Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Asymptotically Optimal Nonlinear Filtering: Theory and Examples with Application to Target State Estimation

Tayfun Çimen*. A. O. Merttopçuoğlu*

** ROKETSAN Missiles Industries Inc., Ankara 06780*
Turkey (Tel: +90538-316-1356; e-mail: tcimen@roketsan.com.tr).

Abstract: The *State-Dependent Riccati Equation* (*SDRE*) filter, which is derived by constructing the dual of the well-known SDRE nonlinear regulator control design technique, has been studied in various papers, with mainly practical investigations of the filter. Until recently, theoretical aspects of the filter had not been fully investigated, leaving many unanswered questions, such as stability and convergence of the filter. The authors conducted an investigation of the conditions under which the state estimate given by this algorithm converges asymptotically to the first order minimum variance estimate given by the extended Kalman filter (EKF). Conditions for determining a region of stability for the SDRE filter were also investigated. The analysis was based on stable manifold theory and Hamilton-Jacobi-Bellman (HJB) equations. In this paper, the motivation for introducing HJB equations is justified with mathematical rigor, which is given by reference to the maximum likelihood approach to deriving the EKF. The application of the SDRE filter is then demonstrated on challenging examples to illustrate the theoretical aspects and design flexibility (additional degrees of freedom) of the algorithm when loss of observability is encountered. In particular, a realistic and detailed evaluation of the filter is carried out for the problem of target state estimation in an advanced tactical missile guidance application for analysis in the optimal guidance problem for air-air engagements using only passive sensor (angle-only) information. Simulation results are presented which show dramatic tracking improvement using the SDRE target tracker.

1. INTRODUCTION

During the past decade, ***State-Dependent Riccati Equation*** (***SDRE***) feedback control for nonlinear regulator systems has become well-known within the control community (see Çimen, 2008 and the references therein). Following the duality between linear-quadratic optimal regulation and linear-quadratic Gaussian estimation, SDRE filters have naturally been suggested in the literature (see Mracek et al., 1996; Pappano and Friedland, 1997; Haessig and Friedland, 1997) for continuous-time nonlinear systems. The algorithm is relatively well-known and involves solving, at a given point in state space, an algebraic state-dependent Riccati equation, or SDRE. The coefficients of this equation vary with the given point in state space. The resulting SDRE filter has the same structure as the infamous continuous steady-state linear Kalman filter. In contrast to the linearized Kalman filter (LKF) (Kalman and Bucy, 1961) and the extended Kalman filter (EKF) (Sage and Melsa, 1971; Bryson and Ho, 1975), which are based on *linearization*, the SDRE filter fully captures the nonlinearities of the system using *parameterization*, and bringing the nonlinear system to a *nonunique* linear structure having *state-dependent coefficients* (*SDCs*). The nonuniqueness of the parameterization creates extra degrees of freedom, which are not available in traditional filtering methods. These additional degrees of freedom can be used to enhance filter performance, avoid singularities, and avoid loss of observability.

The SDRE filter has been studied in various papers, with mainly practical investigations of the filter (Pappano, and Friedland, 1997; Harman and Bar-Itzhack, 1999). Theoretical aspects of the filter have recently been investigated in Çimen, McCaffrey, Harrison and Banks (2007), providing conditions under which the state estimate given by this algorithm converges asymptotically to the first-order minimum variance estimate given by the EKF. Behavioral differences and similarities between the SDRE filter, the LKF and the EKF were then discussed using a simple two-dimensional pendulum problem. Conditions for determining a region of stability for the SDRE filter were also investigated in that paper. The analysis was based on stable manifold theory and Hamilton-Jacobi-Bellman (HJB) equations. The motivation for introducing HJB equations was given by reference to the maximum likelihood approach to deriving the EKF, which is now rigorously justified. The application of the SDRE filter is also demonstrated on challenging examples throughout the paper to illustrate the theoretical aspects and design flexibility (additional degrees of freedom) of the algorithm.

The paper is thus organized as follows. First, the SDRE filter is reviewed in Sections 2. In Section 3, the design flexibility, that is, the additional degrees of freedom, provided by the nonuniqueness of the SDC parameterization is discussed. Asymptotic minimum variance filter has been studied in Çimen *et al.* (2007); the maximum likelihood approach to deriving the EKF is now studied in Section 4. In Section 5, the application of the SDRE filter is illustrated numerically on a challenging problem involving loss of algorithmic observability. Here the properties of the SDRE algorithm are illustrated by outlining the advantages, design flexibility, and implementation aspects. In Section 6, a realistic evaluation of the filter is carried out on a passive tracking application for the problem of target state estimation for advanced missile guidance laws. Simulation results are reported using (only) passive sensor information, where low observability is imminent. Concluding remarks are given in Section 7.

2. THE SDRE FILTER

The filtering counterpart of the SDRE Nonlinear Regulator (see Çimen, 2008) can be obtained by taking the dual of the steady-state linear-regulator and then allowing the coefficient matrices of the dual to be state-dependent. The dual of the steady-state linear regulator is the steady-state continuous Kalman observer, which in the absence of control reduces to the steady-state continuous Kalman filter. This leads to the following formulation. Consider the white-noise driven stochastic, nonlinear, autonomous system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{w}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{1}$$

together with white-noise-corrupted observations

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}) + \mathbf{v}(t) \tag{2}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{z}: \mathbb{R}^n \to \mathbb{R}^l$, $\mathbf{w}(t)$ and $\mathbf{v}(t)$ are zero-mean, white Gaussian noise vectors, uncorrelated with themselves or with $\mathbf{x}(t_0)$, such that for $t > t_0$,

$$\operatorname{cov}\{\mathbf{w}(t), \mathbf{w}(\tau)\} = \mathbf{W}\delta(t-\tau)$$
$$\operatorname{cov}\{\mathbf{v}(t), \mathbf{v}(\tau)\} = \mathbf{V}\delta(t-\tau)$$

where $\mathbf{W} \geq \mathbf{0}$ and $\mathbf{V} > \mathbf{0}$. Let us define the processes $d\boldsymbol{\omega}(t) \triangleq \mathbf{w}(t)dt$, $d\boldsymbol{\upsilon}(t) \triangleq \mathbf{v}(t)dt$ and $d\mathbf{y}(t) \triangleq \mathbf{z}(t)dt$. Then the system of equations (1) and (2) can be written more properly as the following Itô-sense stochastic differential equations

