

Fisher Information Matrix-based Nonlinear System Conversion for State Estimation

Ming Lei, Christophe Baehr and Pierre Del Moral

Abstract—In practical target tracking, a number of improved measurement conversion techniques have been developed and proofed to be superior to the standard (extended) Kalman filtering (KF) in Cartesian coordinates. The framework of conversion technique exhibits fundamental pros and cons and therefore associated with different performance as pointed out in [1]. In this paper, we show that, based on the fisher information matrix (FIM) which can be evaluated approximately using state estimates online, instead of the usual measurement conversion, an equivalent linear dynamics can be reconstructed from a general nonlinear form, thus even the standard KF can be applied theoretically. The proposed approach is explicitly free of the fundamental limitations of traditional measurement conversion. Simulation results are provided by comparison with a state-of-art conversion method with the so-called optimal linear unbiased estimation presented in [1].

Index Terms—Posterior Cramer-Rao lower bound (CRLB), fisher information matrix (FIM), nonlinear system conversion, measurement conversion

I. INTRODUCTION

In actual target tracking implementation, the tracking is performed in Cartesian coordinates while having to employ the measurements collected in polar or spherical, therefore the methods based on measurement-conversion framework are widely used [2]. Its basic idea is to transform the nonlinear measurement model into a so-called pseudo-linear form in Cartesian coordinates, correspondingly, the bias and covariance of the converted equivalent measurement error are estimated, and perform filtering using KF. It can be seen explicitly that there exists neither the bias introduced by local linearization nor the derivation of Jacobian matrix, therefore it is not surprised that the method outperforms the extended KF (EKF) reasonably.

A number of improved variety of measurement conversion techniques have been proposed [5], [6], [7], [8] with only difference in the way of approximating the bias and covariance of the converted error. The pros and cons of these techniques are surveyed in [2]. Especially in literature [1], a novel estimator based on the sense of best linear unbiased estimate (BLUE) is developed, it focus on optimally transferring nonlinear measurement by calculating some statistics

This research was supported by the INRIA Research Foundation under Grant XXXX.

Dr. Ming Lei is with the INRIA, University Bordeaux-I, Talence 33405, France. (minglei.sa@gmail.com)

Prof. Christophe Baehr is with the CNRM/GAME URA1357, French National Centre for Meteorological Research, Toulouse France. (christophe.baehr@meteo.fr)

Prof. Pierre Del Moral is with the INRIA and Mathematics Institut, University Bordeaux-I, Talence 33405, France. (pierre.del_moral@inria.fr)

approximately, essentially which is an expansion of standard KF.

In this paper, a novel method for tracking with nonlinear dynamics and nonlinear measurements is proposed, which is motivated by recognizing that fisher information matrix (FIM) is nothing but an uniquely criterion of low bound of estimated error covariance for arbitrary (linear/ nonlinear) dynamical system, furthermore, although the KF can not optimally handle any nonlinear system directly, we may still be able to develop the recursive BLUE estimator by constructing an equivalent linear dynamics to replace the original nonlinear system on the sense of both having identical FIM.

The rest of the paper is organized as: In Section II, the nonlinear dynamical model and the calculation of recursive FIM are introduced. In Section III, the principle of transformation of equivalent linear dynamics is presented. In Section IV, the filtering algorithm is schemed. Finally a simulation and some conclusions are presented in Section V and VI.

II. BACKGROUND

A. Nonlinear Dynamical System

In Cartesian coordinates, the state of maneuvering target can be assumed as a discrete time nonlinear form and formulated by

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{\Gamma}_k \mathbf{w}_k, \quad (1)$$

where $k = 1, \dots, L$ and L is the length of the simulation time. The definition of state vector $\mathbf{x}_k \triangleq [x_k, y_k, z_k, \dot{x}_k, \dot{y}_k, \dot{z}_k, s_{1 \times (n-6)}]'$, the superscript $[\cdot]'$ stands for transpose of a matrix or vector. $\mathbf{x}_k \in \mathbb{R}^n$ consists of the position and velocity components along x, y and z directions at time k , $s_{1 \times (n-6)}$ is other state components such as acceleration. Process noise \mathbf{w}_k is assumed as a zero-mean Gaussian process with invertible covariance matrix ζ_k , $\mathbf{F}_k \in \mathbb{R}^{n \times n}$ and $\mathbf{\Gamma}_k$, the state transition matrix and noise input matrix, in general are the constant matrices, respectively.

