Recursive Joint Cramér-Rao Lower Bound for Nonlinear Parametric Systems with Colored Noise

Xianqing Li, Zhansheng Duan
Center for Information Engineering Science Research
Xi'an Jiaotong University
Xi'an, Shaanxi 710049, China
Email: lixianqing@stu.xjtu.edu.cn, zsduan@mail.xjtu.edu.cn

Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
Email: uwe.hanebeck@kit.edu

***Abstract*—The performance evaluation for joint state and parameter estimation (JSPE) is of great significance. Joint Cramér-Rao lower bound (JCRLB) has been widely studied for JSPE of nonlinear parametric systems with white noise. However, in practice, the noise is often colored due to high measurement frequency and bandlimited signal channels. In this paper, a recursive JCRLB is developed for JSPE of nonlinear parametric systems with colored noise, characterized by auto-regressive (AR) models. First, we propose a unified recursive JCRLB for JSPE of general nonlinear parametric systems with higher-order autocorrelated process noises and autocorrelated measurement noise simultaneously. Then its relationship with the posterior Cramér-Rao lower bound (PCRLB) for filtering of nonlinear systems with colored noise and the hybrid Cramér-Rao lower bound (HCRLB) for JSPE of regular parametric systems with white noise are provided. Illustrative examples in radar target tracking verify the effectiveness of the proposed JCRLB for the performance evaluation for JSPE of nonlinear parametric systems with colored noise.**

***Index Terms*—JCRLB, Nonlinear parametric systems, Colored noise, Auto-regressive models, Radar target tracking.**

I. Introduction

NONLINEAR filtering technologies have made significant progress for nonlinear systems with white noises. However, due to the complexity of the engineering applications, the assumption of white noise often does not hold. In contrast, the nonlinear systems with colored noise are commonly encountered in engineering applications, such as speech enhancement [1], signal processing [2], target tracking [3], localization and navigation [4], etc. Motivated by these practical applications, it is nontrivial to propose the nonlinear estimators to deal with them. In [5], a state-augmented filter and a measurement-differenced one-step-lag smoother were proposed for systems with higher-order colored noise. Both of these two estimators are optimal under the linear minimum mean square error criterion. A Gaussian approximate filter [6] was proposed for systems with autocorrelated measurement noises, using the measurement difference method. Furthermore, since smoothing is more accurate than the corresponding filtering, fixed-interval, fixed-point, and fixed-lag Gaussian smoothers were developed for nonlinear systems with colored measurement noise [7]. They will be reduced to standard Gaussian smoothers with independent white noise when the correlation is zero.

As is well known, to assess the performance of estimators, the posterior Cramér-Rao lower bound (PCRLB) [8] defined as the inverse of the Fisher information matrix (FIM) has been widely used. It provides a lower bound on the mean square error (MSE) of estimators. Due to the highly complex nature of colored noises, it's more challenging to develop a performance bound for filtering of nonlinear systems with colored noises. By reconstructing the measurement error covariance matrix, a recursive PCRLB was developed for filtering of systems with biased and correlated measurement noise [9]. A unified formula of recursive PCRLB was proposed in [10] for filtering of nonlinear systems with colored noise, which consist of multi-step correlated process noise, multi-step correlated measurement noise and multi-step cross-correlated process and measurement noise simultaneously. Also, it was extended in [11] to nonlinear filters of the nonlinear systems with colored process and measurement noise, characterized by the higher-order auto-regressive (AR) models.

In modern engineering applications, the system models always involve some disturbances, unmodeled effects and unknown inputs, which are usually described as unknown nonrandom parameters. Recently, there has been a surge of interest in joint state and parameter estimation (JSPE) in many fields, such as sensor registration [12], target tracking [13], and signal processing [14], etc. Many representative algorithms have been proposed to deal with the JSPE problems. One of the most well-known schemes used to estimate the state and parameter simultaneously is the joint filter scheme [15], which augments the state vector with the unknown vector parameters. Dual filtering [16] is another hot scheme which uses two separate and parallel filters to estimate them. Hence, this approach usually ignores cross-covariance between states and parameters. Therefore, the joint filter is expected to be more effective than the dual filter for parameter estimation [17]. But they are equivalent when the states and parameters are decoupled. Expectation-maximization (EM) [18] is the third scheme with two iterative steps. In the expectation (E) step, state vectors are estimated by a particular type of filter. Then,

This work is supported in part by National Key Research and Development Plan under Grants 2021YFC2202600 and 2021YFC2202603, and National Natural Science Foundation of China through Grants 62033010, 61773147 and 61673317.

in the maximization (M) step, the nonrandom parameters are identified.

The performance lower bound for JSPE is of importance both for evaluation of estimators and for prediction of the best achievable performance. Thus, a recursive hybrid Cramér-Rao lower bound (HCRLB) was first proposed in [19] for joint kinematic state and nonrandom parameter estimation in ground-moving extended target tracking. In [20], the recursive HCRLB was extended to more general discrete-time Markovian dynamic systems. Furthermore, it was relaxed to nonlinear systems with time-varying measurement parameters in [21]. For JSPE of nonlinear parametric systems with two-adjacent-states dependent (TASD) measurements, a recursive joint Cramér-Rao lower bound (JCRLB) was proposed in [22]. It also provides the recursive JCRLBs for two special TASD systems, where the measurement noises are autocorrelated or cross-correlated with the process noises at one time step apart. For another type of nonlinear parametric systems, in which the process and measurement noises are cross-correlated at the same time, a recursive JCRLB was proposed in [23].

For nonlinear parametric systems with colored noise, the design of joint state and parameter estimators is a great challenge. Meanwhile, the performance evaluation of JSPE for these systems is of equal importance for the design and improvement of estimators. Thus, this paper aims at proposing a performance bound for JSPE of nonlinear parametric systems with colored noises, characterized by autoregressive models including autocorrelated process noise and autocorrelated measurement noise simultaneously. First, we develop a recursive JCRLB for JSPE of the general form of such nonlinear systems. Then, the comparisons of the developed JCRLB with the PCRLB for filtering of nonlinear systems with colored noise and the HCRLB for JSPE of regular parametric systems with white noises are provided. Finally, a numerical example in radar target tracking shows that the proposed JCRLB is effective to evaluate the performance of joint state and parameter estimation for nonlinear parametric systems with colored noise.

The rest of this paper is organized as follows. Section II presents the general form of nonlinear parametric systems with colored noise and formulates the JCRLB problem. Section III develops the recursive JCRLB for such systems. An illustrative example in radar target tracking is provided to verify the effectiveness of the proposed JCRLB in Section IV. Section V concludes this paper.

II. Problem formulation

Consider the following general form of the nonlinear parametric system with colored noise

$$\begin{aligned} \boldsymbol{x}_{k+1} &= f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x) + \boldsymbol{w}_k, \quad (1)\\ \boldsymbol{y}_k &= h_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) + \boldsymbol{v}_k, \quad (2) \end{aligned}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ and $\boldsymbol{y}_k \in \mathbb{R}^m$ are the state and measurement at time k, respectively, $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_z$ are unknown nonrandom parameters, $\langle \boldsymbol{w}_k \rangle$ and $\langle \boldsymbol{v}_k \rangle$ are colored process and measurement noise sequences, satisfying the following pth-order and qth-order AR model, respectively,

$$\boldsymbol{w}_k = \sum_{i=1}^{p} \boldsymbol{\Phi}_{k-i} \boldsymbol{w}_{k-i} + \boldsymbol{\eta}_k, \quad (3)$$

$$\boldsymbol{v}_k = \sum_{j=1}^{q} \boldsymbol{\Psi}_{k-j} \boldsymbol{v}_{k-j} + \boldsymbol{\xi}_k, \quad (4)$$

where $p \geq 1$ and $q \geq 1$, $\langle \boldsymbol{\eta}_k \rangle$ and $\langle \boldsymbol{\xi}_k \rangle$ are mutually independent white noise sequences with probability density functions (PDFs) $p(\boldsymbol{\eta}_k|\boldsymbol{\beta})$ and $p(\boldsymbol{\xi}_k|\boldsymbol{\gamma})$, respectively, which involve unknown nonrandom parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and the initial state x_0 is independent of them with PDF $p(\boldsymbol{x}_0|\boldsymbol{\alpha})$, which incorporates unknown nonrandom parameter vector $\boldsymbol{\alpha}$. The joint estimand (quantity to be estimated) consists of the state $\boldsymbol{x}_k$ and nonrandom parameter vector $\boldsymbol{\theta} = [\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\theta}_x', \boldsymbol{\theta}_z']' \in \mathbb{R}^r$, which collects all the unknown nonrandom parameters.

For simplicity, the following notations are introduced.