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x})dt + \mathbf{G}(\mathbf{x})d\boldsymbol{\omega}(t) \tag{3}$$

$$d\mathbf{y}(t) = \mathbf{h}(\mathbf{x})dt + d\boldsymbol{\upsilon}(t) \tag{4}$$

where $\boldsymbol{\omega}(t)$ and $\boldsymbol{\upsilon}(t)$ are independent Brownian motions uncorrelated with $\mathbf{x}(t_0)$, such that

$$\operatorname{cov}\{\boldsymbol{\omega}(t), \boldsymbol{\omega}(\tau)\} = \mathbf{W}\min(t, \tau)$$
$$\operatorname{cov}\{\boldsymbol{\upsilon}(t), \boldsymbol{\upsilon}(\tau)\} = \mathbf{V}\min(t, \tau).$$

Let us bring (3) and (4) to the SDC form

$$d\mathbf{x}(t) = \mathbf{F}(\mathbf{x})\mathbf{x}dt + \mathbf{G}(\mathbf{x})d\boldsymbol{\omega}(t) \tag{5}$$

$$d\mathbf{y}(t) = \mathbf{H}(\mathbf{x})\mathbf{x}dt + d\boldsymbol{\upsilon}(t)\,. \tag{6}$$

Let $\mathbf{Y}(t) \triangleq \{\mathbf{y}(\tau) : t_0 \leq \tau \leq t\}$ denote the observations up to time t. Let $\hat{\mathbf{x}}(t) \triangleq E\{\mathbf{x}(t) \big| \mathbf{Y}(t)\}$ denote the *conditional mean*, that is, the *minimum variance optimal estimate*, and $\mathbf{P}(t) \triangleq \operatorname{var}\{\mathbf{x}(t) - \hat{\mathbf{x}}(t) \big| \mathbf{Y}(t)\}$ denote the *conditional error variance*. Then, the SDRE filter for estimating the state $\mathbf{x}$ is

$$d\hat{\mathbf{x}}(t) = \mathbf{f}(\hat{\mathbf{x}})dt + \mathbf{K}(\hat{\mathbf{x}})[d\mathbf{y}(t) - \mathbf{h}(\hat{\mathbf{x}})dt] \tag{7}$$

with initial condition $\hat{\mathbf{x}}(t_0) \triangleq E\{\mathbf{x}(t_0)\}$, where $\mathbf{f}(\hat{\mathbf{x}}) = \mathbf{F}(\hat{\mathbf{x}})\hat{\mathbf{x}}$, $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{H}(\hat{\mathbf{x}})\hat{\mathbf{x}}$, and

$$\mathbf{K}(\hat{\mathbf{x}}) = \mathbf{P}(\hat{\mathbf{x}})\mathbf{H}^T(\hat{\mathbf{x}})\mathbf{V}^{-1} \tag{8}$$

is the filter gain, such that $\mathbf{P}(\hat{\mathbf{x}})$ satisfies at $\hat{\mathbf{x}}$ the positive-definite solution of the SDRE

$$\begin{aligned}\mathbf{G}(\hat{\mathbf{x}})\mathbf{W}\mathbf{G}^T(\hat{\mathbf{x}}) &+ \mathbf{F}(\hat{\mathbf{x}})\mathbf{P}(\hat{\mathbf{x}}) + \mathbf{P}(\hat{\mathbf{x}})\mathbf{F}^T(\hat{\mathbf{x}}) \\ &- \mathbf{P}(\hat{\mathbf{x}})\mathbf{H}^T(\hat{\mathbf{x}})\mathbf{V}^{-1}\mathbf{H}(\hat{\mathbf{x}})\mathbf{P}(\hat{\mathbf{x}}) = \mathbf{0}.\end{aligned} \tag{9}$$

3. DESIGN FLEXIBILITY

In a deterministic setting, before stochastic uncertainties are introduced, the SDC parameterization fully captures the nonlinearities of the system. In the multivariable case, the SDC parameterization is not unique and that the parameterization itself can be parameterized. This latter parameterization creates extra degrees of freedom that are not available in traditional filtering methods. These additional degrees of freedom provided by the nonuniqueness of the SDC parameterization can be used to enhance filter performance, avoid singularities, and avoid loss of observability, thus offering a more flexible nonlinear filter policy.

Let us now review the additional degrees of freedom provided by the nonuniqueness of the SDC parameterization. If $\mathbf{A}_1(\mathbf{x})$ and $\mathbf{A}_2(\mathbf{x})$ are two distinct SDC parameterizations, then

$$\mathbf{A}(\mathbf{x}, \alpha) = \alpha\mathbf{A}_1(\mathbf{x}) + (1-\alpha)\mathbf{A}_2(\mathbf{x})$$

is also an SDC parameterization for any α. Note that $\mathbf{A}(\mathbf{x}, \alpha)$ represents an infinite family of SDC parameterizations contained in a line. In general, an SDC parameterization $\mathbf{A}(\mathbf{x}, \alpha)$ can be constructed which is the parametric representation of a hyperplane containing k distinct parameterizations (if they exist), where α is a vector of dimension $k-1$. The introduction of α creates extra degrees of freedom that can be used not only to enhance the performance of the SDRE controller, but also its dual, the SDRE filter.

4. MAXIMUM LIKELIHOOD APPROACH

In the analysis of the minimum variance filter in Çimen *et al.* (2007), the central equations have been those for the EKF. These are derived from a first-order approximate solution to the modified *Fokker-Plank* equation, which describes the evolution of the conditional probability density of $\mathbf{x}(t)$ (see Chapter 9 of Sage & Melsa, 1971). However, the stability proof in Çimen *et al.* (2007) was based on Hamilton-Jacobi equations, which have been introduced in a formal way, and were not directly related to, or required in, the derivation of the equations for the filter dynamics. In this paper, therefore, the maximum likelihood approach will be outlined to estimating $\mathbf{x}(t)$. This approach does lead directly to the Hamilton-Jacobi equations which have been used to prove stability in Çimen *et al.* (2007). Let us start by formulating a least squares version of the problem of estimating $\mathbf{x}(t)$ in (1) given by the observations $\mathbf{z}(t)$ from (2). Minimizing an appropriate error function subject to the dynamical constraint of (1) can be shown to be equivalent to maximizing the conditional probability density function of $\mathbf{x}(t)$. The corresponding estimate $\hat{\mathbf{x}}(t)$ so obtained is the peak or mode of the conditional probability density function and constitutes the maximum likelihood (Bayesian) estimate of $\mathbf{x}(t)$. A Hamilton-Jacobi equation is obtained by using dynamical programming to solve the least squares problem.