There exists a spherical coordinates with its original point stems from the Cartesian coordinates, the measurements collected from this spherical coordinates can be formulated as

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k = [r_k, \theta_k, \varphi_k]' + \mathbf{v}_k, \quad (2)$$

where $\mathbf{h}_k(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a nonlinear measurement mapping and here $m = 3$, that is, the measurement consists of $r_k = (x_k^2 + y_k^2 + z_k^2)^{1/2}$, $\theta_k = \tan^{-1}(y_k/x_k)$ and $\varphi_k = \tan^{-1}[z_k/(x_k^2 + y_k^2)^{1/2}]$. Note that generally the dimension $n > m$. $\mathbf{z}_k, \mathbf{v}_k \in \mathbb{R}^3$ denotes the measurement and corresponding measurement noise assumed as zero-mean Gaussian with

invertible diagonal covariance $\mathbf{R}_k = \text{diag}(\sigma_r^2, \sigma_\theta^2, \sigma_\varphi^2)$. \mathbf{w}_k and \mathbf{v}_k are assumed to be independent mutually.

B. Recursive Computation of FIM

Let $\hat{\mathbf{x}}_k$ be an unbiased state estimate based on measurement set that collected up to and including time k , i.e., $Z^k = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k\}$. Let $p_{z,\mathbf{x}}(Z^k, \mathbf{x}) = p_{z|\mathbf{x}}(Z^k|\mathbf{x})p_{\mathbf{x}}(\mathbf{x})$ be the joint probability density function of the pair (\mathbf{z}, \mathbf{x}) , the lower bound of covariance of $\hat{\mathbf{x}}_k$ can be formulated by

$$\mathbf{C}_k \triangleq E[(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)'] \geq \mathbf{J}_k^{-1}, \quad (3)$$

where \mathbf{J}_k is the $n \times n$ fisher information matrix (FIM) with the elements

$$\mathbf{J}_{ij} = E \left[-\frac{\partial^2 \log p_{z,\mathbf{x}}(Z^k, \mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \right], i, j = 1, \dots, n \quad (4)$$

provided that the derivatives and expectations in (3) and (4) are exist. The inverse of FIM, \mathbf{J}_k^{-1} , is referred to as posterior CRLB. The inequality in (3) means that the difference $\mathbf{C}_k - \mathbf{J}_k^{-1}$ is a positive semi-definite matrix. Let notations ∇ and Δ be operators of the first and second-order partial derivatives, $\nabla_\theta = [\partial/\partial\theta_1, \dots, \partial/\partial\theta_n]'$, $\Delta_\theta^\Phi = \nabla_\theta \nabla_\theta^\Phi$. Tichavsky et al. [3] proposed a Riccati-like recursion for the FIM \mathbf{J}_k calculation:

$$\mathbf{J}_{k+1} = \mathbf{D}_k^{22} - \mathbf{D}_k^{21}(\mathbf{J}_k + \mathbf{D}_k^{11})^{-1}\mathbf{D}_k^{12} \quad (k > 0), \quad (5)$$

$$\mathbf{D}_k^{11} = E[-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_{k+1}|\mathbf{x}_k)], \quad (6)$$

$$\mathbf{D}_k^{12} = E[-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k+1}} \log p(\mathbf{x}_{k+1}|\mathbf{x}_k)] = (\mathbf{D}_k^{21})', \quad (7)$$

$$\mathbf{D}_k^{22} = E[-\Delta_{\mathbf{x}_{k+1}}^{\mathbf{x}_{k+1}} \log p(\mathbf{x}_{k+1}|\mathbf{x}_k) + E[-\Delta_{\mathbf{x}_{k+1}}^{\mathbf{z}_{k+1}} \log p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1})], \quad (8)$$

$$\mathbf{J}_0 = E[-\Delta_{\mathbf{x}_0}^{\mathbf{x}_0} \log p(\mathbf{x}_0)]. \quad (9)$$

The expectation $E[\cdot]$ in (6) and (7) is with respect to \mathbf{x}_k and \mathbf{x}_{k+1} , whereas in (8), it is with respect to \mathbf{x}_{k+1} and \mathbf{z}_{k+1} . Note that matrices of \mathbf{D}_k^{11} , \mathbf{D}_k^{12} and \mathbf{D}_k^{22} are all evaluated at the true value of \mathbf{x}_k and \mathbf{x}_{k+1} theoretically. $p(\mathbf{x}_0)$ is the prior density function of target state. If the first-two order moments statistics of the prior distribution $p(\mathbf{x}_0)$ are known and the covariance given by \mathbf{C}_0 , then $\mathbf{J}_0 = \mathbf{C}_0^{-1}$. The matrix \mathbf{D}_k^{22} defined in (8) indicates the dependency of CRLB on both the state and measurement models.

