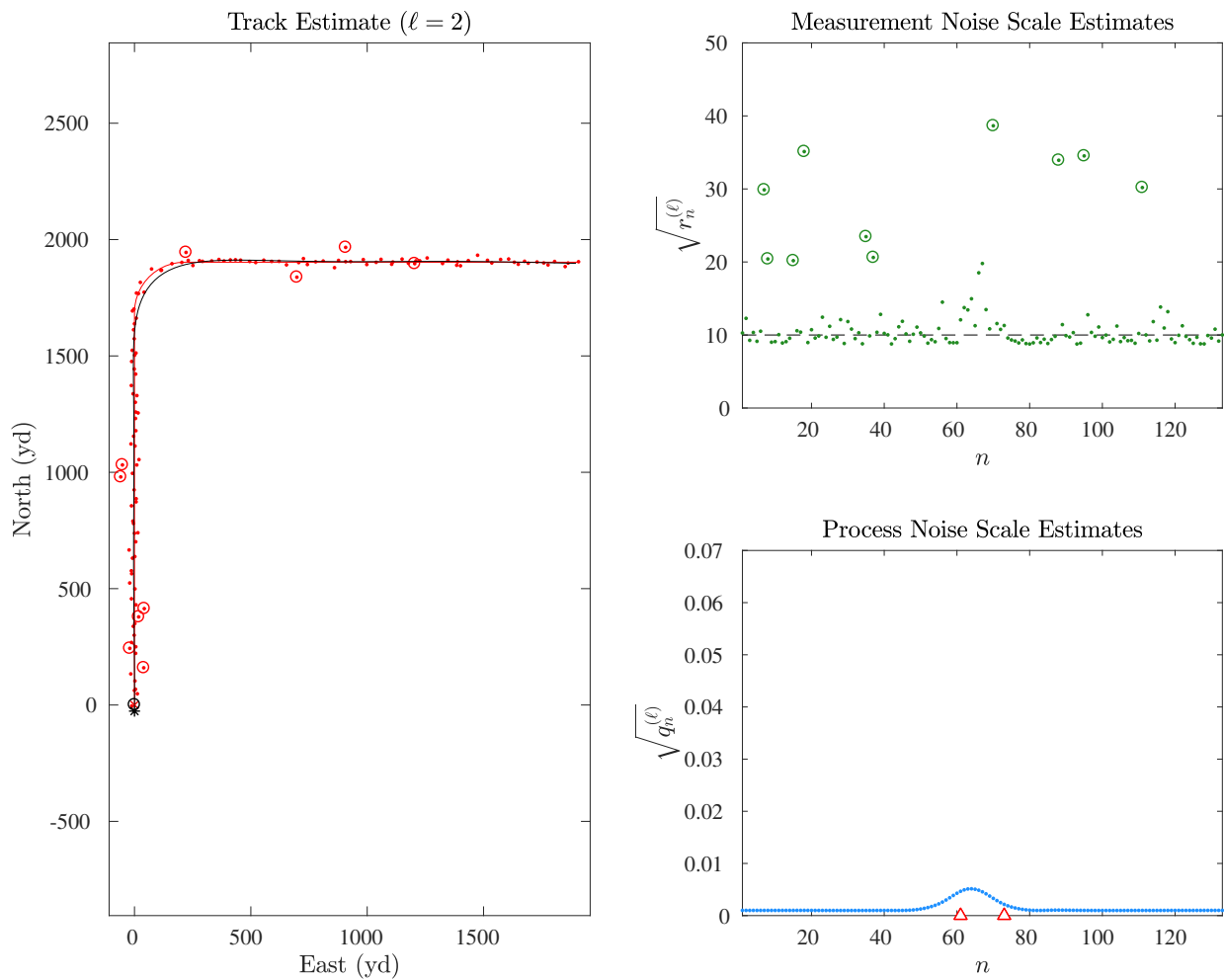


Figure 6. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 2$

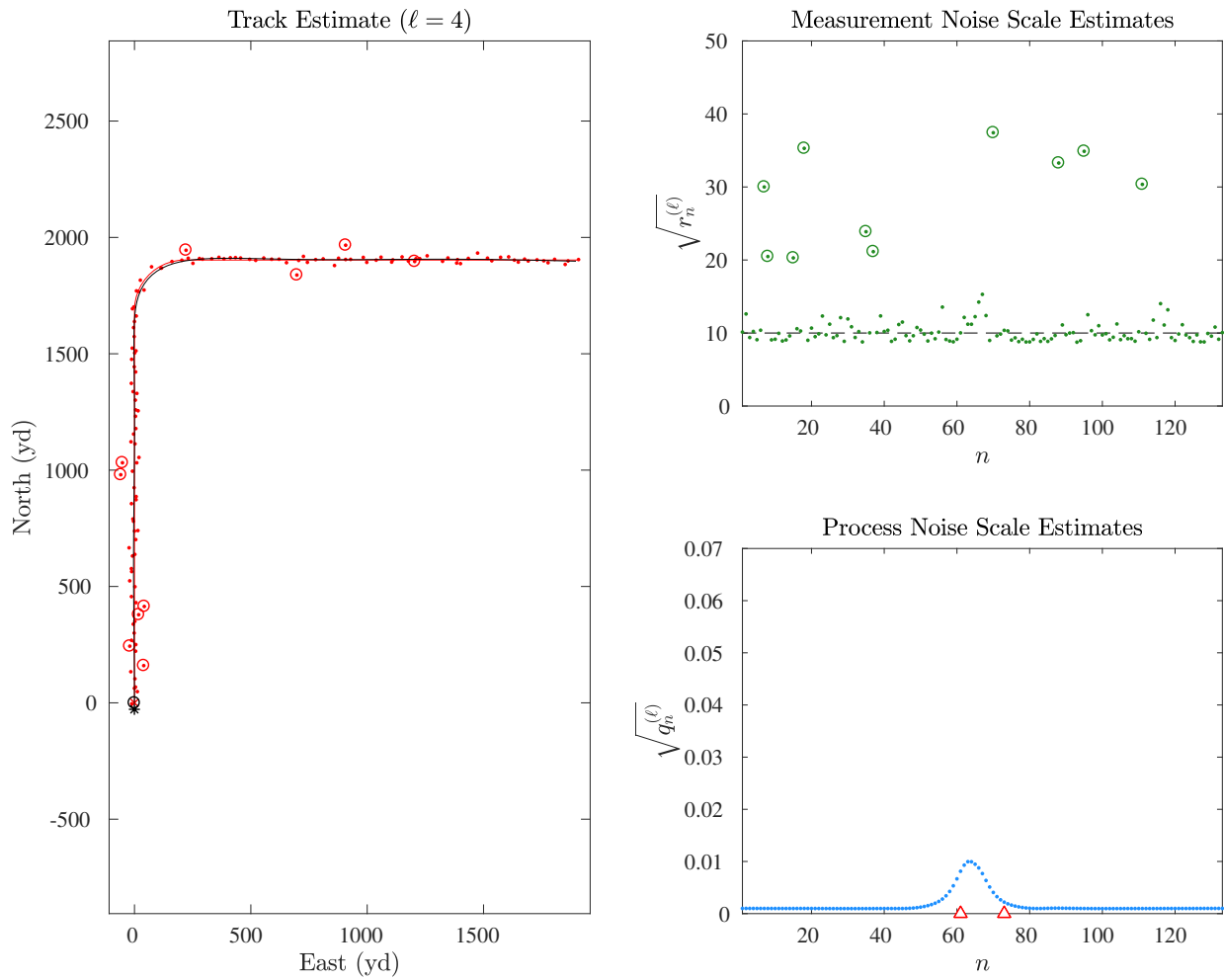


