Received: 7 December 2020 | Revised: 30 January 2021 | Accepted: 11 February 2021

DOI: 10.1049/sil2.12025

IET Signal Processing

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Recursive joint Cramér-Rao lower bound for parametric systems with two-adjacent-states dependent measurements

Xianqing Li[1] | Zhansheng Duan[1] | Uwe D. Hanebeck[2]

[1]Center for Information Engineering Science Research, School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an, China

[2]Intelligent Sensor-Actuator-Systems Laboratory, Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence

Zhansheng Duan, Center for Information Engineering Science Research, School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China.
Email: zsduan@mail.xjtu.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 61673313, 61673317; National Key Research and Development Plan, Grant/Award Number: 2017YFB1301101

Abstract

Joint Cramér-Rao lower bound (JCRLB) is very useful for the performance evaluation of joint state and parameter estimation (JSPE) of non-linear systems, in which the current measurement only depends on the current state. However, in reality, the non-linear systems with two-adjacent-states dependent (TASD) measurements, that is, the current measurement is dependent on the current state as well as the most recent previous state, are also common. First, the recursive JCRLB for the general form of such non-linear systems with unknown deterministic parameters is developed. Its relationships with the posterior CRLB for systems with TASD measurements and the hybrid CRLB for regular parametric systems are also provided. Then, the recursive JCRLBs for two special forms of parametric systems with TASD measurements, in which the measurement noises are autocorrelated or cross-correlated with the process noises at one time step apart, are presented, respectively. Illustrative examples in radar target tracking show the effectiveness of the JCRLB for the performance evaluation of parametric TASD systems.

1 | INTRODUCTION

To assess the performance of parameter estimators, we can resort to the lower bound and the upper bound. They assess estimation performance from two different aspects. The lower bound puts a limit to the best estimation performance, whereas the upper bound [1] puts a limit to the worst estimation performance. For lower bound, the most well-known one is the Cramér-Rao lower bound (CRLB). For upper bound, a lot of work has been investigated. For example, a multi-innovation stochastic gradient algorithm and the corresponding parameter estimation errors upper bound were proposed in [2] for time-invariant stochastic systems. For systems with time-varying parameters, a finite data window stochastic gradient identification algorithm was proposed in [3], in which the minimum parameter estimation error upper bound was obtained by choosing the data window length. A desirable property for estimators is the consistency [4] if their estimation errors converge to zero in a certain sense. For example, a hierarchical least squares identification algorithm was proposed in [5] by decomposing a dual-rate system model into several subsystems model for dual-rate linear systems with noises. This algorithm is consistent and has low computational complexity. For general dual-rate sampled-data systems, a computationally efficient algorithm was proposed in [6] using hierarchical identification principle, in which the parameter estimates can converge to their true values. When the measurements of the systems are scarce, a gradient-based parameter identification method [7] was proposed and its convergence property was also analysed.

Recursive state estimators for non-linear systems have been widely used in signal processing and control. The non-linear filters can be designed by using the Bayesian estimation approach. However, in general, only suboptimal filters can be obtained for non-linear systems due to complicated non-linear multiple integrals. Among the suboptimal filters, the extended Kalman filter (EKF) [8, 9] is probably the earliest and most well-known one. Its main idea is to approximate the non-linear systems as linear systems by first order Taylor series expansion and ignore the higher order terms. So far, a large number of advanced non-linear filters have been developed, for example, the unscented filter (UF) [10, 11], quadrature Kalman filter (QKF) [12, 13], cubature Kalman filter (CKF) [14–16], and so forth. All these non-linear filters aim to approximate the first-two moments used in the linear minimum mean-squared error (LMMSE) estimation. The UF approximates the first-two moments based on the unscented transformation of well-designed sigma-points. The QKF calculates the first-two moments using the Gauss–Hermite quadrature integration rules, but the number of integral points increases exponentially as the dimension of state increases. The CKF transforms the integral problem into a summation problem using a third-degree spherical-radial cubature rule, by which it can approximate the first-two moments needed in the LMMSE. In the recent years, another well-known non-linear filter is the particle filter (PF) [17, 18]. The PF uses sequential Monte Carlo method to generate a large number of samples to characterize the posterior probability density. For the state estimation of special systems, for example, bilinear systems, the standard Kalman filter is not applicable. A state filtering method was developed in [19] by using delta operator to minimize the covariance matrix of state estimation errors for single-input–single-output and multiple-input–multiple-output bilinear systems. For linear systems with multistate delays, extended state Kalman filter will suffer from heavy computational burden. To overcome this, a highly computationally efficient state filter was proposed for such systems in [20] following the same idea of [19].