Let $\bar{\mathbf{x}}(t_0) = E\{\mathbf{x}(t_0)\}$ and $\bar{\mathbf{P}}(t_0) = \mathrm{var}\{\mathbf{x}(t_0)\}$, and consider first the LQ estimation problem (around the equilibrium $\mathbf{x} = \mathbf{0}$) of minimizing

$$J = \tfrac{1}{2}\big(\mathbf{x}(t_0)-\bar{\mathbf{x}}(t_0)\big)^T \bar{\mathbf{P}}^{-1}(t_0)\big(\mathbf{x}(t_0)-\bar{\mathbf{x}}(t_0)\big) + \tfrac{1}{2}\int_{t_0}^{t}\Big\{\big(\mathbf{z}(\tau)-\mathbf{H}(\mathbf{0})\mathbf{x}(\tau)\big)^T \mathbf{V}^{-1}\big(\mathbf{z}(\tau)-\mathbf{H}(\mathbf{0})\mathbf{x}(\tau)\big)+\mathbf{w}^T(\tau)\mathbf{W}^{-1}\mathbf{w}(\tau)\Big\}\, d\tau$$

with respect to $\mathbf{x}(\tau)$ and $\mathbf{w}(\tau)$ subject to the constraint

$$\dot{\mathbf{x}}(t) = \mathbf{F}(\mathbf{0})\mathbf{x}(t) + \mathbf{G}(\mathbf{0})\mathbf{w}(t).$$

This can be thought of as attempting to determine (estimate) $\mathbf{x}(\tau)$ for $t_0 \le \tau \le t$ so that, simultaneously, the errors in the dynamics and in the observations are small. In view of the constraint, it is enough to minimize J with respect to $\mathbf{x}(t_0)$ and $\mathbf{w}(\tau)$ since $\mathbf{x}(\tau)$ is then determined for $t_0 \le \tau \le t$. Let

$$S(\mathbf{x}, t) = \min_{\mathbf{w}(\tau)} J.$$

Then, the dynamic programming equation is given by

$$-\tfrac{\partial S}{\partial t} = \min_{\mathbf{w}(\tau)}\Big\{\tfrac{\partial S^T}{\partial \mathbf{x}}\big[\mathbf{F}(\mathbf{0})\mathbf{x} + \mathbf{G}(\mathbf{0})\mathbf{w}\big] - \tfrac{1}{2}[(\mathbf{z}-\mathbf{H}(\mathbf{0})\mathbf{x})^T\mathbf{V}^{-1}(\mathbf{z}-\mathbf{H}(\mathbf{0})\mathbf{x}) + \mathbf{w}^T\mathbf{W}^{-1}\mathbf{w}]\Big\}.$$

This is minimized by $\mathbf{w} = \mathbf{W}\mathbf{G}^T(\mathbf{0})\tfrac{\partial S}{\partial \mathbf{x}}$, giving

$$\begin{aligned}-\tfrac{\partial S}{\partial t} &= \tfrac{\partial S^T}{\partial \mathbf{x}}\mathbf{F}(\mathbf{0})\mathbf{x} + \tfrac{1}{2}\tfrac{\partial S^T}{\partial \mathbf{x}}\mathbf{G}(\mathbf{0})\mathbf{W}\mathbf{G}^T(\mathbf{0})\tfrac{\partial S}{\partial \mathbf{x}} \\ &\quad -\tfrac{1}{2}(\mathbf{z}-\mathbf{H}(\mathbf{0})\mathbf{x})^T\mathbf{V}^{-1}(\mathbf{z}-\mathbf{H}(\mathbf{0})\mathbf{x}).\end{aligned} \tag{10}$$

The linear filter is obtained by supposing there is a solution of the form

$$S(\mathbf{x}, t) = \tfrac{1}{2}\big(\mathbf{x}-\hat{\mathbf{x}}(t)\big)^T \mathbf{P}^{-1}(t)\big(\mathbf{x}-\hat{\mathbf{x}}(t)\big) + a(t).$$

It can be shown that $-S(\mathbf{x}, t)$ is the exponent of the conditional probability density function of $\mathbf{x}(t)$. Thus $-S(\mathbf{x}, t)$ can be interpreted as the likelihood of the state trajectory passing through $\mathbf{x}$ at time t, given the observations $\mathbf{z}$ made up to time t. This is clearly maximized by $\mathbf{x} = \hat{\mathbf{x}}(t)$, which is therefore the maximum likelihood filtering solution. The equations for $\hat{\mathbf{x}}$ and $\mathbf{P}$ are obtained by calculating $\partial S/\partial t$ and $\partial S/\partial \mathbf{x}$ and substituting in (10). Since the dynamics are linear, $\mathbf{x}(t)$ is normal and so the conditional mean coincides with the conditional mode. In other words, in the linear case, the minimum variance and maximum likelihood solutions coincide and, indeed, it turns out that the equations obtained from (10) are same as those for the LKF:

$$\tfrac{d\hat{\mathbf{x}}(t)}{dt} = \mathbf{F}(\mathbf{0})\hat{\mathbf{x}}(t) + \mathbf{P}(t)\mathbf{H}^T(\mathbf{0})\mathbf{V}^{-1}\{\mathbf{z}(t) - \mathbf{H}(\mathbf{0})\hat{\mathbf{x}}(t)\} \tag{11}$$

$$\begin{aligned}\tfrac{d\mathbf{P}(t)}{dt} &= \mathbf{F}(\mathbf{0})\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(\mathbf{0}) + \mathbf{G}(\mathbf{0})\mathbf{W}\mathbf{G}^T(\mathbf{0}) \\ &\quad -\mathbf{P}(t)\mathbf{H}^T(\mathbf{0})\mathbf{V}^{-1}\mathbf{H}(\mathbf{0})\mathbf{P}(t).\end{aligned} \tag{12}$$

Details of the above are contained in Jazwinski (1970), Section 5.3 and Exs. 7.11-7.12, and will essentially be given in the derivation of the first-order nonlinear solution below. Consider now the nonlinear estimation problem away from $\mathbf{x} = \mathbf{0}$. Thus, consider minimizing

$$\bar{J} = \tfrac{1}{2}\big(\mathbf{x}(t_0)-\bar{\mathbf{x}}(t_0)\big)^T \bar{\mathbf{P}}^{-1}(t_0)\big(\mathbf{x}(t_0)-\bar{\mathbf{x}}(t_0)\big) + \tfrac{1}{2}\int_{t_0}^{t}\Big\{\big(\mathbf{z}(\tau)-\mathbf{h}(\mathbf{x}(\tau))\big)^T \mathbf{V}^{-1}\big(\mathbf{z}(\tau)-\mathbf{h}(\mathbf{x}(\tau))\big)+\mathbf{w}^T(\tau)\mathbf{W}^{-1}\mathbf{w}(\tau)\Big\}\, d\tau$$

subject to

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{G}(\mathbf{x}(t))\mathbf{w}(t).$$