TABLE I
NOMENCLATURE

Notations	Meanings
$\boldsymbol{X}^k = [\boldsymbol{x}_0', \cdots, \boldsymbol{x}_k']'$	accumulated state
$\boldsymbol{Z}^k = [\boldsymbol{z}_1', \cdots, \boldsymbol{z}_k']'$	accumulated measurement
$\boldsymbol{\chi}^k = [(\boldsymbol{X}^k)', \boldsymbol{\theta}']'$	joint estimand of $\boldsymbol{X}^k$ and $\boldsymbol{\theta}$
$\boldsymbol{\chi}_k = [\boldsymbol{x}_k', \boldsymbol{\theta}']'$	joint estimand of $\boldsymbol{x}_k$ and $\boldsymbol{\theta}$
$\hat{\boldsymbol{\chi}}^k$	estimate of $\boldsymbol{\chi}^k$
$\hat{\boldsymbol{\chi}}_k$	estimate of $\boldsymbol{\chi}_k$

For discrete-time nonlinear parametric systems, the joint FIM about $\boldsymbol{\chi}^k$ is defined as

$$\boldsymbol{J}^k \triangleq E[-\Delta_{\boldsymbol{\chi}^k}^{\boldsymbol{\chi}^k} \ln p(\boldsymbol{X}^k, \boldsymbol{Z}^k|\boldsymbol{\theta})]\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad (5)$$

where $p(\boldsymbol{X}^k, \boldsymbol{Z}^k|\boldsymbol{\theta})$ is the joint conditional PDF, Δ denotes the second-order derivative operator, i.e., $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} = \nabla_{\boldsymbol{a}} \nabla_{\boldsymbol{b}}'$, and ∇ denotes the gradient operator, $\boldsymbol{\theta}_0$ is the true value of unknown nonrandom parameter $\boldsymbol{\theta}$.

From the Cramér-Rao inequality, the MSE of the joint estimate $\hat{\boldsymbol{\chi}}^k$ satisfying certain regularity conditions as in [20] is bounded from below by the inverse of the joint FIM $\boldsymbol{J}^k$ [8], [20]

$$E[(\hat{\boldsymbol{\chi}}^k - \boldsymbol{\chi}^k)(\hat{\boldsymbol{\chi}}^k - \boldsymbol{\chi}^k)'] \geq (\boldsymbol{J}^k)^{-1}. \quad (6)$$

Whereas what we need to obtain is the joint FIM about $\boldsymbol{\chi}_k$, i.e., $\boldsymbol{J}_k$, which is defined as the inverse of the $(n+r)\times(n+r)$ right-lower block of $(\boldsymbol{J}^k)^{-1}$. The MSE of a joint estimate $\hat{\boldsymbol{\chi}}_k$ is bounded from below by the inverse of $\boldsymbol{J}_k$ [20], [21]:

$$E[(\hat{\boldsymbol{\chi}}_k - \boldsymbol{\chi}_k)(\hat{\boldsymbol{\chi}}_k - \boldsymbol{\chi}_k)'] \geq \boldsymbol{J}_k^{-1}. \quad (7)$$

Unlike in regular parametric systems, the process noise $\boldsymbol{w}_k$ and measurement noise $\boldsymbol{v}_k$ in nonlinear system (1)-(2) are not mutually independent. They satisfy the higher-order AR model simultaneously. Next, the main goal is to develop a recursion for $\boldsymbol{J}_k$ for this system without manipulating large matrix $(\boldsymbol{J}^k)^{-1}$.

III. JCRLB for JSPE of Systems with Colored Noises

A. JCRLB

To deal with colored noise, we first construct an equivalent system as follows.

***Lemma** 1:* The nonlinear parametric system (1)–(4) is equivalent to the following system

$$\boldsymbol{x}_{k+1} = l_k(\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta}_x) + \boldsymbol{\eta}_k, \tag{8}$$

$$\boldsymbol{z}_k = g_k(\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-q}, \boldsymbol{\theta}_z) + \boldsymbol{\xi}_k, \tag{9}$$

where

$$\begin{aligned} l_k(\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta}_x) &= f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x) + \sum_{i=1}^{p} \boldsymbol{\Phi}_{k-i}(\boldsymbol{x}_{k+1-i} \\ &\quad - f_{k-i}(\boldsymbol{x}_{k-i}, \boldsymbol{\theta}_x)), \\ g_k(\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-q}, \boldsymbol{\theta}_z) &= h_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) - \sum_{j=1}^{q} \boldsymbol{\Psi}_{k-j} h_{k-j}(\boldsymbol{x}_{k-j}, \boldsymbol{\theta}_z), \end{aligned}$$

$$\boldsymbol{z}_k = \boldsymbol{y}_k - \sum_{j=1}^{q} \boldsymbol{\Psi}_{k-j} \boldsymbol{y}_{k-j}.$$

Proof: See Appendix A.

Remark 1: The driven noise sequences $\langle \boldsymbol{\eta}_k \rangle$ and $\langle \boldsymbol{\xi}_k \rangle$ in nonlinear system (8)–(9) are mutually independent white noise, which are not limited to be Gaussian.

In order to obtain the recursion of FIM $\boldsymbol{J}_{k+1}$, some notations are introduced here in advance.

1) $\boldsymbol{L}_k^{i,j}, i,j = 1,2,\cdots,l$ denotes the i-th row and j-th column block of the $l \times l$ martix $\boldsymbol{L}_k^{x,x}$ at time k. If $i \leq 0$, or $j \leq 0$, then $\boldsymbol{L}_k^{i,j} = \boldsymbol{0}$. If $i > l$, or $j > l$, then $\boldsymbol{L}_k^{i,j} = \boldsymbol{0}$.

2) $\boldsymbol{L}_k^{i,\theta}, i = 1,2,\cdots,l$ denotes the i-th row block of the matrix $\boldsymbol{L}_k^{x,\theta}$ at time k. If $i \leq 0$, then $\boldsymbol{L}_k^{i,\theta} = \boldsymbol{0}$. If $i > l$, then $\boldsymbol{L}_k^{i,\theta} = \boldsymbol{0}$.

Then, the recursion of the JCRLB for joint state and parameter estimation of the nonlinear parametric systems (1)-(4) can be obtained as follows

***Theorem** 1:* Partitioning the FIM $\boldsymbol{J}_{k+1}$ about $\boldsymbol{x}_{k+1}$ and $\boldsymbol{\theta}$ as

$$\boldsymbol{J}_{k+1} = \begin{bmatrix} \boldsymbol{J}_{k+1}^{x,x} & \boldsymbol{J}_{k+1}^{x,\theta} \\ \boldsymbol{J}_{k+1}^{\theta,x} & \boldsymbol{J}_{k+1}^{\theta,\theta} \end{bmatrix}, \tag{10}$$

then the recursion of $\boldsymbol{J}_{k+1}$ can be obtained as

$$\begin{cases} \boldsymbol{J}_{k+1}^{x,x} = \boldsymbol{D}_k^{22} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{L}_k^{x,x}]^{-1}\boldsymbol{D}_k^{12} \\ \boldsymbol{J}_{k+1}^{x,\theta} = \boldsymbol{D}_k^{23} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{L}_k^{x,x}]^{-1}(\boldsymbol{D}_k^{13} + \boldsymbol{L}_k^{x,\theta}) \\ \boldsymbol{J}_{k+1}^{\theta,\theta} = \boldsymbol{D}_k^{33} + \boldsymbol{L}_k^{\theta,\theta} - (\boldsymbol{D}_k^{31} + \boldsymbol{L}_k^{\theta,x})[\boldsymbol{D}_k^{11} + \boldsymbol{L}_k^{x,x}]^{-1} \\ \qquad\quad \cdot (\boldsymbol{D}_k^{13} + \boldsymbol{L}_k^{x,\theta}), \end{cases} \tag{11}$$

where $\boldsymbol{J}_{k+1}^{x,\theta} = (\boldsymbol{J}_{k+1}^{\theta,x})'$, $\boldsymbol{L}_k^{x,\theta} = (\boldsymbol{L}_k^{\theta,x})'$ and the involved terms can be obtained as

$$\begin{aligned} \boldsymbol{L}_k^{i,j} &= \boldsymbol{L}_{k-1}^{i+1,j+1} + \boldsymbol{B}_{k-1}^{i+1,j+1} + \boldsymbol{C}_{k-1}^{i+1,j+1} - [\boldsymbol{L}_{k-1}^{i+1,1} + \boldsymbol{B}_{k-1}^{i+1,1} \\ &\quad + \boldsymbol{C}_{k-1}^{i+1,1}][\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,j+1} \\ &\quad + \boldsymbol{B}_{k-1}^{1,j+1} + \boldsymbol{C}_{k-1}^{1,j+1}] \end{aligned}$$