III. FIM-BASED NONLINEAR SYSTEM CONVERSION

A. Derivation Matrices Defined in FIM

From the Gaussian assumption of state \mathbf{x}_k and measurement \mathbf{z}_k as that stated in Section II-A, the log-pdf of \mathbf{x}_{k+1} and \mathbf{z}_{k+1} given \mathbf{x}_k and \mathbf{x}_{k+1} can be written by

$$\begin{aligned} & \ln p(\mathbf{x}_{k+1}|\mathbf{x}_k) \\ &= c_1 - \frac{1}{2} [(\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k))' \mathbf{Q}_k^{-1} (\mathbf{x}_{k+1} - \mathbf{f}_k(\mathbf{x}_k))], \end{aligned} \quad (10)$$

$$\begin{aligned} & \ln p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) \\ &= c_2 - \frac{1}{2} [(\mathbf{z}_{k+1} - \mathbf{h}_{k+1}(\mathbf{x}_{k+1}))' \mathbf{R}_{k+1}^{-1} (\mathbf{z}_{k+1} - \mathbf{h}_{k+1}(\mathbf{x}_{k+1}))], \end{aligned} \quad (11)$$

where c_1 and c_2 are constants. $\mathbf{Q}_k \triangleq \Gamma_k \zeta_k \Gamma_k'$. Then according to the definitions in (6)-(8), it follows directly that

$$\mathbf{D}_k^{11} = E[(\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k)) \mathbf{Q}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k))'], \quad (12)$$

$$\mathbf{D}_k^{12} = -E[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k)] \mathbf{Q}_k^{-1} \quad (13)$$

$$\begin{aligned} \mathbf{D}_k^{22} &= \mathbf{Q}_k^{-1} + E[(\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1})) \\ &\quad \times \mathbf{R}_{k+1}^{-1} (\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1}))']. \end{aligned} \quad (14)$$

B. A Equivalent Form of Nonlinear System With Identical FIM

Without loss of generality, we can find the equivalent matrices $\check{\mathbf{F}}_k$, $\check{\mathbf{H}}_k$, $\check{\mathbf{Q}}_k$ and $\check{\mathbf{R}}_k$ to construct a new linear dynamics like

$$\mathbf{x}_{k+1} = \check{\mathbf{F}}_k \mathbf{x}_k + \delta_k, \quad (15)$$

$$\mathbf{z}_{k+1} = \check{\mathbf{H}}_{k+1} \mathbf{x}_{k+1} + \psi_{k+1}, \quad (16)$$

which has the same fisher information matrix as the original nonlinear system in (1) and (2). Clearly Eqn. (15) is identical to (1) substantially, however we still use different notation to represent it. Similarly we define a new equivalent Eqn.(16) and use it to replace the original nonlinear equation of (2) and map the state to measurement. δ_k and ψ_k correspond to the original process noise $\Gamma_k \mathbf{w}_k$ and measurement noise \mathbf{v}_k in (1) and (2), and are also assumed as the zero-mean Gaussian process noises with zero means and known covariance matrix $\check{\mathbf{Q}}_k$ and $\check{\mathbf{R}}_k$, respectively. Then, based on this constructed linear system, the matrices corresponding to the definitions in (6)-(8) can be derived as,

$$\check{\mathbf{D}}_k^{11} = \check{\mathbf{F}}_k' \check{\mathbf{Q}}_k^{-1} \check{\mathbf{F}}_k, \quad (17)$$

$$\check{\mathbf{D}}_k^{12} = -\check{\mathbf{F}}_k' \check{\mathbf{Q}}_k^{-1}, \quad (18)$$

$$\check{\mathbf{D}}_k^{22} = \check{\mathbf{Q}}_k^{-1} + \check{\mathbf{H}}_{k+1}' \check{\mathbf{R}}_{k+1}^{-1} \check{\mathbf{H}}_{k+1}. \quad (19)$$