Figure 7. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 4$

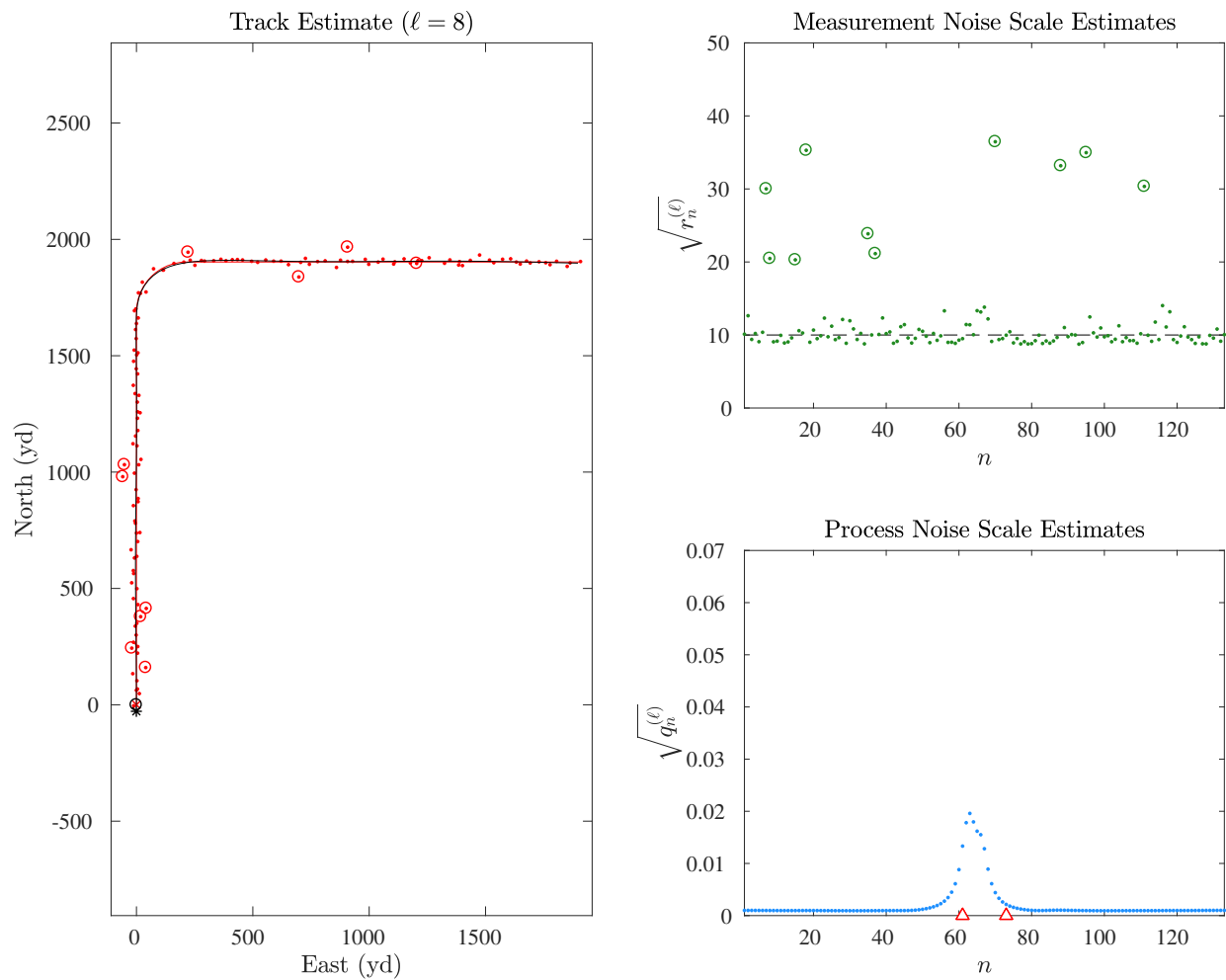


Figure 8. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 8$