All the above estimators mainly deal with the regular non-linear systems, in which the current measurement only depends on the current state. However, in many real applications, the non-linear systems are more complex [19, 20]. For example, non-linear systems with two-adjacent-states dependent (TASD) measurements are also common. Actually, non-linear systems with autocorrelated measurement noises or with cross-correlated process and measurement noises at one time step apart [21] are two typical cases of such systems. To handle these two cases, one way is to reformulate the measurement equation as a new measurement equation which is dependent on the current state as well as the most recent previous state. In many real applications, the whiteness assumption for noise as in the regular non-linear systems cannot be completely satisfied, where the measurement noise is either autocorrelated or cross-correlated with the process noise. In the radar tracking system, the measurement noises are always autocorrelated, which usually influences the maneuvering target tracking performance [22, 23]. In the global navigation satellite system, suffering from signal outages and multi-path error, the measurement noise becomes autocorrelated [24–26]. Parameters estimation is an important subject in modern signal processing where the measurement noises involved is often autocorrelated [27, 28]. The systems with dependent process and measurement noises is also very common in practice [29]. Particle filters were proposed in [30] for these cases in which the process and measurement noises are dependent. Some practical applications of models with dependent noises, for example, sensor fusion and econometrics, have been shown in [31]. To assess the performance of non-linear filters, the posterior CRLB (PCRLB) is proposed. In [32], a recursive approach to compute the PCRLB was developed for filtering of regular non-linear dynamic systems. To evaluate the performance for filtering of non-linear systems with TASD measurements, a recursive CRLB was developed in [33] for non-linear systems with the coloured noise, in which the first-order coloured measurement noise case is a typical TASD system, and a recursive PCRLB was proposed in [34] for the non-linear systems with cross-correlated process and measurement noises at one time step apart. In [35], a conditional PCRLB for TASD systems was proposed, which depends on the measurement history up to now. Because it relies on a specific implementation, it can be used to guide online sensor selection.

Actually, most discrete-time dynamic systems may incorporate some parameters which can be random or non-random. Joint state and parameter estimation (JSPE) has got a growing interest in various applications, such as target tracking [36], signal processing [37], sensor registration [38]. The main approaches for JSPE include joint filter [39, 40], dual filter [41, 42] expectation maximization method [43, 44] and so forth. Joint filter schemes usually augment the parameters into a state vector and use various kinds of filters, for example, EKF, UKF, and so forth, to simultaneously estimate both. Dual filter schemes use two separate filters for state and parameter estimation. It decouples parameters and states and ignores the relevant information between them. Expectation maximization schemes solve the joint estimation problem by two iterative steps, in which the E-step assumes that the parameters are known and estimates the states, and the M-step uses the estimated states to identify the system parameters. To assess the performance for JSPE, hybrid CRLB (HCRLB) is introduced in [45–49]. A unified framework for HCRLB applied to the joint estimation of random and non-random parameters was summarized in [50]. Then the asymptotic tightness of the HCRLB was analysed in [51]. For the purpose of computational efficiency, the recursive form of HCRLB was proposed for the ground moving extended target tracking problem [52], and the more general form of it was obtained for the discrete-time systems involving unknown time invariant deterministic parameters in [53]. Furthermore, the HCRLB was extended to the non-linear systems with time-variant measurement parameters in [54].

Performance evaluation of JSPE for parametric TASD systems has not been studied yet. This article studies this problem and aims at proposing a recursive performance bound for the JSPE of parametric systems with TASD measurements. The contributions of this paper are as follows.