One possible solution of the above linear equation group can be formulated as

$$\check{\mathbf{F}}_k = -(\check{\mathbf{D}}_k^{12})^{-1} \check{\mathbf{D}}_k^{11}, \quad (20)$$

$$\check{\mathbf{Q}}_k = (\check{\mathbf{D}}_k^{12})^{-1} \check{\mathbf{D}}_k^{11} (\check{\mathbf{D}}_k^{21})^{-1}, \quad (21)$$

$$\check{\mathbf{H}}_k = \check{\mathbf{R}}_{k+1}^{1/2} [\check{\mathbf{D}}_k^{22} - \check{\mathbf{D}}_k^{21} (\check{\mathbf{D}}_k^{11})^{-1} \check{\mathbf{D}}_k^{12}]^{1/2}, \quad (22)$$

$$\check{\mathbf{R}}_k = \mathbf{R}_k, \quad (23)$$

where $\mathbf{A}^{1/2}$ stands for the square root of a positive semidefinite matrix \mathbf{A} . All the inverses requested in (20)-(22) are assumed to be existent. In order to make the linear system ((15)(16)) to be a equivalent form of the original nonlinear system ((1)(2)) in sense of having identical information matrix, we have the equalities, $\mathbf{D}_k^{11} = \check{\mathbf{D}}_k^{11}$, $\mathbf{D}_k^{12} = \check{\mathbf{D}}_k^{12}$ and $\mathbf{D}_k^{22} = \check{\mathbf{D}}_k^{22}$. Substitution of (12)-(14) into (20)-(22), together

with (23), implies that

$$\check{\mathbf{F}}_k = \mathbf{Q}_k E \left[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k) \right]^{-1} E \left[(\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k)) \mathbf{Q}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k))' \right], \quad (24)$$

$$\check{\mathbf{Q}}_k = \mathbf{Q}_k E \left[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k) \right]^{-1} \times E \left[(\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k)) \mathbf{Q}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k))' \right] \left(E \left[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k) \right]' \right)^{-1} \mathbf{Q}_k, \quad (25)$$

$$\begin{aligned} \check{\mathbf{H}}_k &= \mathbf{R}_{k+1}^{1/2} \left\{ \mathbf{Q}_k^{-1} \right. \\ &+ E \left[(\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1})) \mathbf{R}_{k+1}^{-1} (\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1}))' \right] \left. \right\}^{1/2} \\ &- \mathbf{R}_{k+1}^{1/2} \left\{ \mathbf{Q}_k^{-1} E \left[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k) \right]' \left(E \left[(\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k)) \right. \right. \right. \\ &\left. \left. \left. \times \mathbf{Q}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k))' \right) \right]^{-1} E \left[\nabla_{\mathbf{x}_k} \mathbf{f}'_k(\mathbf{x}_k) \right] \mathbf{Q}_k^{-1} \right\}^{1/2}. \quad (26) \end{aligned}$$

Note that for a dynamical system with a prior state transition map \mathbf{f}_k , the intensity and distribution of process noise, generally, is unknown, however, the assumption of it in Section II-A, can be accepted comprehensively. According to the theory of CRLB, the matrices \mathbf{D}_k^{11} , \mathbf{D}_k^{12} and \mathbf{D}_k^{22} should be evaluated using the true state though it is not acquired actually, however according to the approximation in [4], the matrices can be computed by using the first two moments of state estimates online.

The term $E[(\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1})) \mathbf{R}_{k+1}^{-1} (\nabla_{\mathbf{x}_{k+1}} \mathbf{h}'_{k+1}(\mathbf{x}_{k+1}))']$ in right-hand of (26) is further derived and see details in Appendix.

IV. SCHEME OF EQUIVALENT LINEAR DYNAMICS FILTERING

For the convenience of implementation, we summarize the scheme of the proposed equivalent linear system filtering (ELSF) as below.

One iteration of filtering for the constructed equivalent dynamics specified by (15) and (16) is as follows:

A. Initializes Parameters (at time $k = 0$)

The covariance matrices of the process and measurement noise, \mathbf{Q}_k and \mathbf{R}_k ; Initial state \mathbf{x}_0 and its error covariance \mathbf{P}_0 ; Sampling interval T ; The length of simulation time step N ; The runs of Monte Carlo simulation M .

B. Parameters Computation (at time $k - 1$)

Calculate the matrices $\check{\mathbf{F}}_{k-1}$, $\check{\mathbf{Q}}_{k-1}$ and $\check{\mathbf{H}}_{k-1}$ using the state estimates such as $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$ (see (24)-(26)).