$$\begin{aligned} \boldsymbol{L}_k^{i,\theta} &= \boldsymbol{L}_{k-1}^{i+1,\theta} + \boldsymbol{E}_{k-1}^{i+1,\theta} + \boldsymbol{F}_{k-1}^{i+1,\theta} - [\boldsymbol{L}_{k-1}^{i+1,1} + \boldsymbol{B}_{k-1}^{i+1,1} \\ &\quad + \boldsymbol{C}_{k-1}^{i+1,1}][\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} \\ &\quad + \boldsymbol{F}_{k-1}^{1,\theta}] = (\boldsymbol{L}_k^{\theta,i})' \end{aligned}$$

$$\begin{aligned} \boldsymbol{L}_k^{\theta,\theta} &= \boldsymbol{L}_{k-1}^{\theta,\theta} + \boldsymbol{E}_{k-1}^{\theta,\theta} + \boldsymbol{F}_{k-1}^{\theta,\theta} - [\boldsymbol{L}_{k-1}^{\theta,1} + \boldsymbol{E}_{k-1}^{\theta,1} + \boldsymbol{F}_{k-1}^{\theta,1}] \\ &\quad \cdot [\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} + \boldsymbol{F}_{k-1}^{1,\theta}] \end{aligned}$$

for $i,j = 1,2,\cdots,l+1$, $l = \max\{p, q-1\}$

$$\boldsymbol{D}_k^{11} = \begin{bmatrix} \boldsymbol{B}_k^{1,1} + \boldsymbol{C}_k^{1,1} & \cdots & \boldsymbol{B}_k^{1,l+1} + \boldsymbol{C}_k^{1,l+1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_k^{l+1,1} + \boldsymbol{C}_k^{l+1,1} & \cdots & \boldsymbol{B}_k^{l+1,l+1} + \boldsymbol{C}_k^{l+1,l+1} \end{bmatrix}$$

$$\boldsymbol{D}_k^{12} = \begin{bmatrix} \boldsymbol{B}_k^{1,l+2} + \boldsymbol{C}_k^{1,l+2} \\ \vdots \\ \boldsymbol{B}_k^{l+1,l+2} + \boldsymbol{C}_k^{l+1,l+2} \end{bmatrix} = (\boldsymbol{D}_k^{21})'$$

$$\boldsymbol{D}_k^{22} = \boldsymbol{B}_k^{l+2,l+2} + \boldsymbol{C}_k^{l+2,l+2}$$

$$\boldsymbol{D}_k^{13} = \begin{bmatrix} \boldsymbol{E}_k^{1,\theta} + \boldsymbol{F}_k^{1,\theta} \\ \vdots \\ \boldsymbol{E}_k^{l+1,\theta} + \boldsymbol{F}_k^{l+1,\theta} \end{bmatrix} = (\boldsymbol{D}_k^{13})'$$

$$\boldsymbol{D}_k^{23} = \boldsymbol{E}_k^{l+2,\theta} + \boldsymbol{F}_k^{l+2,\theta} = (\boldsymbol{D}_k^{23})'$$

$$\boldsymbol{D}_k^{33} = \boldsymbol{E}_k^{\theta,\theta} + \boldsymbol{F}_k^{\theta,\theta}$$

and

$$\boldsymbol{B}_k^{i,j} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta})], \\ \hfill \boldsymbol{\Phi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})], \quad \boldsymbol{\Phi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{C}_k^{i,j} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \\ \hfill \boldsymbol{x}_{k-q+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{E}_k^{i,\theta} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta})], \\ \hfill \boldsymbol{\Phi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})], \quad \boldsymbol{\Phi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{F}_k^{i,\theta} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \\ \hfill \boldsymbol{x}_{k-q+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{E}_k^{\theta,\theta} = \begin{cases} E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta})], & \boldsymbol{\Phi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})], & \boldsymbol{\Phi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{F}_k^{\theta,\theta} = \begin{cases} E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \\ \hfill \boldsymbol{x}_{k-q+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})], \quad \boldsymbol{\Psi}_{k-j} = \boldsymbol{0} \end{cases}$$

Proof: See Appendix C.

Remark 2: The initial Fisher information submatrix $\boldsymbol{L}_0^{x,x}$, $\boldsymbol{L}_0^{x,\theta}$ and $\boldsymbol{L}_0^{\theta,\theta}$ can be obtained from the joint PDF $p(\boldsymbol{X}^l, \boldsymbol{Z}^l|\boldsymbol{\theta})$ and the definitions in Eqs. (43)-(45), where $l = \max\{p, q-1\}$.

B. Relationship with the HCRLB for JSPE of Regular Parametric Systems with White Noise

In the regular parametric system, the process noise $\langle \boldsymbol{w}_k \rangle$ and the measurement noise $\langle \boldsymbol{v}_k \rangle$ are mutually independent white noise sequences, i.e., $\boldsymbol{\Phi}_{k-i} = 0$ and $\boldsymbol{\Psi}_{k-j} = 0$ in the nonlinear parametric system (1)-(4). Then the joint PDF p_{k+1} in (35) will be reduced to

$$\begin{aligned} p_{k+1} &\triangleq p(\boldsymbol{X}^{k+1}, \boldsymbol{Z}^{k+1}|\boldsymbol{\theta}) \\ &= p_k p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta}) p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta}). \end{aligned} \tag{12}$$

Correspondingly, due to $l = 0$, then

$$\begin{aligned} \boldsymbol{L}_k^{x,x} &= \boldsymbol{L}_k^{1,1} \\ &= \boldsymbol{B}_{k-1}^{2,2} + \boldsymbol{C}_{k-1}^{2,2} - \boldsymbol{B}_{k-1}^{1,1}[\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1}]^{-1}\boldsymbol{B}_{k-1}^{1,2} \\ &= \boldsymbol{J}_k^{x,x}. \end{aligned} \tag{13}$$

Similarly, we have $\boldsymbol{L}_k^{x,\theta} = \boldsymbol{J}_k^{x,\theta}$ and $\boldsymbol{L}_k^{\theta,\theta} = \boldsymbol{J}_k^{\theta,\theta}$.

Then ***Theorem 1*** will be reduced to

$$\begin{cases} \boldsymbol{J}_{k+1}^{x,x} = \boldsymbol{D}_k^{22} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1}\boldsymbol{D}_k^{12} \\ \boldsymbol{J}_{k+1}^{x,\theta} = \boldsymbol{D}_k^{23} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1}(\boldsymbol{D}_k^{13} + \boldsymbol{J}_k^{x,\theta}) \\ \boldsymbol{J}_{k+1}^{\theta,\theta} = \boldsymbol{D}_k^{33} + \boldsymbol{J}_k^{\theta,\theta} - (\boldsymbol{D}_k^{31} + \boldsymbol{J}_k^{\theta,x})[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1} \\ \qquad \cdot (\boldsymbol{D}_k^{13} + \boldsymbol{J}_k^{x,\theta}), \end{cases} \tag{14}$$

where

$$\begin{cases} \boldsymbol{D}_k^{11} = \boldsymbol{B}_k^{11}, \boldsymbol{D}_k^{12} = \boldsymbol{B}_k^{12}, \boldsymbol{D}_k^{22} = \boldsymbol{B}_k^{22} + \boldsymbol{C}_k^{22} \\ \boldsymbol{D}_k^{13} = \boldsymbol{E}_k^{1,\theta}, \boldsymbol{D}_k^{23} = \boldsymbol{E}_k^{2,\theta} + \boldsymbol{F}_k^{2,\theta}, \boldsymbol{D}_k^{33} = \boldsymbol{E}_k^{\theta,\theta} + \boldsymbol{F}_k^{\theta,\theta} \end{cases} \tag{15}$$

$$\begin{cases} \boldsymbol{B}_k^{1,1} = E[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{B}_k^{1,2} = E[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{B}_k^{2,2} = E[-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{C}_k^{2,2} = E[-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \\ \boldsymbol{E}_k^{1,\theta} = E[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{E}_k^{2,\theta} = E[-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{F}_k^{2,\theta} = E[-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \\ \boldsymbol{E}_k^{\theta,\theta} = E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta})] \\ \boldsymbol{F}_k^{\theta,\theta} = E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta})]. \end{cases} \tag{16}$$

Eqs. (14)-(16) are exactly the joint FIM for JSPE of regular parametric systems with white noises in [20]. Obviously, it is a special case of the JCRLB proposed in ***Theorem 1***.