Letting

$$\bar{S}(\mathbf{x}, t) = \min_{\mathbf{w}(\tau)} \bar{J},$$

and repeating the above analysis leads to the following Hamilton-Jacobi equation

$$\begin{aligned}-\tfrac{\partial \bar{S}}{\partial t} &= \tfrac{\partial \bar{S}^T}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \tfrac{1}{2}\tfrac{\partial \bar{S}^T}{\partial \mathbf{x}}\mathbf{G}(\mathbf{x})\mathbf{W}\mathbf{G}^T(\mathbf{x})\tfrac{\partial \bar{S}}{\partial \mathbf{x}} \\ &\quad -\tfrac{1}{2}(\mathbf{z}-\mathbf{h}(\mathbf{x}))^T\mathbf{V}^{-1}(\mathbf{z}-\mathbf{h}(\mathbf{x})).\end{aligned} \tag{13}$$

The value of $\mathbf{x}$ which maximizes $-\bar{S}$ is the maximum likelihood estimate, and is denoted $\hat{\mathbf{x}}(t)$. However, it is no longer true that this coincides with the minimum variance estimate. Also, it is not possible to derive an exact equation for $\hat{\mathbf{x}}(t)$ from (13). A first-order approximate equation for $\hat{\mathbf{x}}(t)$ can, however, be derived by expanding the various terms in (13) in Taylor series around $\hat{\mathbf{x}}(t)$ in terms up to order 1. For the sake of brevity, the dependence of $\hat{\mathbf{x}}$ on t will be omitted in the following where $\hat{\mathbf{x}}$ appears inside another function. Since $\hat{\mathbf{x}}(t)$ minimizes $\bar{S}(\mathbf{x}, t)$, close to $\hat{\mathbf{x}}$,

$$\tfrac{\partial \bar{S}(\mathbf{x}, t)}{\partial \mathbf{x}} = \tfrac{\partial^2 \bar{S}(\hat{\mathbf{x}}, t)}{\partial \mathbf{x}^2}(\mathbf{x}-\hat{\mathbf{x}}(t)) = \mathbf{P}^{-1}(t)(\mathbf{x}-\hat{\mathbf{x}}(t)), \tag{14}$$

where $\mathbf{P}^{-1}(t) = \tfrac{\partial^2 \bar{S}(\hat{\mathbf{x}}(t), t)}{\partial \mathbf{x}^2}$. Then, to a first-order approximation,

$$\bar{S}(\mathbf{x}, t) = \tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)(\mathbf{x}-\hat{\mathbf{x}}(t)) + a(t) \tag{15}$$

Substituting (14) into (13) gives

$$\begin{aligned}-\tfrac{\partial \bar{S}}{\partial t} &= (\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\mathbf{f}(\mathbf{x}) - \tfrac{1}{2}(\mathbf{z}-\mathbf{h}(\mathbf{x}))^T\mathbf{V}^{-1}(\mathbf{z}-\mathbf{h}(\mathbf{x})) \\ &\quad +\tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\mathbf{G}(\mathbf{x})\mathbf{W}\mathbf{G}^T(\mathbf{x})\mathbf{P}^{-1}(t)(\mathbf{x}-\hat{\mathbf{x}}(t)).\end{aligned} \tag{16}$$

From (15), however,

$$\begin{aligned}\tfrac{\partial \bar{S}(\mathbf{x}, t)}{\partial t} &= \dot{a}(t) - (\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\tfrac{d\hat{\mathbf{x}}(t)}{dt} \\ &\quad +\tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\tfrac{d\mathbf{P}^{-1}(t)}{dt}(\mathbf{x}-\hat{\mathbf{x}}(t)).\end{aligned} \tag{17}$$

For the other terms in (16), the following are obtained:

$$\left.\begin{aligned}\mathbf{f}(\mathbf{x}) &= \mathbf{f}(\hat{\mathbf{x}}) + \tfrac{\partial \mathbf{f}(\hat{\mathbf{x}})}{\partial \mathbf{x}}\big(\mathbf{x}-\hat{\mathbf{x}}(t)\big) \\ \mathbf{G}(\mathbf{x}) &= \mathbf{G}(\hat{\mathbf{x}}) + \tfrac{\partial \mathbf{G}(\hat{\mathbf{x}})}{\partial \mathbf{x}}\big(\mathbf{x}-\hat{\mathbf{x}}(t)\big) \\ \mathbf{h}(\mathbf{x}) &= \mathbf{h}(\hat{\mathbf{x}}) + \tfrac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \mathbf{x}}\big(\mathbf{x}-\hat{\mathbf{x}}(t)\big).\end{aligned}\right\} \tag{18}$$

Substituting (17) and (18) into (16), and ignoring terms of order higher than two in $\mathbf{x}$, gives

$$\begin{aligned}&-\dot{a}(t) + (\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\tfrac{d\hat{\mathbf{x}}(t)}{dt} - \tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\tfrac{d\mathbf{P}^{-1}(t)}{dt}(\mathbf{x}-\hat{\mathbf{x}}(t)) \\ &= (\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\mathbf{f}(\hat{\mathbf{x}}) + (\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\tfrac{\partial \mathbf{f}(\hat{\mathbf{x}})}{\partial \mathbf{x}}(\mathbf{x}-\hat{\mathbf{x}}(t)) \\ &\quad +\tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\mathbf{P}^{-1}(t)\mathbf{G}(\hat{\mathbf{x}})\mathbf{W}\mathbf{G}^T(\hat{\mathbf{x}})\mathbf{P}^{-1}(t)(\mathbf{x}-\hat{\mathbf{x}}(t)) \\ &\quad -\tfrac{1}{2}\mathbf{z}^T\mathbf{V}^{-1}\mathbf{z} + \mathbf{h}^T(\hat{\mathbf{x}})\mathbf{V}^{-1}(\mathbf{z}-\tfrac{1}{2}\mathbf{h}(\hat{\mathbf{x}})) \\ &\quad +(\mathbf{x}-\hat{\mathbf{x}}(t))^T\tfrac{\partial \mathbf{h}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{V}^{-1}(\mathbf{z}-\mathbf{h}(\hat{\mathbf{x}})) \\ &\quad -\tfrac{1}{2}(\mathbf{x}-\hat{\mathbf{x}}(t))^T\tfrac{\partial \mathbf{h}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{V}^{-1}\tfrac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \mathbf{x}}(\mathbf{x}-\hat{\mathbf{x}}(t)).\end{aligned}$$