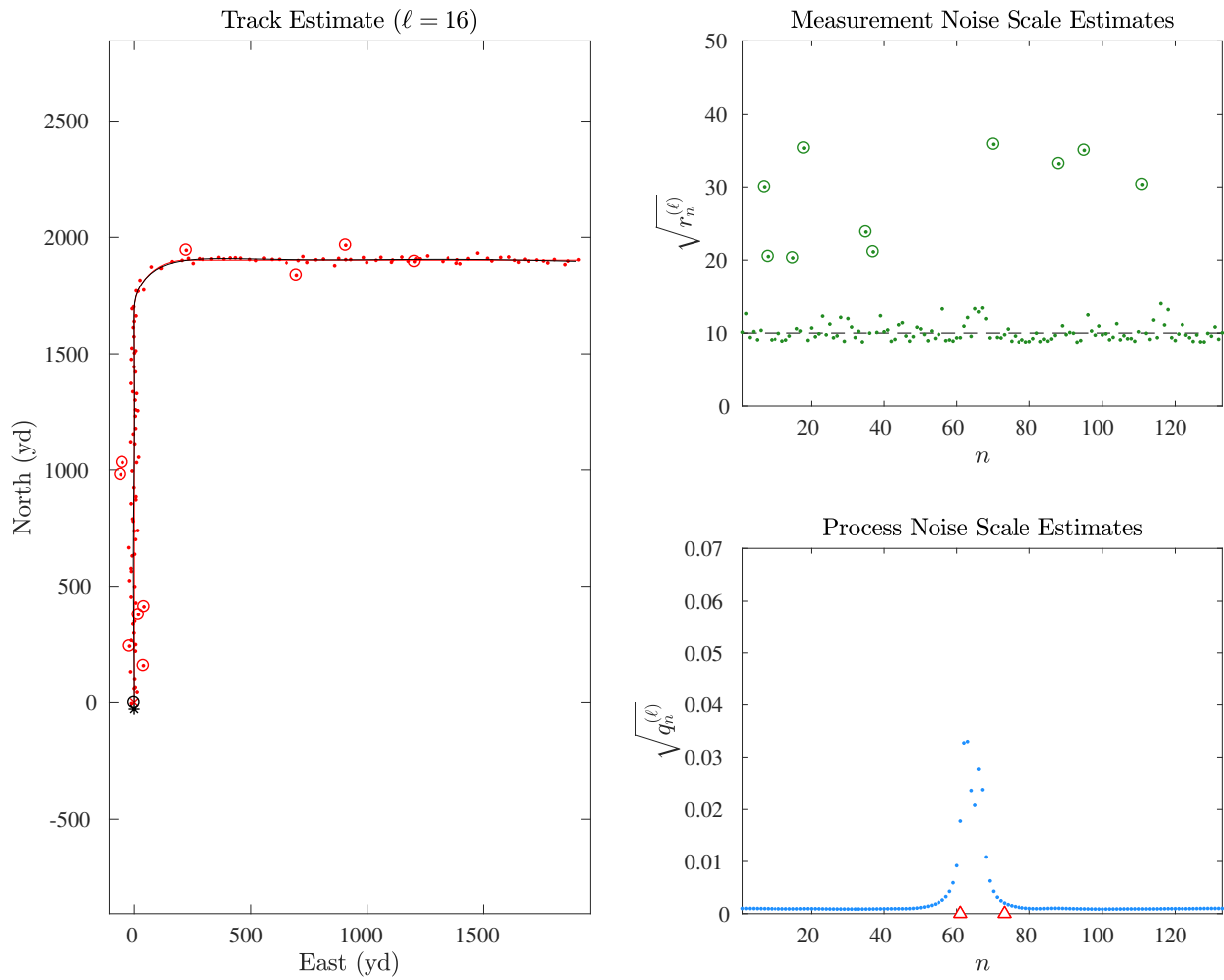


Figure 9. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 16$

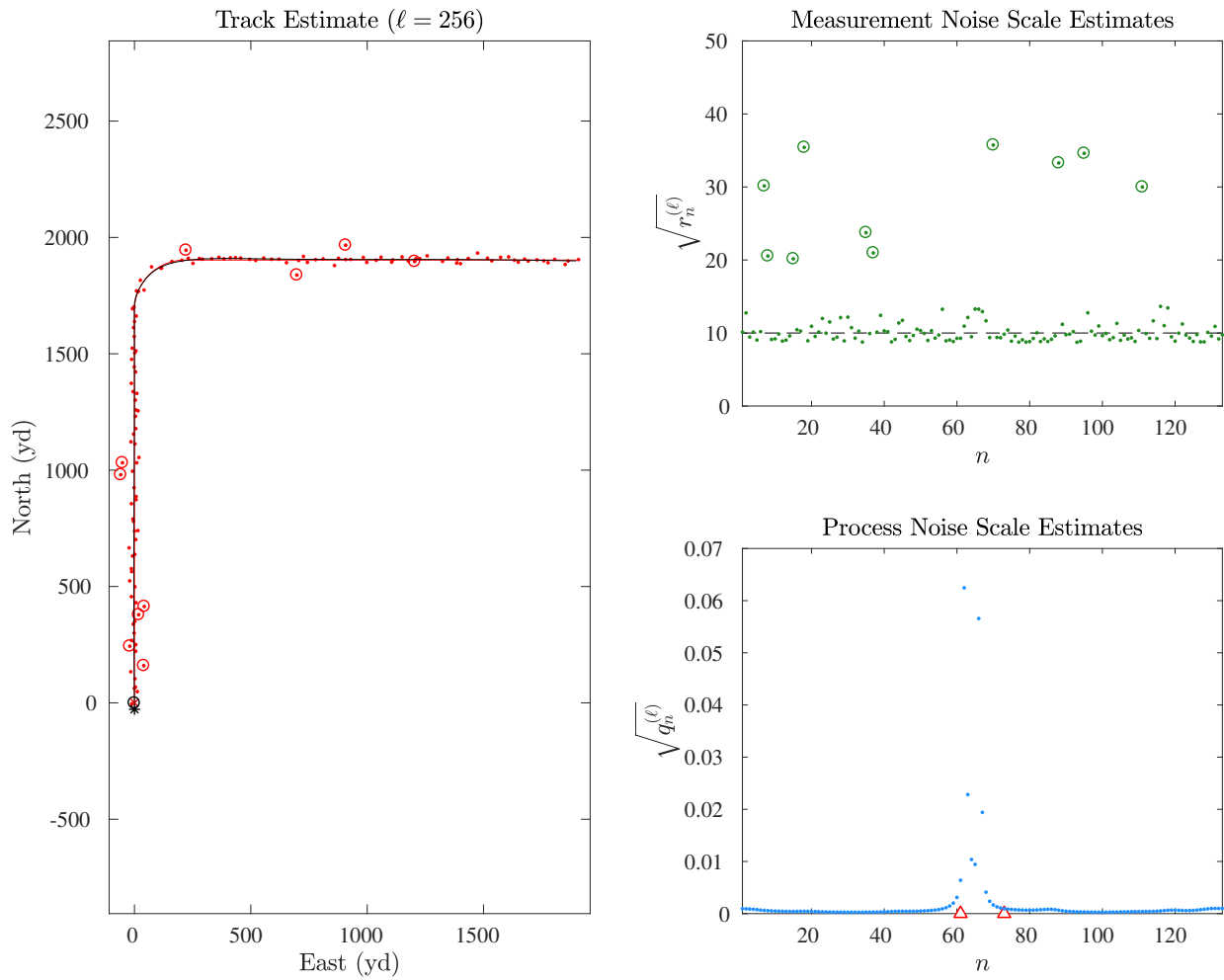


Figure 10. True Track and Measurements (Left, Red), HGMM Track Estimate (Left, Black), Measurement Noise Scale Estimates (Top Right), and Process Noise Scale Estimates (Bottom Right) at Iteration $\ell = 256$