C. Filtering (at time k)

Corresponding to the linear dynamics presented in (15) and (16), the BLUE-based KF can be formulated by

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \check{\mathbf{F}}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{P}_{k|k-1} &= \check{\mathbf{F}}_{k-1} \mathbf{P}_{k-1|k-1} \check{\mathbf{F}}_{k-1}' + \check{\mathbf{Q}}_{k-1}, \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1} \check{\mathbf{H}}_k' (\check{\mathbf{H}}_k \mathbf{P}_{k|k-1} \check{\mathbf{H}}_k' + \mathbf{R}_k)^{-1}, \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \check{\mathbf{H}}_k \hat{\mathbf{x}}_{k|k-1}), \\ \mathbf{P}_{k|k} &= (\mathbf{I}_{n \times n} - \mathbf{K}_k \check{\mathbf{H}}_k) \mathbf{P}_{k|k-1}, \end{aligned}$$

where $\mathbf{I}_{n \times n}$ denotes a $n \times n$ dimensional unit matrix.

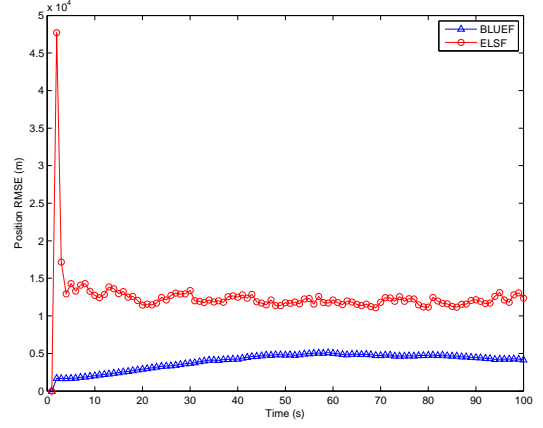


Fig. 1. Comparison of position RMSE.

V. SIMULATION AND DISCUSSION

Considering that it has been shown in [1] that the nonlinear measurement conversion method outperforms the EKF and other methods (e.g., referred to as “fixed-measurement” approaches in [8] and “unbiased/debiased” methods in [6] respectively) explicitly at long range for root-mean-square azimuth of 1.5 or more in terms of the estimation accuracy. We only compare our proposed ELSF with the technique referred as the best linear unbiased estimation filtering (BLUEF) in [1], which is among the top choices in the class of measurement-conversion methods.

Assume that a scenario of a 3-dimensional Cartesian coordinates with a single sensor placed in the origin of the coordinates. The target performs a nearly-constant-velocity (straight line) move. The initial coordinates of target are selected as $[x_0, y_0, z_0] = [-50, 200, 0]km$ with a standard deviation of $1km$. The initial velocity components are also chose from Gaussian distribution with mean $[\dot{x}_0, \dot{y}_0, \dot{z}_0] = [10^3, 0, 0]m/s$ and standard deviation of $10^2m/s$. For the sensor measurement error, it is assumed as a Gaussian with zero-mean and constant covariance $\text{diag}(\sigma_r^2, \sigma_\theta^2, \sigma_\phi^2) = \text{diag}(40^2m^2, 100^2mrad^2, 100^2mrad^2)$. The sensor measurement sampling period is $T = 1s$. The tracking period lasts for 100s. We employ only the root mean square error (RMSE) of position and velocity as comparison criterion and they are defined as $E_k^{\text{posi}} = [\frac{1}{M} \sum_{m=1}^M \sum_{n=1}^3 (\hat{\mathbf{x}}_{k|k,n}^m - \mathbf{x}_{k,n}^m)^2]^{1/2}$ and $E_k^{\text{velo}} = [\frac{1}{M} \sum_{m=1}^M \sum_{n=4}^6 (\hat{\mathbf{x}}_{k|k,n}^m - \mathbf{x}_{k,n}^m)^2]^{1/2}$ respectively, where $k = 1, \dots, N$, time step N here equal to 100. M denotes the runs of Monte Carlo simulation and let $M = 100$. $\hat{\mathbf{x}}_{k|k,n}^m - \mathbf{x}_{k,n}^m$ is the state estimation error of n th component at k th time step and m th run. Fig.1-2 show the comparisons of BLUEF and our ELSF with the process noise \mathbf{w}_k draw from Gaussian distribution with zero mean and a common deviation of $\zeta_k = 1^2 \mathbf{I}_{3 \times 3}$ in x, y, z directions. We can see that there exists an explicit gap between the two approaches. The ELSF corresponding to red circle-line is overall worse than the BLUEF marked by blue triangle-line through all instances whenever for position or velocity. Moreover, ELSF exhibits