1. We develop a recursive joint CRLB (JCRLB) for the general form of parametric TASD systems.
2. Its relationships with the PCRLB for the TASD systems and the HCRLB for parametric regular systems are discussed. It is found that both the PCRLB for TASD systems and the HCRLB for regular parametric systems are special cases of the JCRLB for parametric TASD systems.

3. We present specific JCRLBs for two special cases of parametric TASD systems, including non-linear systems with autocorrelated measurement noises or with cross-correlated process and measurement noises at one time step apart.

The rest of this article is organized as follows. Section 2 formulates the JCRLB problem and presents the non-linear TASD systems with unknown deterministic parameters. Then, a recursive JCRLB to evaluate the performance of JSPE for these systems is proposed in Section 3. Section 4 develops the JCRLBs for two specific non-linear parametric systems. In Section 5, some illustrative examples in radar target tracking are provided to verify the effectiveness of the proposed JCRLB. Section 6 concludes this paper.

2 | PROBLEM FORMULATION

As we know, the classical CRLB is suitable for the estimation of non-random parameters, whereas the Bayesian CRLB is applicable to random parameters. In joint parameter estimation problem, one wishes to estimate an unknown joint parameter vector $\boldsymbol{\chi}=(\boldsymbol{x}',\boldsymbol{\theta}')'$ given the measurement $\boldsymbol{z}$, where $\boldsymbol{x}$ is a random parameter vector and $\boldsymbol{\theta}$ is a non-random parameter vector. The joint Fisher information matrix (JFIM) for $\boldsymbol{\chi}$ is then

$$\boldsymbol{J}=E[-\Delta_{\chi}^{\chi}\ln p(\boldsymbol{x},\boldsymbol{z}\,|\,\boldsymbol{\theta})]$$

where $p(\boldsymbol{x},\ \boldsymbol{z}|\boldsymbol{\theta})$ is the joint conditional probability density function (PDF), Δ denotes the second-order derivative operator, that is, $\Delta_a^b=\nabla_a\nabla_b'$, and ∇ denotes the gradient operator.

Let $\hat{\chi}(\boldsymbol{z})$ be an unbiased estimator of $\boldsymbol{\chi}$. The JCRLB on the mean squared error (MSE) is defined as the inverse of the JFIM [49]:

$$E\{[\hat{\chi}(z)-\chi][\hat{\chi}(z)-\chi]'\}\geq \boldsymbol{J}^{-1}$$

Consider the following general form of non-linear parametric systems with TASD measurements.

$$\boldsymbol{x}_{k+1}=f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x,\boldsymbol{w}_k) \quad (1)$$

$$\boldsymbol{z}_k=h_k(\boldsymbol{x}_k,\boldsymbol{x}_{k-1},\boldsymbol{\theta}_z,\boldsymbol{v}_k) \quad (2)$$

where $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_z$ are unknown as the deterministic parameter vectors, the process noise $\langle\boldsymbol{w}_k\rangle$ and the measurement noise $\langle\boldsymbol{v}_k\rangle$ are mutually independent white noises with PDFs $p(\boldsymbol{w}_k|\boldsymbol{\beta})$ and $p(\boldsymbol{v}_k|\boldsymbol{\gamma})$, respectively, which depend on unknown deterministic parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and the initial state $\boldsymbol{x}_0$ is independent of the process and measurement noises with PDF $p(\boldsymbol{x}_0|\alpha)$, which depends on the unknown deterministic parameter α. The joint estimand (quantity to be estimated) consists of the system state $\boldsymbol{x}_k\in\mathbb{R}^n$ and non-random parameter vector $\boldsymbol{\theta}=[\boldsymbol{\alpha}',\boldsymbol{\beta}',\boldsymbol{\gamma}',\boldsymbol{\theta}_x',\boldsymbol{\theta}_z']'\in\mathbb{R}^m$, composed of all unknown deterministic parameters.

Unlike in regular parametric systems, the measurement $\boldsymbol{z}_k$ in TASD systems depends on both the current state $\boldsymbol{x}_k$ and the most recent previous state $\boldsymbol{x}_{k-1}$. The main goal of this paper is to obtain a recursive JCRLB for filtering of non-linear parametric systems with TASD measurements.