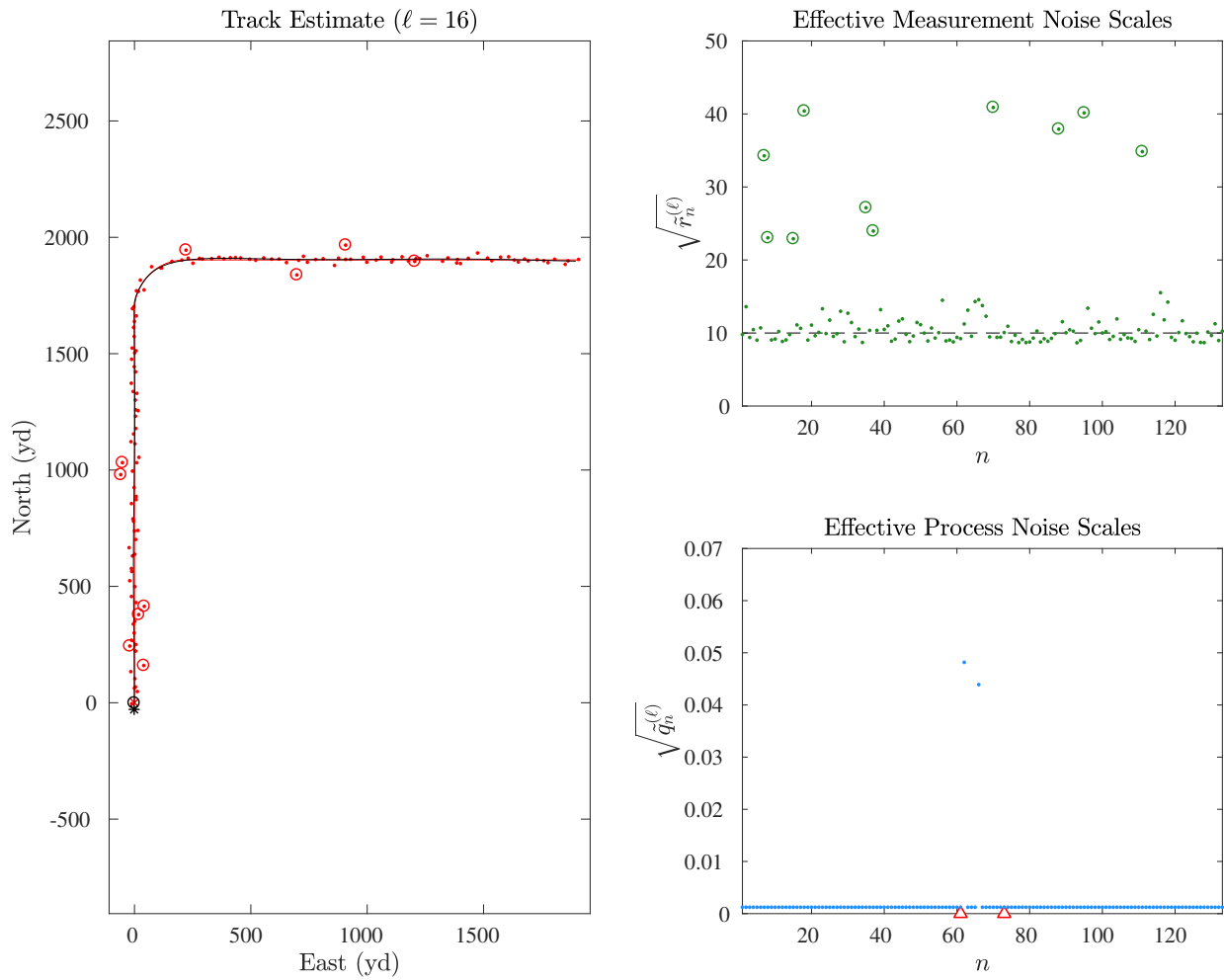


Figure 11. True Track and Measurements (Left, Red), CGMM Track Estimate (Left, Black), Effective Measurement Noise Scales (Top Right), and Effective Process Noise Scales (Bottom Right) at Iteration $\ell = 16$

5. CONCLUDING REMARKS AND FUTURE WORK

Two complementary approaches to the long-term, maneuvering-target tracking problem are derived in this report. The first approach, based on the theory of hidden Gauss-Markov models, treats the target states as missing data and estimates the process and measurement noise scales at each data scan. The second approach treats the process and measurement noise scales as missing data in the context of continuous Gaussian mixture models for the noise terms, and estimates the target states at each data scan. Both approaches are robust to changes in target maneuver behavior and measurement noise characteristics from scan-to-scan by virtue of an effective continuum of noise models applied at each scan. In the HGMM approach, robustness comes from the ability of the HGMM to estimate different process and measurement noise scales within the standard Gaussian noise models at each scan. Robustness in the CGMM approach comes from the ability of the CGMMs for the process and measurement noises at each scan to accommodate non-Gaussian (i.e., heavy-tailed) behavior. Both approaches reduce to iterative Kalman smoothing procedures for the target tracks; batch processing allows for optimal state estimates at all scans based on all available data.