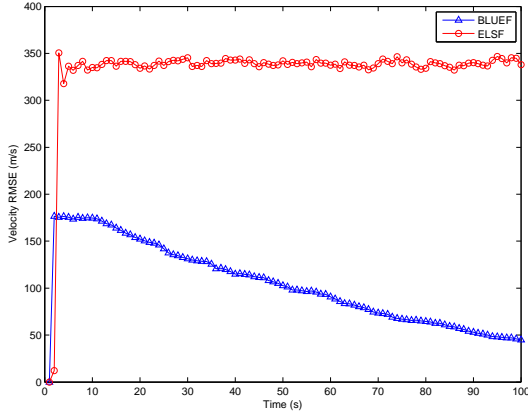


Fig. 2. Comparison of velocity RMSE.

extra higher error in the transient part, then keeps a relatively stability in steady state. In order to have a thoroughly review of ELSF, we compare the two methods by increasing the process noise covariance ζ_k from $1^2 \mathbf{I}_{3 \times 3}$ to $1000^2 \mathbf{I}_{3 \times 3}$ with interval $200^2 \mathbf{I}_{3 \times 3}$, we found the gap of RMSE between the two methods turning to reduce substantially, however the BLUEF is still more accurate than the ELSF, even in the situation of increasing the measurement noise \mathbf{R}_k to 200 percent of current value.

In the simulation conducted, the standard implementation of our proposed ELSF sometimes exhibited numerical problem, we had to use more numerically robust approach. As a comparison, we did not found any numerical problem when performed the BLUEF under the same initial condition and environment.

We can see that the proposed ELSF is valid, but its accuracy is explicitly still inferior to that of the BLUEF with a great quantity. The validity of ELSF depends heavily on an accurate FIM for a nonlinear dynamics, however we made approximation in two place: the one is using the estimates of state instead of true state to calculate the FIM. The other is making only two-order Taylor series expansion of the nonlinear model. These surely introduce much error and reflect in its performance. However the idea of ELSF is thoroughly different from the measurement conversion framework so far, therefore, to improve ELSF by reducing the approximation error will be focused intensively.

VI. CONCLUSIONS

An equivalent linear system filtering (ELSF) in the sense of FIM-based equivalent transformation has been presented for general nonlinear dynamics. The recursive ELSF is valid for general nonlinear system and operates in Cartesian coordinates entirely. Compared with a state-of-art optimal linear unbiased conversion method presented in [1], the simulation results indicate follows.

The proposed ELSF appears to be unacceptably biased in term of accuracy, however, with increasing the process noise or measurement noise or both of them synchronously,

the filtering accuracy can be improved substantially, but still is inferior to the optimal unbiased conversion method. Moreover, ELSF sometimes exhibited numerical problem.

As discussion in Section V, the performance of ELSF depends heavily on how accurate of FIM approximation, and therefore future work will focus on obtaining the well approximation of FIM from the dynamics with high nonlinear and strong noise.

APPENDIX

Derivation of the term $E[(\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k)) \mathbf{R}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k))']$ in right-hand of (26) as follows.

According to definitions in (5), the matrices \mathbf{D}_k^{11} , \mathbf{D}_k^{12} , and \mathbf{D}_k^{22} should be evaluated at the true value of state \mathbf{x}_k and \mathbf{x}_{k+1} theoretically. However, the true state is not available actually, so one feasible choice is to use the state estimate $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{x}}_{k+1|k}$, instead of the true state respectively, to calculate \mathbf{D}_k^{11} , \mathbf{D}_k^{12} , and \mathbf{D}_k^{22} approximately [4]. Therefore we can determine the score function $\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k)$.

From the definition of measurement model (2), The Taylor series expansion of the nonlinear vector-valued measurement $\mathbf{h}_k(\mathbf{x}_k)$ at a "point" of one step prediction of state estimate, $\hat{\mathbf{x}}_{k|k-1}$, while keeps the second-order (quadratic) terms, can be represented by

$$\mathbf{h}_k(\mathbf{x}_k) \approx \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1}) + \hat{\mathbf{G}}_k \tilde{\mathbf{x}}_k + \frac{1}{2} \sum_{i=1}^m \mathbf{e}_i \tilde{\mathbf{x}}_k' \hat{\mathbf{S}}_{k,i} \tilde{\mathbf{x}}_k, \quad (27)$$

where $\mathbf{e}_i \in \mathbb{R}^{m \times 1}$ denotes the i th unit vector. $\tilde{\mathbf{x}}_k \triangleq \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}$ is the state prediction error. $\hat{\mathbf{G}}_k$ denotes Jacobian matrix at $\hat{\mathbf{x}}_{k|k-1}$, which represents the linear (first-order) dependence of the state residual, and show that