Now, equating terms of order 1 and 2 in $(\mathbf{x}-\hat{\mathbf{x}}(t))$ gives the first-order approximate equations for $\hat{\mathbf{x}}$ and $\mathbf{P}^{-1}$ as follows:

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = \mathbf{f}(\hat{\mathbf{x}}) + \mathbf{P}(t)\frac{\partial \mathbf{h}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{V}^{-1}(\mathbf{z}(t) - \mathbf{h}(\hat{\mathbf{x}})) \tag{19}$$

$$\begin{aligned}-\frac{d\mathbf{P}^{-1}(t)}{dt} &= \mathbf{P}^{-1}(t)\frac{\partial \mathbf{f}(\hat{\mathbf{x}})}{\partial \mathbf{x}} + \frac{\partial \mathbf{f}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{P}^{-1}(t) \\ &+ \mathbf{P}^{-1}(t)\mathbf{G}(\hat{\mathbf{x}})\mathbf{W}\mathbf{G}^T(\hat{\mathbf{x}})\mathbf{P}^{-1}(t) - \frac{\partial \mathbf{h}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{V}^{-1}\frac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \mathbf{x}}.\end{aligned} \tag{20}$$

The initial conditions are obtained from the boundary condition for (13), that is,

$$\bar{S}(\mathbf{x}(t_0), t_0) = \tfrac{1}{2}(\mathbf{x}(t_0) - \bar{\mathbf{x}}(t_0))^T \bar{\mathbf{P}}^{-1}(t_0)(\mathbf{x}(t_0) - \bar{\mathbf{x}}(t_0)).$$

For $\bar{S}$ of the form (15), this is satisfied by $\hat{\mathbf{x}}(t_0) = \bar{\mathbf{x}}(t_0)$ and $\mathbf{P}^{-1}(t_0) = \bar{\mathbf{P}}^{-1}(t_0)$. Implementing the algorithm (19) and (20) involves inverting $\mathbf{P}^{-1}(t)$ at each step. An explicit equation for $\mathbf{P}(t)$ is obtained from (20) by noting that $\mathbf{P}\mathbf{P}^{-1} = \mathbf{I}$, which implies $\frac{d\mathbf{P}}{dt} = -\mathbf{P}\frac{d\mathbf{P}^{-1}}{dt}\mathbf{P}$, and so

$$\begin{aligned}\frac{d\mathbf{P}(t)}{dt} &= \frac{\partial \mathbf{f}(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{P}(t) + \mathbf{P}(t)\frac{\partial \mathbf{f}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}} \\ &+ \mathbf{G}(\hat{\mathbf{x}})\mathbf{W}\mathbf{G}^T(\hat{\mathbf{x}}) - \mathbf{P}(t)\frac{\partial \mathbf{h}^T(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{V}^{-1}\frac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \mathbf{x}}\mathbf{P}(t).\end{aligned} \tag{21}$$

Eqs. (19) and (21) thus constitute a first-order approximation to the maximum likelihood solution to the filtering problem (1) and (2). Note that these equations are the same as those for the EKF which gives the first-order minimum variance solution. As noted above, these solutions do *not* generally coincide and so one can expect the higher order approximations to diverge from one another. Also, note that the SDRE algorithm in (7)-(9) can be thought of as an approximation to the steady-state form of (19) and (21), should this exist. It is pointed out in Jazwinski (1970), Section 5.3, that maximum likelihood estimation is of questionable value unless it is known in advance that the conditional probability density function of $\mathbf{x}(t)$ is unimodal and concentrated near the mode. The point of this section, however, was to indicate where the Hamilton-Jacobi equations that were used to prove stability in Çimen *et al.* (2007) came from and how they are related to the derivation of the equations for the EKF.

In order to analyze the performance of the SDRE filter and illustrate the theory developed, a numerical example from Mracek et al. (1996) has been reconsidered in Çimen *et al.* (2007), using a simple pendulum operating in the nonlinear regime. A comparative study by Monte Carlo simulations has been carried out between the SDRE filter, the LKF, and the EKF on this simple problem. The asymptotic behavior of the SDRE filter was shown to yield much improved performance and convergence properties compared with these local approximations, coping with highly misleading initial states, and converging rapidly to the true states. In the sequel, a more challenging problem with loss of algorithmic observability is considered.

5. LOSS OF OBSERVABILITY PROBLEM

Consider the following example with the system of equations (Mracek, Cloutier and D'Souza, 1996; Ewing, 2000)

$$\dot{x}_1(t) = w_1(t), \quad \dot{x}_2(t) = w_2(t)$$

and measurement equations

$$y_1(t) = x_1(t) + x_2(t) + v_1(t)$$
$$y_2(t) = x_1(t)x_2(t) + v_2(t),$$

where $w_1(t)$, $w_2(t)$, $v_1(t)$ and $v_2(t)$ are zero-mean Gaussian-distributed random variables. Using the complementary estimation terms, this system can be shown to be weakly observable everywhere except at the origin. Now, the continuous EKF uses a Taylor series expansion to linearize the measurement about the current state estimate as

$$\frac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}} = \begin{bmatrix} 1 & 1 \\ \hat{x}_2 & \hat{x}_1 \end{bmatrix}.$$

Thus, $\frac{\partial \mathbf{h}(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}}$ loses rank whenever $\hat{x}_1 = \hat{x}_2$, at which point the system becomes unobservable to the EKF algorithm. Therefore, for the LKF filter, where the coefficient matrix is given by $\frac{\partial \mathbf{h}(\mathbf{0})}{\partial \hat{\mathbf{x}}}$, loss of observability is encountered for all $\hat{\mathbf{x}}$. There is no solution to the problem with loss of observability using the LKF or the EKF. Using the SDRE filter, however, there are two distinct parameterizations

$$\mathbf{H}_1(\mathbf{x}) = \begin{bmatrix} 1 & 1 \\ x_2 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{H}_2(\mathbf{x}) = \begin{bmatrix} 1 & 1 \\ 0 & x_1 \end{bmatrix},$$

which can be combined to form the parameterized SDC measurement matrix as

$$\mathbf{H}(\mathbf{x}, \alpha) = \begin{bmatrix} 1 & 1 \\ \alpha x_2 & (1-\alpha)x_1 \end{bmatrix}.$$

Using this SDC parameterization, α can be chosen such that loss of rank is avoided.