3 | JCRLB FOR PARAMETRIC TASD SYSTEMS

3.1 | JCRLB

From Equations (1) and (2), the joint conditional probability distribution of $\boldsymbol{x}^{k+1}=[\boldsymbol{x}_0',\ldots,\boldsymbol{x}_{k+1}']'$ and $\boldsymbol{z}^{k+1}=[\boldsymbol{z}_1',\ldots,\boldsymbol{z}_{k+1}']'$ given the parameter $\boldsymbol{\theta}$ at arbitrary time $k+1$ is

$$\begin{aligned}p_{k+1}&\triangleq p\left(\boldsymbol{x}^{k+1},\boldsymbol{z}^{k+1}\middle|\boldsymbol{\theta}\right)\\&=p\left(\boldsymbol{x}^k,\boldsymbol{z}^k\middle|\boldsymbol{\theta}\right)\cdot p\left(\boldsymbol{x}_{k+1}|\boldsymbol{x}^k,\boldsymbol{z}^k,\boldsymbol{\theta}\right)\\&\quad\cdot p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}^k,\boldsymbol{z}^k,\boldsymbol{\theta})\\&=p_k\cdot p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta})\cdot p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})\end{aligned} \quad (3)$$

Definition 1 Define $\hat{\chi}^k$ and $\hat{\chi}_k$ as estimates of joint estimands $\chi^k=[(\boldsymbol{x}^k)',\ \boldsymbol{\theta}']'$ and $\chi_k=[(\boldsymbol{x}_k)',\ \boldsymbol{\theta}']'$, respectively.

Definition 2 The MSE of $\hat{\chi}^k$ at time k is defined as

$$\mathrm{MSE}(\hat{\chi}^k)\triangleq E[\tilde{\chi}^k(\tilde{\chi}^k)']=\int\tilde{\chi}^k(\tilde{\chi}^k)'p_k d_{\boldsymbol{x}^k}d_{\boldsymbol{z}^k}$$

and the MSE of $\hat{\chi}_k$ at time k is defined as

$$\mathrm{MSE}(\hat{\chi}_k)\triangleq E[\tilde{\chi}_k(\tilde{\chi}_k)']=\int\tilde{\chi}_k(\tilde{\chi}_k)'p(\boldsymbol{x}_k,\boldsymbol{z}^k|\boldsymbol{\theta})d_{\boldsymbol{x}_k}d_{\boldsymbol{z}^k}$$

where $\tilde{\chi}^k=\hat{\chi}^k-\chi^k$ and $\tilde{\chi}_k=\hat{\chi}_k-\chi_k$ are the associated estimation errors, $p(\boldsymbol{x}_k,\ \boldsymbol{z}^k|\boldsymbol{\theta})$ is the joint conditional probability distribution of $\boldsymbol{x}_k$ and $\boldsymbol{z}^k$ given $\boldsymbol{\theta}$

Definition 3 Define the JFIM $\boldsymbol{J}^k$ about the joint estimand sequence χ^k as

$$\boldsymbol{J}^k\triangleq E[-\Delta_{\chi^k}^{\chi^k}\ln p_k]\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}=-\int(\Delta_{\chi^k}^{\chi^k}\ln p_k)p_k d_{\boldsymbol{x}^k}d_{\boldsymbol{z}^k}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

where $\boldsymbol{\theta}_0$ is the true value of the unknown non-random parameter $\boldsymbol{\theta}$.

Lemma 1 *The MSE of a joint unbiased estimate* $\hat{\chi}^k$ *is bounded from below by the inverse of the* $\boldsymbol{J}^k$ [49] as

$$\mathrm{MSE}(\hat{\chi}^k) \triangleq E[\tilde{\chi}^k(\tilde{\chi}^k)'] \geq (\boldsymbol{J}^k)^{-1}$$

Definition 4 Define $\boldsymbol{J}_k^{-1}$ as the $(n+m) \times (n+m)$ right-lower block of $(\boldsymbol{J}^k)^{-1}$ and $\boldsymbol{J}_k$ as the JFIM about $\boldsymbol{x}_k$ and $\boldsymbol{\theta}$,

Lemma 2 *The MSE of a joint unbiased estimate* $\hat{\chi}_k$ *is bounded from below by the inverse of* $\boldsymbol{J}_k$ [51, 53] as

$$\mathrm{MSE}(\hat{\chi}_k) \triangleq E[\tilde{\chi}_k(\tilde{\chi}_k)'] \geq \boldsymbol{J}_k^{-1}$$

We aim to obtain a recursive form to calculate $\boldsymbol{J}_k^{-1}$ without manipulating the large matrix $\boldsymbol{J}^k$.