This report focused on the theory and derivation of the HGMM and CGMM approaches. Future work will include performance evaluations based on numerical experiments. Performance in terms of tracking error and consistency will be emphasized, where the consistency of an approach is a measure of its estimate of track error—as quantified by its estimated state covariance matrices—versus actual track error. State covariance matrices for the HGMM and CGMM target state estimates are not derived in this report; these derivations are left for future work. For the HGMM approach, it is reasonable to expect the state covariance matrices to be given by the covariance matrices, $P_{n|N}^{(\ell^*)}$, in equation (30) obtained from the E-step of the final (ℓ^* th) EM iteration, conditioned on the system parameters $\mathcal{S}(\mathcal{Q}^{(\ell^*)}, \mathcal{R}^{(\ell^*)})$. However, it remains to demonstrate that these state covariance estimates are indeed consistent. For the CGMM approach, state covariance matrices may be obtained in a variety of ways (see reference 12, chapter 4). Future work will investigate the observed information approach discussed in references 25 and 26, which forms the covariance matrix from the inverse of the difference between the complete-data and missing-data information matrices. Based on the results of reference 26, it is reasonable to anticipate the state covariance matrices obtained from the Kalman smoother for the final iteration of the CGMM approach are the bases for the complete-data information matrix for this problem.

Future work will also focus on practical implementations of HGMM and CGMM batch algorithms for long-term maneuvering-target tracking. As derived here, both approaches assume a fixed batch length with a fixed number (N) of data scans. This requires N scans of data to accumulate before the algorithms can produce a track estimates. To eliminate this lag, both approaches can be implemented as expanding batches, where each new data scan is added to the leading edge of the batch, and the entire batch is reprocessed to obtain optimally smoothed estimates at each scan in the batch. Outputs from the previous batch can then be used to help initialize the estimation procedure for the next batch, and potentially greatly reduce the number of EM iterations required for convergence. Though the computational complexity of both algorithms is linear in the number of data scans, expanding the batch indefinitely may be impractical for

problems with very long time scales, high data rates, or insufficient computational resources. In these cases, the expanding batch may be switched to a sliding batch, where each new data scan is added to the leading edge of the batch, while the oldest scan is dropped from the batch. Again, outputs from the previous batch can then be used to help initialize the estimation procedure for the current batch, and potentially greatly reduce the number of EM iterations required for convergence.

Finally, future work will address the application of the converted measurement techniques developed in reference 20 to the HGMM and CGMM approaches for tracking problems with nonlinear measurement equations. As presented in reference 20, these techniques are applied to the sequential filtering problem, though their extension to the batch smoothing problem addressed here would appear to be straightforward. Converted measurements are correlated, yielding measurement covariance matrices that are non-diagonal. However, both the HGMM and CGMM approaches are consistent with the noise model of the converted measurement techniques because these approaches make no assumptions on the structure of the unscaled measurement covariance matrices.

APPENDIX

CLOSED-FORM EXPRESSIONS FOR CGMM INTEGRALS

The integrals required by the CGMM approach are one-dimensional integrals that may be efficiently evaluated by standard numerical techniques such as those developed in reference 24. A single, simplified expression for these integrals is derived in this appendix. Furthermore, in at least two cases of practical interest, it is shown this integral may be evaluated in closed-form in terms of well-known special functions.

Upon substituting equation (92) into equation (89) and rearranging terms, the integral required by the calculation of effective process noise scales, \tilde{q}_n , in equation (89) takes the form

$$\int_0^\infty \frac{\phi_n(q)}{q} dq = \frac{1}{\int_0^\infty \pi(\tau) \mathcal{N}(x_n | A_n x_{n-1}, \tau \bar{Q}_n) d\tau} \int_0^\infty \frac{\pi(q)}{q} \mathcal{N}(x_n | A_n x_{n-1}, q \bar{Q}_n) dq. \quad (\text{A-1})$$

Let η_n be defined as the quadratic form

$$\eta_n = \frac{1}{2} \|x_n - A_n x_{n-1}\|_{\bar{Q}_n^{-1}}^2 \quad (\text{A-2})$$

$$\equiv \frac{1}{2} (x_n - A_n x_{n-1})^T \bar{Q}_n^{-1} (x_n - A_n x_{n-1}). \quad (\text{A-3})$$

Then the Gaussian PDF in equation (A-1) may be written in terms of η_n as

$$\mathcal{N}(x_n | A_n x_{n-1}, q \bar{Q}_n) = \frac{c_n}{q^{L/2}} \exp\left(-\frac{\eta_n}{q}\right), \quad (\text{A-4})$$

where c_n is an irrelevant constant (independent of q). Substituting equation (A-4) into equation (A-1) and simplifying terms yields

$$\int_0^\infty \frac{\phi_n(q)}{q} dq = \frac{1}{\int_0^\infty \frac{\pi(\tau)}{\tau^{L/2}} \exp\left(-\frac{\eta_n}{\tau}\right) d\tau} \int_0^\infty \frac{\pi(q)}{q^{L/2+1}} \exp\left(-\frac{\eta_n}{q}\right) dq. \quad (\text{A-5})$$