$$\begin{aligned} \hat{\mathbf{G}}_k &\triangleq [\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k)]' |_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}} = [\partial h_{k,i}(\mathbf{x}_k) / \partial x_j]_{m \times n} \\ &= \begin{pmatrix} x/l_2 & y/l_2 & z/l_2 & 0 & 0 & 0 \\ -y/l_1^2 & x/l_1^2 & 0 & 0 & 0 & 0 \\ -xz/l_1 l_2^2 & -yz/l_1 l_2^2 & l_1/l_2^2 & 0 & 0 & 0 \end{pmatrix}_k, \quad (28) \end{aligned}$$

where $l_1 = \sqrt{x^2 + y^2}$, $l_2 = \sqrt{x^2 + y^2 + z^2}$. $\hat{\mathbf{S}}_{k,i} \triangleq [\nabla_{\mathbf{x}_k} [\nabla_{\mathbf{x}_k} h_{k,i}(\mathbf{x}_k)']] |_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}$ denotes the Hessian matrix of the i th element $h_{k,i}(\mathbf{x}_k)$ of $\mathbf{h}_k(\mathbf{x}_k)$, when $i = 1, 2, 3$, it can be formulated directly as

$$\begin{aligned} \hat{\mathbf{S}}_{k,1} &= \text{diag}[\hat{\mathbf{S}}_{k,1}^*, \mathbf{0}_{3 \times 3}] |_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}, \\ \hat{\mathbf{S}}_{k,1}^* &= \frac{1}{l_2^2} \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix}_k, \\ \hat{\mathbf{S}}_{k,2} &= \text{diag}[\hat{\mathbf{S}}_{k,2}^*, \mathbf{0}_{4 \times 4}] |_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}, \\ \hat{\mathbf{S}}_{k,2}^* &= \frac{1}{l_1^4} \begin{pmatrix} 2xy & y^2 - x^2 \\ y^2 - x^2 & -2xy \end{pmatrix}_k, \\ \hat{\mathbf{S}}_{k,3} &= \text{diag}[\hat{\mathbf{S}}_{k,3}^*, \mathbf{0}_{3 \times 3}] |_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}, \\ \hat{\mathbf{S}}_{k,3}^* &= \frac{1}{l_1^4 l_2^4} \begin{pmatrix} -z[2x^4 + x^2 y^2 - y^2(y^2 + z^2)] & & & & & \\ -xyz[3(x^2 + y^2) + z^2] & & & & & \\ x(x^2 + y^2 - z^2) l_1^2 & & & & & \\ -xyz[3(x^2 + y^2) + z^2] & x(x^2 + y^2 - z^2) l_1^2 & & & & \\ z[x^4 - 2y^4 + x^2(z^2 - y^2)] & y(x^2 + y^2 - z^2) l_1^2 & & & & \\ y(x^2 + y^2 - z^2) l_1^2 & & & & & 2z l_1^4 \end{pmatrix}_k, \end{aligned}$$

where $\mathbf{0}_{3 \times 3}$ and $\mathbf{0}_{4 \times 4}$ stands for 3×3 and 4×4 -dimensional zero matrices. The definition of l_1 and l_2 are same with that in (28). Then from $\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k) \approx \hat{\mathbf{G}}'_k + \sum_{i=1}^m \mathbf{e}'_i \hat{\mathbf{S}}_{k,i} \tilde{\mathbf{x}}_k$ we have

$$\begin{aligned} & E[(\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k)) \mathbf{R}_k^{-1} (\nabla_{\mathbf{x}_k} \mathbf{h}'_k(\mathbf{x}_k))'] \\ & \approx E \left[\left(\hat{\mathbf{G}}'_k + \sum_{i=1}^m \mathbf{e}'_i \hat{\mathbf{S}}_{k,i} \tilde{\mathbf{x}}_k \right) \mathbf{R}_k^{-1} \left(\hat{\mathbf{G}}_k + \sum_{j=1}^m \tilde{\mathbf{x}}_k \hat{\mathbf{S}}_{k,j} \mathbf{e}_j \right) \right] \\ & = \hat{\mathbf{G}}'_k \mathbf{R}_k^{-1} \hat{\mathbf{G}}_k + E \left[\left(\sum_{i=1}^m \mathbf{e}'_i \hat{\mathbf{S}}_{k,i} \tilde{\mathbf{x}}_k \right) \mathbf{R}_k^{-1} \left(\sum_{j=1}^m \tilde{\mathbf{x}}_k \hat{\mathbf{S}}_{k,j} \mathbf{e}_j \right) \right] \\ & = \hat{\mathbf{G}}'_k \mathbf{R}_k^{-1} \hat{\mathbf{G}}_k + \sum_{i=1}^m \sum_{j=1}^m \mathbf{R}_k^{-1} (\hat{\mathbf{S}}_{k,i} \mathbf{P}_{k|k-1} \hat{\mathbf{S}}_{k,j}), \end{aligned} \quad (29)$$