The EKF and the SDRE filter were coded and evaluated numerically in MATLAB®, simulating the problem for 10 seconds at an update rate of 100Hz. The *dispersion of the estimation errors* over 50 runs is illustrated in Fig. 1 for initial positions $x_1(0) = 0.1$ and $x_2(0) = 0.3$, with initial variance set to $\mathbf{P}_0 = \mathbf{I}_{2\times2}$. In the simulations, the value of α has been taken as 0.8.

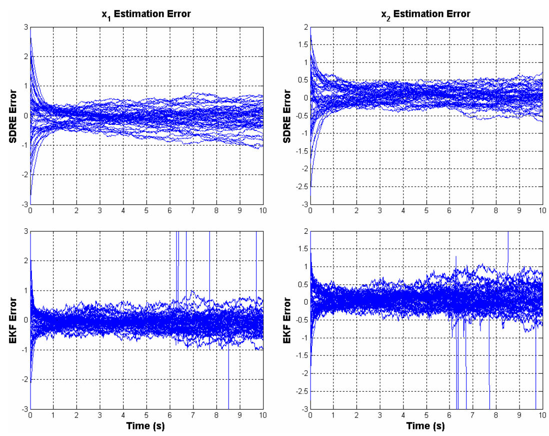

Fig. 1. Error distribution between true and estimated states over 50 Monte Carlo runs when $x_1(0) = 0.1$, $x_2(0) = 0.3$, $\mathbf{P}_0 = \mathbf{I}_{2\times2}$, $\mathbf{W} = \mathbf{I}_{2\times2}$, and $\mathbf{V} = 0.1\mathbf{I}_{2\times2}$

When the state estimates cross the algorithmically unobservable condition of $x_1 = x_2$, numerical instability is encountered in the EKF estimates, which is clearly seen from the corresponding error plots. Note, however, how the SDRE filter still provides accurate nonlinear estimates of the true states, even in cases of algorithmic loss of observability. This simple academic problem illustrates the advantage of the SDRE filter when loss of algorithmic observability is an issue, which cannot be prevented using the LKF or the EKF. Using the SDRE filter formulation, however, the SDC parameterizations can be adjusted to avoid these conditions, so that when the state estimates cross the algorithmically unobservable condition, filter convergence can still be accomplished. Note also that, when the initial variance is increased, such that the filter initial conditions highly differ from the actual ones, the EKF estimates in each run end up diverging from the actual states. The SDRE filter estimates, on the other hand, still continue to converge to the true states.

6. TARGET STATE ESTIMATION

Estimation of the state trajectories of a moving target, also known as *target tracking*, has been a problem of active concern to practitioners in both military and civilian applications since the 1960's. There are several nonlinear techniques available for maneuvering target tracking. From a computational standpoint, however, minimum-variance estimation is the most common choice of implementation because of its simplicity, compared to other techniques such as statistical linearization and batch least squares. The simplest implementation of minimum-variance estimation is the EKF, which has been applied extensively to interception problems. Specifically, this technology has been used to develop tracking algorithms that extract the maximum amount of information about a target trajectory from homing sensor data and to derive advanced guidance laws that optimize the use of this information in directing the missile towards the selected target. There are, however, severe problems associated with the synthesis of target trackers in intercept applications such as air-to-air missiles (Hepner and Geering, 1991). These problems include filter divergence due to *lack of complete observability*, *modeling errors*, and the restriction of the *computation time* due to high sampling rates. It is well-known that low observability always occurs when bearing angles or bearing rates are the only measurements available about the missile-target relative motion. These measurements are common in short-range missiles. In this section, the SDRE filter is used for addressing the critical problem of estimating the missile-to-target position and velocity, and target acceleration (required by optimal guidance laws) when only *passive seeker* (*angle only*) information is available onboard a missile. The estimation effectiveness of the SDRE filter is evaluated to determine its influence on tactical missile guidance when coupled with an *optimal linear-quadratic* (LQ) guidance law. The tracking performance of this concept, implemented for a passive tracking system, is evaluated by simulations.

6.1 Filter Dynamics

In target tracking, the state vector $\mathbf{x}(t)$ of the target dynamics usually contains position, velocity, and acceleration as state variables. Target acceleration modeling is a critical design factor in the filter. In practical applications, it is impossible to model the target acceleration accurately because target motion cannot be exactly known, particularly a flying target. Due to a lack of knowledge of its dynamics, this unknown input is often modeled as a *random process*. One of the simplest models used in maneuvering target tracking is the *Wiener-process acceleration model*. Thus, when a target is treated as a point object, this maneuvering motion is described in continuous time by the vector-valued equation $\dot{\mathbf{a}}_T(t) = \mathbf{w}_T(t)$, where the *acceleration derivative* $\dot{\mathbf{a}}_T(t)$ (also known as "*jerk*") is an independent process (white noise) $\mathbf{w}_T(t)$ with power spectral density $\mathbf{W}_T$ that accounts for unpredictable modeling errors in $\mathbf{a}_T(t)$. In the literature, this model is also referred to simply as the continuous-time "*nearly-constant-acceleration model*".

The nonlinear filtering problem is now set up as a *nine-state* filter with the system model in Cartesian coordinates. The nine filter states are the three components of the *relative position* vector ($\mathbf{r}$), *relative velocity* vector ($\dot{\mathbf{v}}$), and *target acceleration* vector, all with respect to inertial coordinates. Therefore, using a nearly-constant-acceleration model to improve missile performance, a minimal representation for the state vector includes $\mathbf{x} = [\mathbf{r}^T \quad \mathbf{v}^T \quad \mathbf{a}_T{}^T]^T$ as the nine variables, which represent coordinates of a Cartesian coordinate system in three dimensions used to describe the missile-to-target relative motion and target acceleration. Based on equations of missile-target relative motion, the corresponding state-space representation can be derived as

$$\begin{bmatrix} \dot{\mathbf{r}}(t) \\ \dot{\mathbf{v}}(t) \\ \dot{\mathbf{a}}_T(t) \end{bmatrix}_{9\times1} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix}_{9\times9} \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{v}(t) \\ \mathbf{a}_T(t) \end{bmatrix}_{9\times1} + \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix}_{9\times9} \begin{bmatrix} \mathbf{0}_{3\times1} \\ -\mathbf{a}_M(t) \\ \mathbf{0}_{3\times1} \end{bmatrix}_{9\times1} + \begin{bmatrix} \mathbf{0}_{3\times1} \\ \mathbf{w}_M(t) \\ \mathbf{w}_T(t) \end{bmatrix}_{9\times1},$$

where the commanded acceleration $\mathbf{a}_M(t)$ is determined using the linear-quadratic (LQ) optimal guidance law

$$\mathbf{a}_M(t) = \tfrac{3}{\hat{t}_{go}{}^2}\left[\hat{\mathbf{r}}(t) + \hat{t}_{go}\hat{\mathbf{v}}(t) + \tfrac{1}{2}\hat{t}_{go}{}^2\hat{\mathbf{a}}_T(t)\right]_{3\times1},$$

with the estimate of *time-to-go* constructed from $\hat{r}(t) = \|\hat{\mathbf{r}}(t)\|$ and $\hat{\dot{r}}(t)$ using the approximate relation $\hat{t}_{go} \approx -\hat{r}/\hat{\dot{r}}$.