From this expression, it follows the integrals in equation (A-1) take the following general form:

$$I(k, \varepsilon; f) \equiv \int_0^\infty \frac{f(\tau)}{\tau^k} \exp\left(-\frac{\varepsilon}{\tau}\right) d\tau, \quad (\text{A-6})$$

for $k \in \{j/2 : j = 1, 2, \dots\}$, $\varepsilon \geq 0$, and f a non-negative function on $(0, \infty)$. Using the results of equations (A-1) through (A-6), the effective process noise covariance scales given by equation (89) can be written simply as

$$\tilde{q}_n^{(\ell-1)} = \frac{I\left(\frac{L}{2}, \eta_n^{(\ell-1)}; \pi\right)}{I\left(\frac{L}{2} + 1, \eta_n^{(\ell-1)}; \pi\right)}, \quad (\text{A-7})$$

with

$$\eta_n^{(\ell-1)} = \frac{1}{2} \|x_n^{(\ell-1)} - A_n x_{n-1}^{(\ell-1)}\|_{\bar{Q}_n^{-1}}^2. \quad (\text{A-8})$$

Likewise, using a similar argument it is straightforward to show the effective measurement noise covariance scales given by equation (91) can be written simply as

$$\tilde{r}_n^{(\ell-1)} = \frac{I\left(\frac{M}{2}, \zeta_n^{(\ell-1)}; \chi\right)}{I\left(\frac{M}{2} + 1, \zeta_n^{(\ell-1)}; \chi\right)}, \quad (\text{A-9})$$

with

$$\zeta_n^{(\ell-1)} = \frac{1}{2} \|z_n - B_n x_n^{(\ell-1)}\|_{\bar{R}_n^{-1}}^2. \quad (\text{A-10})$$

Lastly, the general integral expression given by equation (A-6) can be used to simplify the expression for the incomplete-data log-PDF in equation (96). Indeed, expanding the expressions for the Gaussian PDFs in equation (96), using the expressions in equations (A-6), (A-8), and (A-10), and dropping terms that do not depend on the target state estimates, $X^{(\ell)}$, yields

$$\begin{aligned} \log p(X^{(\ell)}, Z) &= \frac{1}{2} \|x_0^{(\ell)} - \mu_0\|_{P_0^{-1}}^2 + \left[\sum_{n=1}^N \log I\left(\frac{L}{2}, \eta_n^{(\ell)}; \pi\right) \right] \\ &\quad + \left[\sum_{n=1}^N \log I\left(\frac{M}{2}, \zeta_n^{(\ell)}; \chi\right) \right]. \end{aligned} \quad (\text{A-11})$$

The integral $I(k, \varepsilon; f)$ may be evaluated in closed-form in terms of well-known special functions in at least two cases of practical interest. Specifically, in the case where f is the inverse-gamma PDF given by equation (56), with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, the integral in equation (A-6) is given simply in terms of the gamma function as

$$\begin{aligned} I(k, \varepsilon; f) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{1}{\tau^{k+\alpha+1}} \exp\left(-\frac{\varepsilon + \beta}{\tau}\right) d\tau, \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} (\varepsilon + \beta)^{-(k+\alpha)} \Gamma(k + \alpha). \end{aligned} \quad (\text{A-12})$$

In the case where f is the uniform PDF given by equation (55), defined on the closed interval $[a, b] \subset (0, \infty)$, the closed-form of the integral $I(k, \varepsilon; f)$ depends on the value of k . Specifically, in this case,

$$\begin{aligned} I(k, \varepsilon; f) &= \frac{1}{b-a} \int_a^b \frac{1}{\tau^k} \exp\left(-\frac{\varepsilon}{\tau}\right) d\tau, \\ &= \frac{I_b(k, \varepsilon) - I_a(k, \varepsilon)}{b-a}, \end{aligned} \quad (\text{A-13})$$

with

$$I_\tau(k, \varepsilon) = \begin{cases} 2\sqrt{\pi\varepsilon} \operatorname{erf}(\varepsilon/\tau) + 2\sqrt{\tau} \exp(-\varepsilon/\tau), & k = 1/2, \\ \operatorname{Ei}(\varepsilon/\tau), & k = 1, \\ \Gamma(k - 1, \varepsilon/\tau)/\varepsilon^{k-1}, & k > 1, \end{cases} \quad (\text{A-14})$$

where $\operatorname{erf}(\cdot)$ denotes the error function, $\operatorname{Ei}(\cdot)$ the exponential integral, and $\Gamma(\cdot, \cdot)$ the incomplete gamma function defined in chapters 7, 5, and 6, respectively, of reference 27. To avoid numerical instabilities, care must be taken when evaluating the expressions in equation (A-14) at values of ε near zero. In practice, a linear approximation gives accurate values of the integral for values of ε in the region $0 < \varepsilon < \rho(k) \ll 1$:

$$I_\tau(k, \varepsilon) \approx \frac{I_\tau(k, \rho(k)) - I_\tau(k, 0)}{\rho(k)} \varepsilon + I_\tau(k, 0), \quad (\text{A-15})$$

with

$$I_\tau(k, 0) \equiv \begin{cases} \log \tau, & k = 1, \\ \frac{1}{1-k} \tau^{1-k}, & \text{otherwise.} \end{cases} \quad (\text{A-16})$$

For particular values of k , which will be problem specific but fixed, the upper bounds $\rho(k)$ may be determined by comparing plots of equations (A-14) and (A-15) and identifying the values of ε at which equation (A-14) becomes unstable.