Theorem 1 *Given the Fisher information matrix* $\boldsymbol{J}_k$ about $\boldsymbol{x}_k$ and $\boldsymbol{\theta}$ as

$$\boldsymbol{J}_k = \begin{bmatrix} \boldsymbol{J}_k^{x,x} & \boldsymbol{J}_k^{x,\theta} \\ \boldsymbol{J}_k^{\theta,x} & \boldsymbol{J}_k^{\theta,\theta} \end{bmatrix} \tag{4}$$

the recursion from $\boldsymbol{J}_k$ to $\boldsymbol{J}_{k+1}$ can be obtained as

$$\begin{cases} \boldsymbol{J}_{k+1}^{x,x} = \boldsymbol{D}_k^{22} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1}\boldsymbol{D}_k^{12} \\ \boldsymbol{J}_{k+1}^{x,\theta} = \boldsymbol{D}_k^{23} - \boldsymbol{D}_k^{21}[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1}(\boldsymbol{D}_k^{13} + \boldsymbol{J}_k^{x,\theta}) = (\boldsymbol{J}_{k+1}^{\theta,x})' \\ \boldsymbol{J}_{k+1}^{\theta,\theta} = \boldsymbol{D}_k^{33} + \boldsymbol{J}_k^{\theta,\theta} - (\boldsymbol{D}_k^{31} + \boldsymbol{J}_k^{\theta,x})[\boldsymbol{D}_k^{11} + \boldsymbol{J}_k^{x,x}]^{-1}(\boldsymbol{D}_k^{13} + \boldsymbol{J}_k^{x,\theta}) \end{cases} \tag{5}$$

where

$$\begin{cases} \boldsymbol{D}_k^{11} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \\ \boldsymbol{D}_k^{12} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \\ \boldsymbol{D}_k^{13} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \\ \boldsymbol{D}_k^{22} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \\ \boldsymbol{D}_k^{23} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{\theta}}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \\ \boldsymbol{D}_k^{33} = E_{p_{k+1}}\{-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}}[\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\} \end{cases} \tag{6}$$

and $\boldsymbol{D}_k^{12} = (\boldsymbol{D}_k^{21})'$, $\boldsymbol{D}_k^{13} = (\boldsymbol{D}_k^{31})'$, $\boldsymbol{D}_k^{23} = (\boldsymbol{D}_k^{32})'$. The initial information $\boldsymbol{J}_0$ can be obtained from the prior probability distribution $p(\boldsymbol{x}_0|\alpha)$ as

$$\begin{cases} \boldsymbol{J}_0^{x,x} = E[-\Delta_{\boldsymbol{x}_0}^{\boldsymbol{x}_0}\ln p(\boldsymbol{x}_0|\boldsymbol{\alpha})] \\ \boldsymbol{J}_0^{x,\theta} = E[-\Delta_{\boldsymbol{x}_0}^{\boldsymbol{\theta}}\ln p(\boldsymbol{x}_0|\boldsymbol{\alpha})] = (\boldsymbol{J}_0^{\theta,x})' \\ \boldsymbol{J}_0^{\theta,\theta} = E[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}}\ln p(\boldsymbol{x}_0|\boldsymbol{\alpha})] \end{cases} \tag{7}$$

Proof See Appendix 1.

Remark 1 If $p(\boldsymbol{x}_0|\alpha)$ is not available, then $\boldsymbol{J}_0^{x,x} = \boldsymbol{0}$, $\boldsymbol{J}_0^{x,\theta} = (\boldsymbol{J}_0^{\theta,x})' = \boldsymbol{0}$, and $\boldsymbol{J}_0^{\theta,\theta} = \boldsymbol{0}$. Next, we discuss two commonly encountered cases in practice, including the additive Gaussian noise case and linear additive Gaussian noise case.