C. Relationship with the PCRLB for Filtering of Systems with Colored Noise

Suppose that the discrete-time system (1)-(4) is reduced to the following nonlinear system

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{w}_k), \tag{17}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \boldsymbol{v}_k), \tag{18}$$

$$\boldsymbol{w}_k = \sum_{i=1}^{p} \boldsymbol{\Phi}_{k-i}\boldsymbol{w}_{k-i} + \boldsymbol{\eta}_k, \tag{19}$$

$$\boldsymbol{v}_k = \sum_{j=1}^{q} \boldsymbol{\Psi}_{k-j}\boldsymbol{v}_{k-j} + \boldsymbol{\xi}_k, \tag{20}$$

where $\langle \boldsymbol{\eta}_k \rangle$ and $\langle \boldsymbol{\xi}_k \rangle$ are mutually independent white noise sequences and both of them are independent of the initial state $\boldsymbol{x}_0$. The nonlinear system (17)-(20) does not depend on any nonrandom parameters, i.e., $\boldsymbol{\theta} \in \varnothing$. Then the joint PDFs p_{k+1} in (35) will be reduced to

$$\begin{aligned} p_{k+1} &\triangleq p(\boldsymbol{X}^{k+1}, \boldsymbol{Z}^{k+1}) \\ &= p(\boldsymbol{X}^k, \boldsymbol{Z}^k) p(\boldsymbol{x}_{k+1}|\boldsymbol{X}^k, \boldsymbol{Z}^k) p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{X}^k, \boldsymbol{Z}^k) \\ &= p_k p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}) p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_{k-q+1}). \end{aligned} \tag{21}$$

Due to $\boldsymbol{\theta} \in \varnothing$, we have $\boldsymbol{J}_k^{x,\theta} = 0$ and $\boldsymbol{J}_k^{\theta,\theta} = 0$. Correspondingly, ***Theorem 1*** will be reduced to

$$\boldsymbol{J}_{k+1}^{x,x} = \boldsymbol{D}_k^{22} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{L}_k^{x,x}]^{-1}\boldsymbol{D}_k^{12}, \tag{22}$$

where the involved terms can be obtained as

$$\boldsymbol{D}_k^{11} = \begin{bmatrix} \boldsymbol{B}_k^{1,1} + \boldsymbol{C}_k^{1,1} & \cdots & \boldsymbol{B}_k^{1,l+1} + \boldsymbol{C}_k^{1,l+1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_k^{l+1,1} + \boldsymbol{C}_k^{l+1,1} & \cdots & \boldsymbol{B}_k^{l+1,l+1} + \boldsymbol{C}_k^{l+1,l+1} \end{bmatrix}$$

$$\boldsymbol{D}_k^{12} = \begin{bmatrix} \boldsymbol{B}_k^{1,l+2} + \boldsymbol{C}_k^{1,l+2} \\ \vdots \\ \boldsymbol{B}_k^{l+1,l+2} + \boldsymbol{C}_k^{l+1,l+2} \end{bmatrix}$$

$$\boldsymbol{D}_k^{22} = \boldsymbol{B}_k^{l+2,l+2} + \boldsymbol{C}_k^{l+2,l+2}$$

$$\begin{aligned} \boldsymbol{L}_k^{i,j} &= \boldsymbol{L}_{k-1}^{i+1,j+1} + \boldsymbol{B}_{k-1}^{i+1,j+1} + \boldsymbol{C}_{k-1}^{i+1,j+1} - [\boldsymbol{L}_{k-1}^{i+1,1} + \boldsymbol{B}_{k-1}^{i+1,1} \\ &\quad + \boldsymbol{C}_{k-1}^{i+1,1}][\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,j+1} + \boldsymbol{B}_{k-1}^{1,j+1} \\ &\quad + \boldsymbol{C}_{k-1}^{1,j+1}] \end{aligned}$$

for $i, j = 1, 2, \cdots, l+1, \quad l = \max\{p, q-1\}$,

where

$$\boldsymbol{B}_k^{i,j} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p})], \\ \qquad\qquad \boldsymbol{\Phi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)], \quad \boldsymbol{\Phi}_{k-j} = \boldsymbol{0} \end{cases}$$

$$\boldsymbol{C}_k^{i,j} = \begin{cases} E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots \\ \qquad\qquad \boldsymbol{x}_{k-q+1})], \quad \boldsymbol{\Psi}_{k-j} \neq \boldsymbol{0} \\ E[-\Delta_{\boldsymbol{x}_{k-l-1+i}}^{\boldsymbol{x}_{k-l-1+j}} \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1})], \quad \boldsymbol{\Psi}_{k-j} = \boldsymbol{0} \end{cases}$$

The above recursion of FIM in (22) is exactly the FIM for filtering of nonlinear systems with colored noises in [11]. Obviously, it is a special case of the JCRLB proposed in ***Theorem 1***.

IV. Illustrative Example

In this section, a numerical example in radar target tracking is provided to demonstrate the utility of the proposed JCRLB for JSPE of the nonlinear parametric systems with colored noise.

Consider a target with coordinated turn motion in a 2D plane [11], [26]. Its motion equation is modeled as

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} & 0 & \frac{\cos \omega T - 1}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \frac{1 - \cos \omega T}{\omega} & 1 & \frac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix} \boldsymbol{x}_k + \boldsymbol{w}_k, \tag{23}$$

where $\boldsymbol{x}_k = [\mathrm{x}_k, \dot{\mathrm{x}}_k, \mathrm{y}_k, \dot{\mathrm{y}}_k]'$ is the state vector, $T = 1$s is the sampling interval, $\omega = 2^{\circ}\mathrm{s}^{-1}$ is the turning rate, a known nonrandom parameter, and $\langle \boldsymbol{w}_k \rangle$ is the process noise.

A 2D radar is mounted at the origin of the 2D plane to measure the range r_k and the bearing θ_k. Then the measurement equation can be modeled as

$$\boldsymbol{z}_k^m = \begin{bmatrix} r_k^m \\ \theta_k^m \end{bmatrix} = \begin{bmatrix} \sqrt{\mathrm{x}_k^2 + \mathrm{y}_k^2} \\ \tan^{-1}(\frac{\mathrm{y}_k}{\mathrm{x}_k}) \end{bmatrix} + \begin{bmatrix} \Delta_r \\ \Delta_\theta \end{bmatrix} + \boldsymbol{v}_k, \tag{24}$$

where Δ_r and Δ_θ are range and bearing measurement biases, which are unknown nonrandom parameters with ground truth 20 m and 5 mrad, respectively, and $\langle \boldsymbol{v}_k \rangle$ is the measurement noise.

In this example, we assume that $\boldsymbol{w}_k$ in (23) is a first-order colored noise and modeled as

$$\boldsymbol{w}_k = 0.9 \boldsymbol{I} \boldsymbol{w}_{k-1} + \boldsymbol{\eta}_k, \tag{25}$$

where $\boldsymbol{I}$ is a identity matrix, $\boldsymbol{\eta}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$ is a driven noise with

$$\boldsymbol{Q}_k = S_w \begin{bmatrix} \frac{2(\omega T - \sin \omega T)}{\omega^3} & \frac{1-\cos \omega T}{\omega^2} & 0 & \frac{(\omega T - \sin \omega T)}{\omega^2} \\ \frac{1-\cos \omega T}{\omega^2} & T & -\frac{(\omega T - \sin \omega T)}{\omega^2} & 0 \\ 0 & -\frac{(\omega T - \sin \omega T)}{\omega^2} & \frac{2(\omega T - \sin \omega T)}{\omega^3} & \frac{1-\cos \omega T}{\omega^2} \\ \frac{(\omega T - \sin \omega T)}{\omega^2} & 0 & \frac{1-\cos \omega T}{\omega^2} & T \end{bmatrix}, \tag{26}$$

where $S_w = 0.1\ \mathrm{m}^2\mathrm{s}^{-3}$ is the power spectral density.

The measurement noise $\boldsymbol{v}_k$ in (24) is also first-order colored and modeled as

$$\boldsymbol{v}_k = 0.8 \boldsymbol{I} \boldsymbol{v}_{k-1} + \boldsymbol{\xi}_k, \tag{27}$$

where $\boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k)$ is a driven white noise with $\boldsymbol{R}_k = \mathrm{diag}(\sigma_r^2, \sigma_\theta^2)$, $\langle \boldsymbol{\eta}_k \rangle$ and $\langle \boldsymbol{\xi}_k \rangle$ are mutually independent, and both of them are independent of the initial state $\boldsymbol{x}_0$. $\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{P}_0)$ with

$$\bar{\boldsymbol{x}}_0 = [2000\ \mathrm{m}, 100\ \mathrm{ms}^{-1}, 4000\ \mathrm{m}, 10\ \mathrm{ms}^{-1}]',$$
$$\boldsymbol{P}_0 = \mathrm{diag}(100^2\ \mathrm{m}^2, 100^2\ \mathrm{m}^2\mathrm{s}^{-2}, 100^2\ \mathrm{m}^2, 10^2\ \mathrm{m}^2\mathrm{s}^{-2}).$$

For this radar target tracking example with first-order colored process and measurement noises simultaneously, we can obtain the following equivalent parametric system as

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k) + \boldsymbol{\Phi}_{k-1}(\boldsymbol{x}_k - f_{k-1}(\boldsymbol{x}_{k-1})) + \boldsymbol{\eta}_k, \tag{28}$$

$$\begin{aligned} \boldsymbol{z}_k &= \boldsymbol{z}_k^m - \boldsymbol{\Psi}_{k-1} \boldsymbol{z}_{k-1}^m \\ &= h_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) - \boldsymbol{\Psi}_{k-1} h_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z) + \boldsymbol{\xi}_k. \end{aligned} \tag{29}$$

To show how hard it is to jointly estimate the target state and radar measurement biases using the proposed JCRLB, we consider three different settings for σ_r and σ_θ in Table II. The JCRLBs are obtained over 500 Monte Carlo runs.