REFERENCES

1. H. A. P. Blom and Yaakov Bar-Shalom, "The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, 1988, pp. 780–783.
2. Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, 1995.
3. Yaakov Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, New York, 1988.
4. B. Chen and J. K. Tugnait, "Tracking of Multiple Maneuvering Targets in Clutter using IMM/JPDA Filtering and Fixed-Lag Smoothing," *Automatica*, vol. 37, no. 2, 2001, pp. 239–249.
5. Peter Willett, Yanhua Ruan, and Roy L. Streit, "The PMHT for Maneuvering Targets," *Proceedings of SPIE*, July 1998.
6. L. E. Baum and T. Petrie, "Statistical Inference for Probabilistic Functions of Finite State Markov Chains," *Annals of Mathematical Statistics*, vol. 37, 1966, pp. 1554–1563.
7. L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *Annals of Mathematical Statistics*, vol. 41, no. 1, 1967, pp. 164–171.
8. Phillip L. Ainsleigh, "Theory of Continuous-State Hidden Markov Models and Hidden Gauss-Markov Models," NUWC-NPT Technical Report 11,274, Naval Undersea Warfare Center Division, Newport, RI, 20 March 2001.
9. Phillip L. Ainsleigh, N. Kehtarnavaz, and Roy L. Streit, "Hidden Gauss-Markov models for Signal Classification," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, 2002, pp. 1355–1367.
10. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society Series B*, vol. 39, 1977, pp. 1–38.
11. R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*, 2nd edition, John Wiley & Sons, Hoboken, NJ, 2002.
12. Geoffrey J. McLachlan and Thriyambakam Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, 1997.
13. Graham W. Pulford and Barbara F. La Scala, "MAP Estimation of Target Manoeuvre Sequence with the Expectation-Maximization Algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, April 2002, pp. 367–377.

14. Roy L. Streit and Tod E. Luginbuhl, "Maximum Likelihood Method for Probabilistic Multi-Hypothesis Tracking," *SPIE Conference on Signal and Data Processing of Small Targets*, Orlando, FL, 1994.
15. Roy L. Streit and Tod E. Luginbuhl, "Probabilistic Multi-Hypothesis Tracking," NUWC-NPT Technical Report 10,428, Naval Undersea Warfare Center Division, Newport, RI, 15 February 1995.
16. Andrew Logothetis, Vikram Krishnamurthy, and Jan Holst, "On Maneuvering Target Tracking via the PMHT," *Proceedings of the 36th Conference on Decision & Control*, San Diego, CA, December 1997, pp. 5024–5029.
17. Yanhua Ruan and Peter K. Willett, "Maneuvering PMHTs," *Proceedings of SPIE*, 2001, vol. 4473, pp. 186–197.
18. Tod E. Luginbuhl, Phillip Ainsleigh, Sunil Mathews, and Roy L. Streit, "Accurate Likelihood Evaluation for Multiple Model PMHT Algorithms," *IEEE Aerospace Conference*, 2008.
19. Robert A. Singer, "Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-6, no. 4, July 1970, pp. 473–483.
20. Steven Bordonaro, Peter Willett, and Yaakov Bar-Shalom, "Decorrelated Unbiased Converted Measurement Kalman Filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, April 2014, pp. 1431–1444.
21. Y. Bar-Shalom and X.-R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, Boston, 1993.
22. C. F. J. Wu, "On the Convergence Properties of the EM Algorithm," *Annals of Statistics*, vol. 11, 1983, pp. 95–103.
23. Anthony O'Hagan, *Kendall's Advanced Theory of Statistics*, vol. 2B: Bayesian Inference, Arnold, 1994.
24. Philip J. Davis and Philip Rabinowitz, *Methods of Numerical Integration*, 2nd edition, Academic Press, 1984.
25. T. A. Louis, "Finding the Observed Information Matrix when Using the EM Algorithm," *Journal of the Royal Statistical Society B*, vol. 44, 1982, pp. 226–233.
26. Michael J. Walsh, "Computing the Observed Information Matrix for Dynamic Mixture Models," NUWC-NPT Technical Report 11,768, Naval Undersea Warfare Center Division, Newport, RI, 25 September 2006.
27. Milton Abramowitz and Irene A. Stegun, eds., *Handbook of Mathematical Functions*, Dover Publications, New York, 1972.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Defense Technical Information Center	1
Naval Undersea Warfare Center Division Newport (Corporate Research and Information Center)	1