Corollary 1 *Suppose that the discrete-time non-linear system (Equations (1) and (2)) is driven by additive Gaussian noises as*

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x) + \boldsymbol{w}_k \tag{8}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z) + \boldsymbol{v}_k \tag{9}$$

where the noises $\langle \boldsymbol{w}_k \rangle$ and $\langle \boldsymbol{v}_k \rangle$ are mutually independent white Gaussian processes with zero mean and invertible covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively. Then Equation (6) will be simplified to

$$\begin{cases} \boldsymbol{D}_k^{11} = E\{[\nabla_{\boldsymbol{x}_k} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{x}_k} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]'\} \\ \quad + E\{[\nabla_{\boldsymbol{x}_k} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \times \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{x}_k} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{12} = -E\{[\nabla_{\boldsymbol{x}_k} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]\}\boldsymbol{Q}_k^{-1} \\ \quad + E\{[\nabla_{\boldsymbol{x}_k} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \cdot \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{x}_{k+1}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{13} = E\{[\nabla_{\boldsymbol{x}_k} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{\theta}} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]'\} \\ \quad + E\{[\nabla_{\boldsymbol{x}_k} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \times \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{\theta}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{22} = \boldsymbol{Q}_k^{-1} + E\{[\nabla_{\boldsymbol{x}_{k+1}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \cdot \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{x}_{k+1}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{23} = -\boldsymbol{Q}_k^{-1}E\{[\nabla_{\boldsymbol{\theta}} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]'\} \\ \quad + E\{[\nabla_{\boldsymbol{x}_{k+1}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \cdot \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{\theta}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{33} = E\{[\nabla_{\boldsymbol{\theta}} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{\theta}} f_k'(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]'\} \\ \quad + E\{[\nabla_{\boldsymbol{\theta}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)] \\ \quad \times \boldsymbol{R}_{k+1}^{-1}[\nabla_{\boldsymbol{\theta}} h_{k+1}'(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]'\} \end{cases} \tag{10}$$

where $\boldsymbol{\theta} = [\boldsymbol{\theta}_x', \theta_z']'$

Proof See Appendix 2.

Corollary 2 *Suppose that the discrete-time system (Equations (8) and (9)) is further reduced to a linear system with additive Gaussian noises as*

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{F}_k^{\theta} \boldsymbol{\theta}_x + \boldsymbol{w}_k \tag{11}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{C}_{k-1} \boldsymbol{x}_{k-1} + \boldsymbol{H}_k^{\theta} \boldsymbol{\theta}_z + \boldsymbol{v}_k \tag{12}$$

where the noises $\langle \boldsymbol{w}_k \rangle$ and $\langle \boldsymbol{v}_k \rangle$ are mutually independent white Gaussian processes with zero mean and invertible covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively. Then Equation (10) will be simplified to

$$\begin{cases} \boldsymbol{D}_k^{11} = \boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} \boldsymbol{F}_k + \boldsymbol{C}_k' \boldsymbol{R}_{k+1}^{-1} \boldsymbol{C}_k \\ \boldsymbol{D}_k^{12} = -\boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} + \boldsymbol{C}_k' \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} = (\boldsymbol{D}_k^{21})' \\ \boldsymbol{D}_k^{13} = \boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] + \boldsymbol{C}_k' \boldsymbol{R}_{k+1}^{-1} [\boldsymbol{0}, \boldsymbol{H}_{k+1}^{\theta}] = (\boldsymbol{D}_k^{31})' \\ \boldsymbol{D}_k^{22} = \boldsymbol{Q}_k^{-1} + \boldsymbol{H}_{k+1}' \boldsymbol{R}_{k+1}^{-1} \boldsymbol{H}_{k+1} \\ \boldsymbol{D}_k^{23} = -\boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] + \boldsymbol{H}_{k+1}' \boldsymbol{R}_{k+1}^{-1} [\boldsymbol{0}, \boldsymbol{H}_{k+1}^{\theta}] = (\boldsymbol{D}_k^{32})' \\ \boldsymbol{D}_k^{33} = [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}]' \boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] + [\boldsymbol{0}, \boldsymbol{H}_{k+1}^{\theta}]' \boldsymbol{R}_{k+1}^{-1} [\boldsymbol{0}, \boldsymbol{H}_{k+1}^{\theta}] \end{cases} \tag{13}$$

where $\boldsymbol{0}$'s are zero matrices with appropriate dimensions.