where $\mathbf{P}_{k|k-1} \triangleq E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k']$, $\mathbf{R}_{k,ij}$ is the (i, j) th element of matrix \mathbf{R}_k , and the term

$$\hat{\mathbf{G}}'_k \mathbf{R}_k^{-1} \hat{\mathbf{G}}_k = \text{diag} \left(\left(\begin{array}{ccc} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{array} \right), \mathbf{0}_{3 \times 3} \right)_k,$$

where

$$\begin{aligned} g_{11} &= \frac{y^2}{\sigma_{\theta_1}^2 l_1^4} + \frac{1}{l_2^2} \left[x^2 (l_2^2 / \sigma_r^2 + z^2 / (\sigma_\varphi^2 l_1^2)) \right], \\ g_{12} &= xy \left[\frac{1}{\sigma_r^2 l_2^2} + \frac{1}{l_1^4} \left((z^2 l_1^2) / (\sigma_\varphi^2 l_2^4) - 1 / \sigma_\theta^2 \right) \right], \\ g_{22} &= \frac{x^2}{\sigma_{\theta_1}^2 l_1^4} + \frac{1}{l_2^2} y^2 \left[l_2^2 / \sigma_r^2 + z^2 / (\sigma_\varphi^2 l_1^2) \right], \\ g_{13} &= \frac{1}{l_2^2} xz (l_2^2 / \sigma_r^2 - 1 / \sigma_\varphi^2), \\ g_{23} &= \frac{1}{l_2^2} yz (l_2^2 / \sigma_r^2 - 1 / \sigma_\varphi^2), \\ g_{33} &= \frac{1}{l_2^2} \left[(l_2^2 z^2) / \sigma_r^2 + l_1^2 / \sigma_\varphi^2 \right]. \end{aligned}$$

REFERENCES

- [1] Zhanlue Z., X. Rong Li, Vesselin P. Jilkov, "Best Linear Unbiased Filtering with Nonlinear Measurements for Target Tracking", *IEEE Trans. on Aerospace and Electronic Systems*, Vol.40, No. 4, 2004, pp.1324-1336.
- [2] Li, X. R., and Jilkov, V. P., "A survey of maneuvering target tracking-Part III: Measurement models", *In Proceedings of the 2001 SPIE Conference on Signal and Data Processing of Small Targets*, Vol. 4473, San Diego, CA, July-Aug. 2001, pp: 423-446.
- [3] P. Tichavsky, C. Muravchik, A. Nehorai, "Posterior Cramer-Rao Bounds for Discrete-Time Nonlinear Filtering", *IEEE Trans. on Signal Processing*, vol.46, No. 5, May 1998, pp: 1386-1396.
- [4] Ming Lei, Barend J. van Wyk and Yong Qi, "Online Estimation of the Approximate Posterior Cramer-Rao Lower Bound for Discrete-time Nonlinear Filtering", *IEEE Trans. on Aerospace and Electronic Systems*, to appear.
- [5] Lerro, D., and Bar-Shalom, Y. "Tracking with debiased consistent converted measurements vs. EKF", *IEEE Trans. on Aerospace and Electronic Systems*, Vol.29, No.3, July 1993, pp: 1015-1022.
- [6] Mo, L., Song, X., Zhou, Y., Sun, Z., and Bar-Shalom, Y., "Unbiased converted measurements for tracking", *IEEE Trans. on Aerospace and Electronic Systems*, Vol.34, No.3, 1998, pp: 1023-1026.
- [7] Suchomski, P., "Explicit expressions for debiased statistics of 3D converted measurements", *IEEE Trans. on Aerospace and Electronic Systems*, Vol.35, No.1, 1999, pp: 368-370.

- [8] Miller, M. D., and Drummond, O. E., "Comparison of methodologies for mitigating coordinate transformation bias in target tracking", *In Proceedings of the 2000 SPIE Conference on Signal and Data Processing of Small Targets*, Vol. 4048, Orlando, FL, Apr. 2000, pp: 414-427.