As mentioned before, estimating position, velocity and acceleration from angle-only (or rate-only) information is a very difficult problem because bearing angle-only or bearing rate-only measurements do not guarantee complete observability of the state $\mathbf{x}$. Since the axial acceleration component is unobservable, the implementation of guidance laws that use this quantity is useless. However, it is simple to see that lateral and normal components of target acceleration are sufficient information to keep the missile on the homing path, thus ensuring intercept. Therefore, only estimates of the observable quantities are required in the LQ guidance law. The continuous-time vector-valued white noise process

$\mathbf{w}(t) = [\mathbf{0}_{1\times3} \quad \mathbf{w}_M{}^T(t) \quad \mathbf{w}_T{}^T(t)]^T$ is used to characterize missile acceleration and target modeling errors. Its corresponding continuous-time process noise intensity matrix

$$\mathbf{W} = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{W}_M & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{W}_T \end{bmatrix}_{9\times9},$$

which is equivalent to the power spectral density matrix when time-invariant.

6.2 Measurement Model

Sensors used for target tracking provide measurements of a target in a natural sensor (3D spherical or 2D polar) coordinate system or frame. Therefore, while target motion models are best described by target state in *Cartesian coordinates*, measurements of the target state are directly available in the original *sensor* (non-Cartesian) *coordinates*. For *passive sensors* onboard the missile and *range-denial countermeasures*, the sensor only provides a measure of the bearing (or azimuth) angle θ and elevation angle ϕ. As such, in the sensor coordinates, measurements for angle-only passive tracking from a *strapdown* (body-fixed) seeker are generally modeled in body axes as the spherical angles of the LOS vector $\mathbf{r}$ in the following form of additive noise:

$$\phi_k = \tan^{-1}\left(\frac{z_k}{\sqrt{x_k{}^2 + y_k{}^2}}\right) + v_{\phi_k}, \quad \theta_k = \tan^{-1}\left(\frac{y_k}{x_k}\right) + v_{\theta_k},$$

where x, y, and z are the three components of relative position $\mathbf{r}$ in inertial coordinates, and $\mathbf{v}_k \sim N(\mathbf{0}, \mathbf{V}_k)$. Measurements are, therefore, nonlinear functions of the true state, corrupted by additive errors with Gaussian statistical properties. For tracking, however, these measurements in the sensor coordinates are converted to the Cartesian coordinates, using the spherical-to-Cartesian transformation. A unit vector in the line-of-sight direction is:

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_x & \lambda_y & \lambda_z \end{bmatrix}^T = \begin{bmatrix} \cos\theta\cos\phi & \sin\theta\cos\phi & \sin\phi \end{bmatrix}^T.$$

The LOS rates are determined by differentiating $\boldsymbol{\lambda}$. Tracking is then performed entirely in the Cartesian coordinates as this provides equations with an attractive "linear-like" structure, which is desirable for the SDRE filter. From kinematic relations, for bearing-only and bearing-rate-only measurements, the state-dependent measurement sensitivity matrix $\mathbf{H}(\mathbf{x})$ for SDRE filtering becomes

$$\mathbf{H}(\mathbf{x}_k) = \begin{bmatrix} \frac{1}{\|\mathbf{r}_k\|}\mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ -\frac{\alpha \mathbf{r}_k{}^T \mathbf{v}_k}{\|\mathbf{r}_k\|^3}\mathbf{I}_{3\times3} & \frac{\alpha}{\|\mathbf{r}_k\|}\mathbf{I}_{3\times3} + \frac{(1-\alpha)\left(\mathbf{I}_{3\times3} - \boldsymbol{\lambda}_k \boldsymbol{\lambda}_k{}^{\mathrm{T}}\right)}{\|\mathbf{r}_k\|} & \mathbf{0}_{3\times3} \end{bmatrix}_{6\times9},$$

where $\boldsymbol{\lambda} = \mathbf{r}/r$, and $0 \le \alpha \le 1$ becomes a design parameter.

6.3 Noise statistics

The process noise covariance $\mathbf{W}_T$ associated with the process noise $\mathbf{w}_T$ for target acceleration is selected so as to inform the filter of the estimated variance on the *lateral* and *normal* target acceleration components. For example, with an estimated variance of $7.8\times10^3\ \mathrm{m^2/s^4}$, the filter is informed that the target acceleration is within ±9g of its true value. Random noise effects are included in the process noise $\mathbf{w}_M$ for missile acceleration to account for accelerometer noise. With an RMS error of 0.4g for an accelerometer, the corresponding process noise covariance is $[0.4(9.81\ \mathrm{m/s^2})]^2$. The statistics of the measurement noise process $\mathbf{v}_k$ are selected to represent angle measurement errors. Measurement noise is assumed to be normally distributed with zero mean and with corresponding standard deviations of $0.01\mathrm{rad}$. Measurement update frequency is assumed to be 100Hz and the model propagation time step is set to $\Delta t = 0.01\mathrm{s}$. Because of the transformation of measurements from spherical to Cartesian coordinates, it must be emphasized that the components of the measurement noise covariance matrix in the Cartesian coordinates become correlated. Due to space limitations, however, this particular aspect of the tracking model is not addressed in the paper.

6.4 Simulation results

Now consider a short-range air-to-air missile equipped with a passive seeker, and a highly maneuvering target aircraft. In the sequel, it will be assumed that only bearing angle information is available about the missile-target relative motion. Bearing rates are obtained by differentiating the noisy angular measurements in the Cartesian coordinates. The main goal is the synthesis of a tracking filter that supplies estimates of range, relative velocity, and target acceleration to the LQ guidance law of the missile. The basic approach is the SDRE filter employing a nearly-constant-acceleration target model coupled with the LQ guidance law during a single flyout in the presence of process and measurement noise, where measurements are related to the target states in the Cartesian coordinates in a nonlinear way.