Proof See Appendix 3.

3.2 | Relationship with the PCRLB for TASD systems

Corollary 3 *Suppose that the discrete-time system (Equations (1) and (2)) is reduced to the following non-parametric non-linear TASD system*

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{w}_k) \tag{14}$$

$$\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \boldsymbol{v}_k) \tag{15}$$

where the process noise $\langle \boldsymbol{w}_k \rangle$ and the measurement noise $\langle \boldsymbol{v}_k \rangle$ are mutually independent white noise sequences. The initial state $\boldsymbol{x}_0$ is independent of the process and measurement noises. Then the recursion for $\boldsymbol{J}_k$ in Theorem 1 will be reduced to

$$\boldsymbol{J}_{k+1} = \boldsymbol{D}_k^{22} - \boldsymbol{D}_k^{21} [\boldsymbol{D}_k^{11} + \boldsymbol{J}_k]^{-1} \boldsymbol{D}_k^{12} \tag{16}$$

where

$$\begin{cases} \boldsymbol{D}_k^{11} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k)] \} \\ \boldsymbol{D}_k^{12} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k)] \} \\ \quad = (\boldsymbol{D}_k^{21})' \\ \boldsymbol{D}_k^{22} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k)] \} \end{cases} \tag{17}$$

$$p_{k+1} \triangleq p(\boldsymbol{x}^{k+1}, \boldsymbol{z}^{k+1}) \tag{18}$$

This is exactly the FIsM for non-parametric TASD systems proposed in [33, 34]. Obviously, Equation (16) is a special case of the JCRLB in Theorem 1.

3.3 | Relationship with the HCRLB for regular parametric systems

In regular parametric systems, the current measurement $\boldsymbol{z}_k$ only depends on the current state $\boldsymbol{x}_k$, that is, $\boldsymbol{z}_k = h_k(\boldsymbol{x}_k, \boldsymbol{v}_k, \theta_z)$. This is a special case of Equation (2). For such systems, the likelihood PDF $p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta})$ in Equation (3) will be reduced to $p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{\theta})$. As a result, in Equation (6) we have

$$\begin{aligned} & E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} [\ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta})] \} \\ & \quad = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} [\ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \} = \boldsymbol{0} \end{aligned}$$

Similarly,

$$E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}} [\ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta})] \} = \boldsymbol{0}$$
$$E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} [\ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta})] \} = \boldsymbol{0}$$

Then, Equation (6) will be simplified to

$$\begin{cases} \boldsymbol{D}_k^{11} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta})] \} \\ \boldsymbol{D}_k^{12} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta})] \} \\ \boldsymbol{D}_k^{13} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta})] \} \\ \boldsymbol{D}_k^{22} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \} \\ \boldsymbol{D}_k^{23} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{\theta}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \} \\ \boldsymbol{D}_k^{33} = E_{p_{k+1}} \{ -\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} [\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1} | \boldsymbol{x}_{k+1}, \boldsymbol{\theta})] \} \end{cases} \tag{19}$$

The JFIM $\boldsymbol{J}_k$ in Equation (5) with Equation (6) replaced by Equation (19) is exactly the FIM for regular parametric systems proposed in [53]. Therefore, Equation (19) is a special case of the JCRLB in Theorem 1.

4 | JCRLB FOR TWO SPECIAL TYPES OF NON-LINEAR SYSTEMS

The non-linear parametric systems, in which the measurement noises are either autocorrelated or cross-correlated with the process noises at one time step apart, are two special types of TASD systems. Specific JCRLBs for these two types of systems will be presented next.