We note that the uncertainty of the measurement in equivalent system (28)-(29) is only dependent on the driven noise $\langle \boldsymbol{\xi}_k \rangle$. Thus if the covariance $\boldsymbol{R}_k$ of driven noise increases, then the uncertainty of the measurement increases. So we expect that the accuracy of the estimation will be reduced as $\boldsymbol{R}_k$ increases.

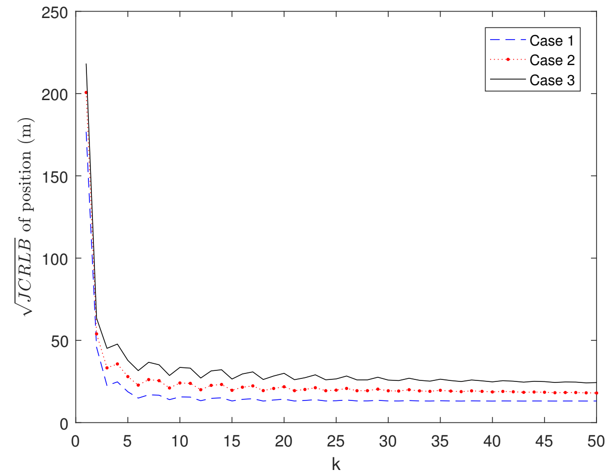

Fig. 1. $\sqrt{\text{JCRLB}}$ for position under different driven noise level.

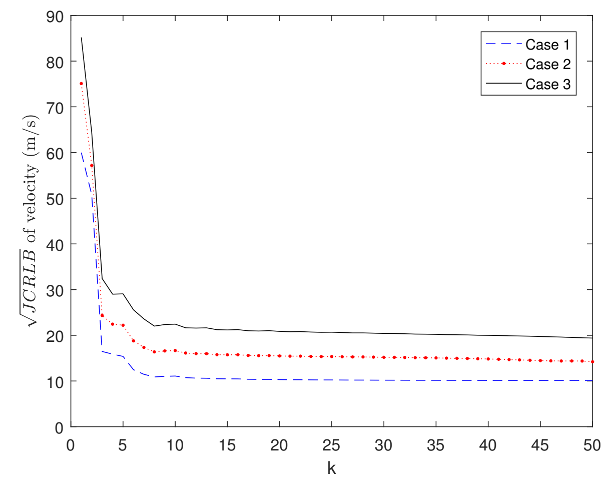

Fig. 2. $\sqrt{\text{JCRLB}}$ for velocity under different driven noise level.

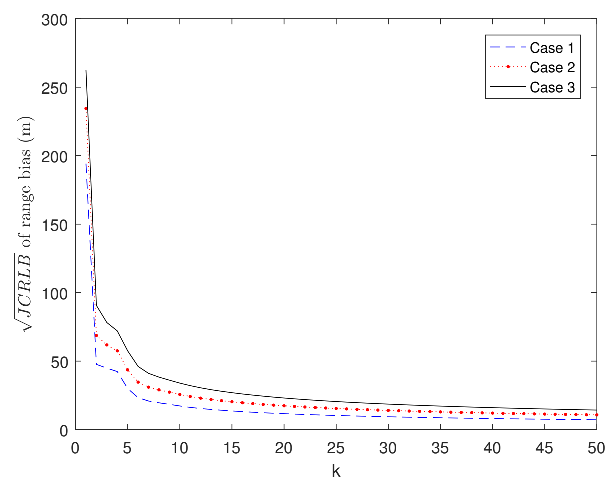

Fig. 3. $\sqrt{\text{JCRLB}}$ for range bias under different driven noise level.

TABLE II
THE STANDARD DEVIATION OF $\boldsymbol{\xi}_k$

	Range	Bearing
Case 1	$\sigma_r = 10$ m	$\sigma_\theta = 3$ mrad
Case 2	$\sigma_r = 15$ m	$\sigma_\theta = 5$ mrad
Case 3	$\sigma_r = 20$ m	$\sigma_\theta = 8$ mrad

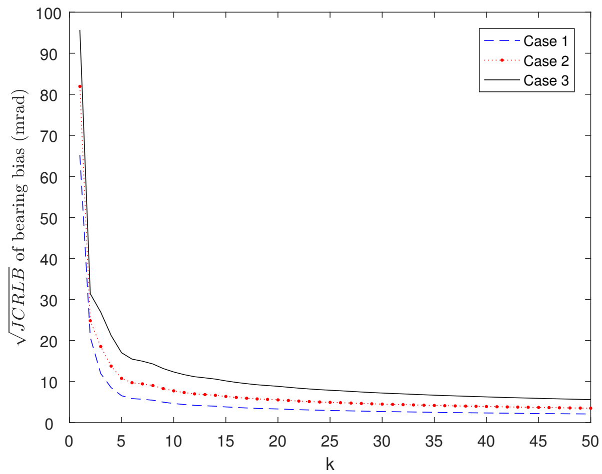

Fig. 4. $\sqrt{\text{JCRLB}}$ for bearing bias under different driven noise level.

Figs. 1-4 show that the $\sqrt{\text{JCRLB}}$s of position, velocity, range bias Δ_r and bearing bias Δ_θ increase as the covariance $\boldsymbol{R}_k$ of driven noise increases. It indicates that the JCRLBs for the nonlinear system (23)-(24) get larger when $\boldsymbol{R}_k$ gets larger. That is, the larger the driven noise level is, the more difficult it is to jointly estimate the target motion state and radar measurement biases. Obviously, this is consistent with our expectations.

V. Conclusion

The performance of joint state and parameter estimators is of importance and can be lower bounded by the JCRLB. In this paper, we have proposed a recursive JCRLB for joint state and parameter estimation of the nonlinear parametric systems, where the process noises and measurement noises are both autocorrelated satisfying the higher-order auto-regressive models. Its relationship with HCRLB for JSPE of regular parametric systems with white noises and PCRLB for filtering of nonlinear systems with colored noises have been provided as well. It is found that both of them are special cases of the newly developed JCRLB.

Appendix A
Proof of Lemma 1

From (1)-(2), we have

$$\boldsymbol{w}_{k-i} = \boldsymbol{x}_{k+1-i} - f_{k-i}(\boldsymbol{x}_{k-i}, \boldsymbol{\theta}_x),\ i = 1, 2, \cdots, p, \quad (30)$$

$$\boldsymbol{v}_{k-j} = \boldsymbol{z}_{k-j} - h_{k-j}(\boldsymbol{x}_{k-j}, \boldsymbol{\theta}_z),\ j = 1, 2, \cdots, q. \quad (31)$$

Then (3)-(4) can be rewritten as

$$\boldsymbol{w}_k = \sum_{i=1}^{p} \boldsymbol{\Phi}_{k-i}(\boldsymbol{x}_{k+1-i} - f_{k-i}(\boldsymbol{x}_{k-i}, \boldsymbol{\theta}_x)) + \boldsymbol{\eta}_k, \quad (32)$$

$$\boldsymbol{v}_k = \sum_{j=1}^{q} \boldsymbol{\Psi}_{k-j}(\boldsymbol{z}_{k-j} - h_{k-j}(\boldsymbol{x}_{k-j}, \boldsymbol{\theta}_z)) + \boldsymbol{\xi}_k. \quad (33)$$

Substituting (32)-(33) into (1)-(2), we can obtain the equivalent nonlinear system (8)-(9). This completes the proof.