In order to investigate the quick response capability of the SDRE algorithm to lateral maneuver acceleration, a 2D simulation is carried out for a particular engagement scenario, where the motion takes place in a horizontal plane. The missile is launched from the origin of the Cartesian coordinate system with velocity $\mathbf{v}_M = [500 \quad 0]^T\ \mathrm{m/s}$, and with zero lateral acceleration. The target, on the other hand, flies at a constant speed of $270\,\mathrm{m/s}$ with velocity vector $\mathbf{v}_T = [250 \quad 100]^T\ \mathrm{m/s}$. With initial position of the target $\mathbf{r}_T = [8000 \quad 500]^T\ \mathrm{m}$, initial range is around $8\,\mathrm{km}$. During the engagement, the target starts with zero lateral acceleration. The first maneuver takes place in the middle of the engagement when $r = 6000\,\mathrm{m}$, and is sustained until $r = 2000\,\mathrm{m}$ with magnitude 5.5g's. The second maneuver occurs when $r = 2000\,\mathrm{m}$ and lasts until intercept with a stronger magnitude of 8.8g's in the opposite direction. Since range is unknown to the missile, the filter cannot be initialized at launch with the true relative position values. Furthermore, no knowledge about target velocity and acceleration is assumed to be provided at the beginning. As such, initial range estimate is overestimated as 10km, whereas initial target velocity and acceleration estimates are both set to zero. Consequently, initial value of relative velocity is initialized using missile velocity from pursuer INS. Simulation results are of the engagement are displayed in Fig. 2, showing the exact and estimated histories of range, relative velocity, and target acceleration, along with the corresponding missile-target trajectories using these estimated quantities. The simulation results shown here are typical of a large number of runs made in a number of scenarios that verified the stable and exceptionally qualified performance of the SDRE filter during rapid maneuvers. Note that only the estimates in the y-coordinate have been given, because the components in the x-coordinate are unobservable with bearing-angle only measurements, and the SDRE algorithm cannot give the correct estimates of these quantities, albeit there is no divergence.

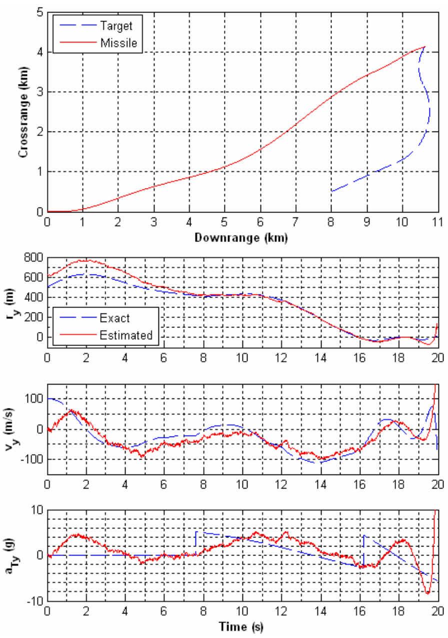

Fig. 2. Exact and estimated states with missile-target trajectories of the engagement scenario

The accuracy of the state estimates obtained to date using the EKF leaves much room for improvement. A method of selecting the initial values for the target variables in the tracking equation is extremely important for an EKF since the equations are linearized. If the selection of these values is in error by a large margin, the linearized elements of the tracking equation may produce a divergent track. This is in contrast to the SDRE filter, which does a good job of estimating the states with only passive seeker information in a highly dynamic environment. Large initial errors in range and velocity are quickly corrected soon after the engagement begins. When the target begins to maneuver, the SDRE filter is able to maintain track. With frequent updates but inaccurate measurements of LOS angle, excellent SDRE filter performance is demonstrated. The results ascertain the insensitivity of the SDRE filter to measurement noise, initialization error, and robustness in the presence of stressing target maneuvers.

7. CONCLUSIONS

As predicted from theoretical considerations, it is quite evident from simulation results that the SDRE filter becomes significant for increasing values of initial error and noise variances. The filter is capable of tracking the true values of the states, with little sensitivity to the selection of the statistics, or even to severe differences in the initial state estimates when compared with the EKF. Analysis of the SDRE filter also demonstrated an advantage over linearization techniques such as LKF and EKF when confronted with the problem of loss of algorithmic observability. The SDRE filter offers the advantage of multiple parameterizations in hopes of improving performance. This offers a distinct advantage over the EKF in successfully working the observability problem. These results also proved valuable in target state estimation for advanced missile guidance problems. The SDRE filter was compared to, and shown to outperform, both the LKF and the EKF throughout the authors' studies, and is proved to be a viable candidate for nonlinear estimation.

REFERENCES

Çimen, T., D. McCaffrey, R.F. Harrison and S.P. Banks (2007). Asymptotically optimal nonlinear filtering. In: *Proc. of the 17th IFAC Symposium on Automatic Control in Aerospace*, Toulouse, France.

Çimen, T. (2008). State-dependent Riccati equation (SDRE) control: A survey. In: *Proc. of the 17th IFAC World Congress*, Seoul, South Korea.

Ewing, C.M. (2000). An analysis of the state dependent Riccati equation method nonlinear estimation technique. In: *Proc. of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, Denver, CO, USA, pp. 1093 - 1103.

Haessig, D. and B. Friedland (1997). A method for simultaneous state and parameter estimation in nonlinear systems. In: *Proc. of the Amer. Cont. Conf.*, Albuquerque, New Mexico, pp. 947 - 951.

Harman, R.R. and I.Y. Bar-Itzhack (1999). Pseudolinear and state-dependent Riccati equation filters for angular rate estimation. *AIAA J. of Guid., Cont., and Dynamics*, **22**, 723-725.

Hepner, S.A.R. and H.P. Geering (1991). Adaptive two-time-scale tracking filter for target acceleration estimation, *J. of Guidance, Control, and Dynamics*, **14**, 581-588.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Kalman, R.E. and R. Bucy (1961). New results in linear filtering and prediction. Transactions of ASME, 83D.

Mracek, C.P., J.R. Cloutier and C.A. D'Souza (1996). A new technique for nonlinear estimation. In: *Proc. of the 1996 IEEE International Conference on Control Applications*, Dearborn, MI, USA, pp. 338 - 343.

Pappano, V. and B. Friedland (1997). SDRE observer for an induction machine. In: *Proc. of Amer. Cont. Conf.*, Albuquerque, New Mexico.

Sage, A.P. and J.L. Melsa (1971). *Estimation Theory with Applications to Communications and Control*. McGraw-Hill, New York.