Appendix B

***Lemma** 2:* Let $\boldsymbol{B} = [\boldsymbol{B}_{11}, \boldsymbol{B}_{12}]$, $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$, $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} \\ \boldsymbol{C}_{21} \end{bmatrix}$, and the Schur complement $\boldsymbol{D} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}(\boldsymbol{A}_{11})^{-1}\boldsymbol{A}_{12}$. Suppose that $\boldsymbol{A}$, $\boldsymbol{D}$ are invertible, then we have

$$\begin{aligned} \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{C} = \boldsymbol{B}_{11}\boldsymbol{A}_{11}^{-1}\boldsymbol{C}_{11} + (\boldsymbol{B}_{12} - \boldsymbol{B}_{11}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})\boldsymbol{D}^{-1} \\ \cdot (\boldsymbol{C}_{21} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{C}_{11}). \end{aligned} \quad (34)$$

Appendix C
Proof of Theorem 1

From (8) and (9), the joint PDF can be decomposed as

$$\begin{aligned} p_{k+1} &\triangleq p(\boldsymbol{X}^{k+1}, \boldsymbol{Z}^{k+1}|\boldsymbol{\theta}) \\ &= p(\boldsymbol{X}^k, \boldsymbol{Z}^k|\boldsymbol{\theta})p(\boldsymbol{x}_{k+1}|\boldsymbol{X}^k, \boldsymbol{Z}^k, \boldsymbol{\theta})p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{X}^k, \boldsymbol{Z}^k, \boldsymbol{\theta}) \\ &= p_k p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta})p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_{k-q+1}, \boldsymbol{\theta}) \end{aligned} \quad (35)$$

When $\boldsymbol{\Phi}_{k-j} = \boldsymbol{0}$, the transfer function in (35) is simplified to

$$p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \cdots, \boldsymbol{x}_{k-p}, \boldsymbol{\theta}) = p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta}) \quad (36)$$

When $\boldsymbol{\Psi}_{k-j} = \boldsymbol{0}$, the likelihood function in (35) is simplified to

$$p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \cdots, \boldsymbol{x}_{k-q+1}, \boldsymbol{\theta}) = p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{\theta}) \quad (37)$$

Partition $\boldsymbol{\chi}^k$ as $\boldsymbol{\chi}^k = [(\boldsymbol{X}^{k-l-1})', (\boldsymbol{x}_{k-l})', \cdots, (\boldsymbol{x}_k)', \boldsymbol{\theta}']'$ and $l = \max\{p, q-1\}$. Then $\boldsymbol{J}^k$ can be similarly partitioned as

$$\begin{aligned} \boldsymbol{J}^k &= \left[\begin{array}{c|ccc|c} \boldsymbol{A}_k^{1,1} & \boldsymbol{A}_k^{1,2} & \cdots & \boldsymbol{A}_k^{1,l+2} & \boldsymbol{A}_k^{1,\theta} \\ \hline \boldsymbol{A}_k^{2,1} & \boldsymbol{A}_k^{2,2} & \cdots & \boldsymbol{A}_k^{2,l+2} & \boldsymbol{A}_k^{2,\theta} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{A}_k^{l+2,1} & \boldsymbol{A}_k^{l+2,2} & \cdots & \boldsymbol{A}_k^{l+2,l+2} & \boldsymbol{A}_k^{l+2,\theta} \\ \hline \boldsymbol{A}_k^{\theta,1} & \boldsymbol{A}_k^{\theta,2} & \cdots & \boldsymbol{A}_k^{\theta,l+2} & \boldsymbol{A}_k^{\theta,\theta} \end{array}\right] \\ &= \left[\begin{array}{c|c|c} \boldsymbol{J}_k^{11} & \boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{13} \\ \hline \boldsymbol{J}_k^{21} & \boldsymbol{J}_k^{22} & \boldsymbol{J}_k^{23} \\ \hline \boldsymbol{J}_k^{31} & \boldsymbol{J}_k^{32} & \boldsymbol{J}_k^{33} \end{array}\right] \end{aligned} \quad (38)$$

where

$$\begin{aligned} \boldsymbol{A}_k^{1,1} &= E(-\Delta_{\boldsymbol{X}^{k-l-1}}^{\boldsymbol{X}^{k-l-1}} \ln p_k) \\ \boldsymbol{A}_k^{i,1} &= E(-\Delta_{\boldsymbol{x}_{k-l-2+i}}^{\boldsymbol{X}^{k-l-1}} \ln p_k) = (\boldsymbol{A}_k^{1,i})' \\ \boldsymbol{A}_k^{1,\theta} &= E(-\Delta_{\boldsymbol{X}^{k-l-1}}^{\boldsymbol{\theta}} \ln p_k) = (\boldsymbol{A}_k^{\theta,1})' \end{aligned}$$

$$
\begin{aligned}
\boldsymbol{A}_k^{i,j} &= E(-\Delta_{\boldsymbol{x}_{k-l-2+i}}^{\boldsymbol{x}_{k-l-2+j}} \ln p_k) \\
\boldsymbol{A}_k^{i,\theta} &= E(-\Delta_{\boldsymbol{x}_{k-l-2+i}}^{\boldsymbol{\theta}} \ln p_k) = (\boldsymbol{A}_k^{\theta,i})' \\
\boldsymbol{A}_k^{\theta,\theta} &= E(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p_k)
\end{aligned}
$$

for $i, j = 2, 3, \cdots, l+2$.

Define $\boldsymbol{L}_k$ as

$$
\begin{aligned}
\boldsymbol{L}_k &= \begin{bmatrix} \boldsymbol{J}_k^{22} & \boldsymbol{J}_k^{23} \\ \boldsymbol{J}_k^{32} & \boldsymbol{J}_k^{33} \end{bmatrix} - \begin{bmatrix} \boldsymbol{J}_k^{21} \\ \boldsymbol{J}_k^{31} \end{bmatrix} (\boldsymbol{J}_k^{11})^{-1} \begin{bmatrix} \boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{13} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{J}_k^{22} - \boldsymbol{J}_k^{21}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{23} - \boldsymbol{J}_k^{21}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{13} \\ \boldsymbol{J}_k^{32} - \boldsymbol{J}_k^{31}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{33} - \boldsymbol{J}_k^{31}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{13} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{L}_k^{x,x} & \boldsymbol{L}_k^{x,\theta} \\ \boldsymbol{L}_k^{\theta,x} & \boldsymbol{L}_k^{\theta,\theta} \end{bmatrix}
\end{aligned} \tag{39}
$$

Similarly, Partition $\boldsymbol{\chi}^{k+1}$ as $\boldsymbol{\chi}^{k+1} = [(\boldsymbol{X}^{k-l-1})', (\boldsymbol{x}_{k-l})', \cdots, (\boldsymbol{x}_k)', (\boldsymbol{x}_{k+1})', \boldsymbol{\theta}']'$. Then $\boldsymbol{J}^{k+1}$ can be partitioned as

$$
\boldsymbol{J}^{k+1} = \begin{bmatrix}
\boldsymbol{A}_k^{1,1} & \boldsymbol{A}_k^{1,2} & \cdots & \boldsymbol{A}_k^{1,l+2} & \boldsymbol{0} & \boldsymbol{A}_k^{1,\theta} \\
\boldsymbol{A}_k^{2,1} & \boldsymbol{A}_k^{2,2} + \boldsymbol{B}_k^{1,1} + \boldsymbol{C}_k^{1,1} & \cdots & \boldsymbol{A}_k^{2,l+2} + \boldsymbol{B}_k^{1,l+1} + \boldsymbol{C}_k^{1,l+1} & \boldsymbol{B}_k^{1,l+2} + \boldsymbol{C}_k^{1,l+2} & \boldsymbol{A}_k^{2,\theta} + \boldsymbol{E}_k^{1,\theta} + \boldsymbol{F}_k^{1,\theta} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
\boldsymbol{A}_k^{l+2,1} & \boldsymbol{A}_k^{l+2,2} + \boldsymbol{B}_k^{l+1,1} + \boldsymbol{C}_k^{l+1,1} & \cdots & \boldsymbol{A}_k^{l+2,l+2} + \boldsymbol{B}_k^{l+1,l+1} + \boldsymbol{C}_k^{l+1,l+1} & \boldsymbol{B}_k^{l+1,l+2} + \boldsymbol{C}_k^{l+1,l+2} & \boldsymbol{A}_k^{l+2,\theta} + \boldsymbol{E}_k^{l+1,\theta} + \boldsymbol{F}_k^{l+1,\theta} \\
\boldsymbol{0} & \boldsymbol{B}_k^{l+2,1} + \boldsymbol{C}_k^{l+2,1} & \cdots & \boldsymbol{B}_k^{l+2,l+1} + \boldsymbol{C}_k^{l+2,l+1} & \boldsymbol{B}_k^{l+2,l+2} + \boldsymbol{C}_k^{l+2,l+2} & \boldsymbol{E}_k^{l+2,\theta} + \boldsymbol{F}_k^{l+2,\theta} \\
\boldsymbol{A}_k^{\theta,1} & \boldsymbol{A}_k^{\theta,2} + \boldsymbol{E}_k^{\theta,1} + \boldsymbol{F}_k^{\theta,1} & \cdots & \boldsymbol{A}_k^{\theta,l+2} + \boldsymbol{E}_k^{\theta,l+1} + \boldsymbol{F}_k^{\theta,l+1} & \boldsymbol{E}_k^{\theta,l+2} + \boldsymbol{F}_k^{\theta,l+2} & \boldsymbol{A}_k^{\theta,\theta} + \boldsymbol{E}_k^{\theta,\theta} + \boldsymbol{F}_k^{\theta,\theta}
\end{bmatrix} \tag{40}
$$

$$
= \left[\begin{array}{cc|cc}
\boldsymbol{J}_k^{11} & \boldsymbol{J}_k^{12} & \boldsymbol{0} & \boldsymbol{J}_k^{13} \\
\boldsymbol{J}_k^{21} & \boldsymbol{J}_k^{22} + \boldsymbol{D}_k^{11} & \boldsymbol{D}_k^{12} & \boldsymbol{J}_k^{23} + \boldsymbol{D}_k^{13} \\
\hline
\boldsymbol{0} & \boldsymbol{D}_k^{21} & \boldsymbol{D}_k^{22} & \boldsymbol{D}_k^{23} \\
\boldsymbol{J}_k^{31} & \boldsymbol{J}_k^{32} + \boldsymbol{D}_k^{31} & \boldsymbol{D}_k^{32} & \boldsymbol{J}_k^{33} + \boldsymbol{D}_k^{33}
\end{array}\right] \tag{41}
$$

where $\boldsymbol{B}_k^{i,j}$, $\boldsymbol{C}_k^{i,j}$, $\boldsymbol{E}_k^{i,\theta}$, $\boldsymbol{F}_k^{i,\theta}$, $i, j = 1, 2, \cdots, l+2$, and $\boldsymbol{D}_k^{ij}$, $i, j = 1, 2, 3$ are as defined in Theorem 1.

Since $\boldsymbol{J}_{k+1}^{-1}$ equals the $(n+r)\times(n+r)$ right-lower block of $(\boldsymbol{J}^{k+1})^{-1}$, by using partitioned matrix inversion lemma [25] we have

$$
\begin{aligned}
\boldsymbol{J}_{k+1} &= \begin{bmatrix} \boldsymbol{D}_k^{22} & \boldsymbol{D}_k^{23} \\ \boldsymbol{D}_k^{32} & \boldsymbol{J}_k^{33} + \boldsymbol{D}_k^{33} \end{bmatrix} - \begin{bmatrix} \boldsymbol{0} & \boldsymbol{D}_k^{21} \\ \boldsymbol{J}_k^{31} & \boldsymbol{J}_k^{32} + \boldsymbol{D}_k^{31} \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} \boldsymbol{J}_k^{11} & \boldsymbol{J}_k^{12} \\ \boldsymbol{J}_k^{21} & \boldsymbol{J}_k^{22} + \boldsymbol{D}_k^{11} \end{bmatrix}^{-1} \begin{bmatrix} 0 & \boldsymbol{J}_k^{13} \\ \boldsymbol{J}_k^{13} & \boldsymbol{J}_k^{23} + \boldsymbol{D}_k^{13} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{J}_{k+1}^{x,x} & \boldsymbol{J}_{k+1}^{x,\theta} \\ \boldsymbol{J}_{k+1}^{\theta,x} & \boldsymbol{J}_{k+1}^{\theta,\theta} \end{bmatrix}
\end{aligned} \tag{42}
$$

Substituting (39) into (42) and using Lemma 2, thus the recursion can be obtain as in Eq. (11), and the blocks of $\boldsymbol{L}_k^{x,x}$, $\boldsymbol{L}_k^{x,\theta}$, $\boldsymbol{L}_k^{\theta,\theta}$ can be obtained as

$$
\boldsymbol{L}_k^{i,j} = \boldsymbol{A}_k^{i+1,j+1} - \boldsymbol{A}_k^{i+1,1}(\boldsymbol{A}_k^{1,1})^{-1}\boldsymbol{A}_k^{1,1+j} \tag{43}
$$

$$
\boldsymbol{L}_k^{i,\theta} = \boldsymbol{A}_k^{i+1,\theta} - \boldsymbol{A}_k^{i+1,1}(\boldsymbol{A}_k^{1,1})^{-1}\boldsymbol{A}_k^{1,\theta} \tag{44}
$$

$$
\boldsymbol{L}_k^{\theta,\theta} = \boldsymbol{A}_k^{\theta,\theta} - \boldsymbol{A}_k^{\theta,1}(\boldsymbol{A}_k^{1,1})^{-1}\boldsymbol{A}_k^{1,\theta} \tag{45}
$$

for $i, j = 1, 2, \cdots, l+1$

Using the definitions of $\boldsymbol{B}_k^{i,j}$, $\boldsymbol{C}_k^{i,j}$, $\boldsymbol{A}_k^{i,j}$ and (35)-(37), we have

$$
\begin{bmatrix} \boldsymbol{A}_k^{1,1} & \boldsymbol{A}_k^{1,1+j} \\ \boldsymbol{A}_k^{i+1,1} & \boldsymbol{A}_k^{i+1,j+1} \end{bmatrix} = \begin{bmatrix}
\boldsymbol{A}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{1,2} & \boldsymbol{A}_{k-1}^{1,j+2} \\
\boldsymbol{A}_{k-1}^{2,1} & \boldsymbol{A}_{k-1}^{2,2} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{2,j+2} + \boldsymbol{B}_{k-1}^{1,j+1} + \boldsymbol{C}_{k-1}^{1,j+1} \\
\boldsymbol{A}_{k-1}^{i+2,1} & \boldsymbol{A}_{k-1}^{i+2,2} + \boldsymbol{B}_{k-1}^{i+1,1} + \boldsymbol{C}_{k-1}^{i+1,1} & \boldsymbol{A}_{k-1}^{i+2,j+2} + \boldsymbol{B}_{k-1}^{i+1,j+1} + \boldsymbol{C}_{k-1}^{i+1,j+1}
\end{bmatrix} \tag{46}
$$

Note that $\boldsymbol{L}_k^{i,j}$ in (43) is the Schur complement of the matrix on the left hand side of (46), Then it can be obtained by the Schur complement of the matrix on the right hand side of (46) as

$$
\begin{aligned}
\boldsymbol{L}_k^{i,j} &= \boldsymbol{A}_{k-1}^{i+2,j+2} + \boldsymbol{B}_{k-1}^{i+1,j+1} + \boldsymbol{C}_{k-1}^{i+1,j+1} - \\
&\quad \begin{bmatrix} \boldsymbol{A}_{k-1}^{i+2,1} & \boldsymbol{A}_{k-1}^{i+2,2} + \boldsymbol{B}_{k-1}^{i+1,1} + \boldsymbol{C}_{k-1}^{i+1,1} \end{bmatrix} \\
&\quad \cdot \begin{bmatrix} \boldsymbol{A}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{1,2} \\ \boldsymbol{A}_{k-1}^{2,1} & \boldsymbol{A}_{k-1}^{2,2} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1} \end{bmatrix}^{-1} \\
&\quad \cdot \begin{bmatrix} \boldsymbol{A}_{k-1}^{1,j+2} \\ \boldsymbol{A}_{k-1}^{2,j+2} + \boldsymbol{B}_{k-1}^{1,j+1} + \boldsymbol{C}_{k-1}^{1,j+1} \end{bmatrix}
\end{aligned} \tag{47}
$$

Using Lemma 2 and the definition of $\boldsymbol{L}_k^{i,j}$, we can simplify (47) to obtain the following recursion

$$
\begin{aligned}
\boldsymbol{L}_k^{i,j} = \boldsymbol{L}_{k-1}^{i+1,j+1} + \boldsymbol{B}_{k-1}^{i+1,j+1} + \boldsymbol{C}_{k-1}^{i+1,j+1} - (\boldsymbol{L}_{k-1}^{i+1,1} + \boldsymbol{B}_{k-1}^{i+1,1} \\
+ \boldsymbol{C}_{k-1}^{i+1,1})(\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1})^{-1}(\boldsymbol{L}_{k-1}^{1,j+1} + \boldsymbol{B}_{k-1}^{1,j+1} \\
+ \boldsymbol{C}_{k-1}^{1,j+1})
\end{aligned}
$$

Similarly, we have

$$
\begin{bmatrix} \boldsymbol{A}_k^{1,1} & \boldsymbol{A}_k^{1,\theta} \\ \boldsymbol{A}_k^{i+1,1} & \boldsymbol{A}_k^{i+1,\theta} \end{bmatrix} = \begin{bmatrix}
\boldsymbol{A}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{1,2} & \boldsymbol{A}_{k-1}^{1,\theta} \\
\boldsymbol{A}_{k-1}^{2,1} & \boldsymbol{A}_{k-1}^{2,2} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{2,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} + \boldsymbol{F}_{k-1}^{1,\theta} \\
\boldsymbol{A}_{k-1}^{i+2,1} & \boldsymbol{A}_{k-1}^{i+2,2} + \boldsymbol{B}_{k-1}^{i+1,1} + \boldsymbol{C}_{k-1}^{i+1,1} & \boldsymbol{A}_{k-1}^{i+2,\theta} + \boldsymbol{E}_{k-1}^{i+1,\theta} + \boldsymbol{F}_{k-1}^{i+1,\theta}
\end{bmatrix} \tag{48}
$$

$$
\begin{bmatrix} \boldsymbol{A}_k^{1,1} & \boldsymbol{A}_k^{1,\theta} \\ \boldsymbol{A}_k^{\theta,1} & \boldsymbol{A}_k^{\theta,\theta} \end{bmatrix} = \begin{bmatrix}
\boldsymbol{A}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{1,2} & \boldsymbol{A}_{k-1}^{1,\theta} \\
\boldsymbol{A}_{k-1}^{2,1} & \boldsymbol{A}_{k-1}^{2,2} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1} & \boldsymbol{A}_{k-1}^{2,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} + \boldsymbol{F}_{k-1}^{1,\theta} \\
\boldsymbol{A}_{k-1}^{\theta,1} & \boldsymbol{A}_{k-1}^{\theta,2} + \boldsymbol{E}_{k-1}^{\theta,1} + \boldsymbol{F}_{k-1}^{\theta,1} & \boldsymbol{A}_{k-1}^{\theta,\theta} + \boldsymbol{E}_{k-1}^{\theta,\theta} + \boldsymbol{F}_{k-1}^{\theta,\theta}
\end{bmatrix} \tag{49}
$$

Then, using Schur complement and Lemma 2, we can obtain the recursion for $\boldsymbol{L}_k^{i,\theta}$ and $\boldsymbol{L}_k^{\theta,\theta}$ as

$$\begin{aligned}\boldsymbol{L}_k^{i,\theta} = & \boldsymbol{L}_{k-1}^{i+1,\theta} + \boldsymbol{E}_{k-1}^{i+1,\theta} + \boldsymbol{F}_{k-1}^{i+1,\theta} - [\boldsymbol{L}_{k-1}^{i+1,1} + \boldsymbol{B}_{k-1}^{i+1,1} \\ & + \boldsymbol{C}_{k-1}^{i+1,1}][\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} \\ & + \boldsymbol{F}_{k-1}^{1,\theta}] \end{aligned} \tag{50}$$

$$\begin{aligned}\boldsymbol{L}_k^{\theta,\theta} = & \boldsymbol{L}_{k-1}^{\theta,\theta} + \boldsymbol{E}_{k-1}^{\theta,\theta} + \boldsymbol{F}_{k-1}^{\theta,\theta} - [\boldsymbol{L}_{k-1}^{\theta,1} + \boldsymbol{E}_{k-1}^{\theta,1} + \boldsymbol{F}_{k-1}^{\theta,1}] \\ & \cdot [\boldsymbol{L}_{k-1}^{1,1} + \boldsymbol{B}_{k-1}^{1,1} + \boldsymbol{C}_{k-1}^{1,1}]^{-1}[\boldsymbol{L}_{k-1}^{1,\theta} + \boldsymbol{E}_{k-1}^{1,\theta} + \boldsymbol{F}_{k-1}^{1,\theta}] \end{aligned} \tag{51}$$

This completes the proof.

References

[1] K. Y. Lee, B. G. Lee, and S. Ann, "Adaptive filtering for speech enhancement in colored noise," *IEEE Signal Process. Lett.*, vol. 4, no. 10, pp. 277-279, 1997.

[2] A. Mahmoudi, M. Karimi, and H. Amindavar, "Parameter estimation of autoregressive signals in presence of colored AR(1) noise as a quadratic eigenvalue problem," *Signal Process.*, vol. 92, no. 4, pp. 1151-1156, 2012.

[3] W. R. Wu and D. C. Chang, "Maneuvering target tracking with colored noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 4, pp. 1311-1320, 1996.

[4] L. Mihaylova, D. Angelova, D. Bull, et al., "Localization of mobile nodes in wireless networks with correlated in time measurement noise," *IEEE Trans. Mob. Comput.*, vol. 10, no. 1, pp. 44-53, 2011.

[5] D. L. Zhang, Z. S. Duan, P. C. Wang and Y. H. Zhang "Spacecraft State Estimation with Multichannel Higher-order ARMA Colored Noises," in *Proc. 10th Int. Conf. Control, Autom. Inf. Sci.*, Xi'an, China, Oct. 2021, pp. 602-607.

[6] X. X. Wang and Q. Pan, "Nonlinear Gaussian filter with the colored measurement noise," in *Proc. 17th Int. Conf. Inf. Fusion*, Salamanca, Spain, Jul. 2014, pp. 1-7.

[7] X. X. Wang, Y. Liang, Q. Pan, C. H. Zhao, and F. Yang, "Nonlinear Gaussian smoothers with colored measurement noise," *IEEE Trans. Autom. Control*, vol.60, no. 3, pp. 870-876, 2015.

[8] H. V. Trees and K. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Hoboken, NJ: Wiley-IEEE, 2007.

[9] H. C. Lambert, "Cramér-Rao bounds for target tracking problems involving colored measurement noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 620-636, 2012.

[10] Z. G. Wang and X. J. Shen, "Posterior Cramŕ-Rao bounds for discrete-time nonlinear filtering with finitely correlated noises," in *Proc. 34th Chinese Control Conf.*, Hangzhou, China, Jul. 2015, pp. 4541-4546.

[11] Z. G. Wang, X. J. Shen and Y. M. Zhu, "Posterior Cramér-Rao bounds for nonlinear dynamic system with colored noises," *J. Syst. Sci. Complex.*, vol. 32, no. 6, pp. 1526-1543, 2019.

[12] L. Gao, G. Battistelli and L. Chisci, "Distributed joint sensor registration and multitarget tracking via sensor network," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1301-1317, 2020.

[13] C. Nemeth, P. Fearnhead, L. Mihaylova and D. Vorley, "Bearings-only tracking with particle filtering for joint parameter learning and state estimation," in *Proc. 15th Int. Conf. Inf. Fusion*, Singapore, Singapore, Jul. 2012, pp. 824-831.

[14] S. H. Li, X. X. Feng and F. Pan, "Joint state and parameter estimation for hidden wide-sense stationary ARMA processes under unknown noise," in *Proc. 20th Int. Conf. Inf. Fusion*, Xi'an, China, Jul. 2017, pp. 1228-1235.

[15] A. Nikoofard, U. J. F. Aarsnes, et al., "State and parameter estimation of a drift-flux model for underbalanced drilling operations," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 6, pp. 2000-2009, 2017.

[16] A. X. Yu, Y. Liu, J. B. Zhu and Z. Dong, "An improved dual unscented Kalman filter for state and parameter estimation," *Asian J. Control*, vol. 18, no. 4, pp. 1427-1440, 2016.

[17] R.Van Der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, Dept. Fac. OGI School Sci. Eng., Oregon Health Sci. Univ., Hillsboro, OR, USA, 2004.

[18] D. L. Huang and H. Leung, "An expectation maximization-based interacting multiple model approach for cooperative driving systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 206-228, 2005.

[19] L. F. Xu and X. R. Li, "Hybrid Cramér-Rao lower bound on tracking ground moving extended target," in *Proc. 12th Int. Conf. Inf. Fusion*, Seattle, WA, USA, Jul. 2009, pp. 304-307.

[20] C. Ren, J. Galy, and E. Chaumette, et al., "Recursive hybrid Cramér-Rao bound for discrete-time Markovian dynamic systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1543-4547, 2015.

[21] J. Galy, A. Renaux and E. Chaumette, et al., "Recursive hybrid CRB for Markovian systems with time-variant measurement parameters," in *Proc. IEEE Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process.*, Cancun, Mexico, Dec. 2015, pp. 473-476.

[22] X. Q. Li, Z. S. Duan, and U. D. Hanebeck, "Recursive joint Cramér-Rao lower bound for parametric systems with two-adjacent-states dependent measurements," *IET Signal Process.*, vol. 15, no. 4, pp. 221–237, 2021.

[23] X. Q. Li and Z. S. Duan, "Joint Cramér-Rao bound lower bound for nonlinear parametric systems with cross-correlated noises," *IEEE Signal Process. Lett.*, vol. 28, pp. 977–981, 2021.

[24] Y. S. Shmaliy, S. Y. Zhao, et al., "Optimal and unbiased filtering with colored process noise using state differencing," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 548–551, 2019.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd revised ed., New York: Cambridge Univ. Press, 2012.

[26] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333-1364, 2003.