4.1 | JCRLB for non-linear systems with autocorrelated measurement noises

Consider the following non-linear system:

$$\boldsymbol{x}_{k+1} = f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x) + \boldsymbol{w}_k \tag{20}$$

$$\boldsymbol{y}_k = l_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) + \boldsymbol{e}_k \tag{21}$$

where $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_z$ are unknown deterministic parameter vectors, the process noise $\langle \boldsymbol{w}_k \rangle$ is white, the measurement noise $\langle \boldsymbol{e}_k \rangle$ is autocorrelated satisfying the following first-order autoregressive(AR) model [23].

$$\boldsymbol{e}_k = \boldsymbol{\Psi}_{k-1}\boldsymbol{e}_{k-1} + \boldsymbol{\xi}_{k-1} \tag{22}$$

where the driven noise $\langle \boldsymbol{\xi}_{k-1} \rangle$ is white, $\langle \boldsymbol{w}_k \rangle$ and $\langle \boldsymbol{\xi}_{k-1} \rangle$ are mutually independent, and both independent of the initial state $\boldsymbol{x}_0$ as well.

The difference of two adjacent measurements yields

$$\boldsymbol{z}_k = \boldsymbol{y}_k - \boldsymbol{\Psi}_{k-1}\boldsymbol{y}_{k-1} \tag{23}$$

By treating $\boldsymbol{z}_k$ as a pseudo measurement, we have the following TASD measurement equation

$$\begin{aligned}\boldsymbol{z}_k &= l_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) - \boldsymbol{\Psi}_{k-1} l_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z) + \boldsymbol{\xi}_{k-1} \\ &= h_k(\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z) + \boldsymbol{v}_k\end{aligned} \tag{24}$$

where

$$h_k(\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z) = l_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z) - \boldsymbol{\Psi}_{k-1} l_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_z)$$
$$\boldsymbol{v}_k = \boldsymbol{\xi}_{k-1}$$

Obviously the pseudo measurement noise $\langle \boldsymbol{v}_k \rangle$ is white, and independent of $\langle \boldsymbol{w}_k \rangle$ and $\boldsymbol{x}_0$.

Applying Theorem 1 to the equivalent TASD system (Equations (20) and (24)), we can obtain the recursive JCRLB for non-linear systems with autocorrelated measurement noises.

Next we discuss some specific forms of the JFIM for non-linear systems with autocorrelated measurement noises when the noises are Gaussian and the systems are linear.

Theorem 2 *For the non-linear system (Equations (20) and (22)), if $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$ and $\boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k)$, then Equation (6) in Theorem 1 will be simplified to*

$$\begin{cases} \boldsymbol{D}_k^{11} = E\{[\nabla_{\boldsymbol{x}_k} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{x}_k} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]'\} \\ \quad + E\{[\nabla_{\boldsymbol{x}_k} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]\boldsymbol{R}_k^{-1}[\nabla_{\boldsymbol{x}_k} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]'\} \\ \boldsymbol{D}_k^{12} = -E\{[\nabla_{\boldsymbol{x}_k} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]\}\boldsymbol{Q}_k^{-1} \\ \quad - E\{[\nabla_{\boldsymbol{x}_k} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]\boldsymbol{R}_k^{-1}[\nabla_{\boldsymbol{x}_{k+1}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{13} = E\{[\nabla_{\boldsymbol{x}_k} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{\theta}} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]'\} \\ \quad - E\{[\nabla_{\boldsymbol{x}_k} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]\boldsymbol{R}_k^{-1}\nabla_{\boldsymbol{\theta}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z) \\ \quad - \nabla_{\boldsymbol{\theta}} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k{}'\} \\ \boldsymbol{D}_k^{22} = \boldsymbol{Q}_k^{-1} + E[\nabla_{\boldsymbol{x}_{k+1}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z)] \\ \quad \cdot \boldsymbol{R}_k^{-1}[\nabla_{\boldsymbol{x}_{k+1}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z)]'\} \\ \boldsymbol{D}_k^{23} = -\boldsymbol{Q}_k^{-1}E\{[\nabla_{\boldsymbol{\theta}} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]'\} + E[\nabla_{\boldsymbol{x}_{k+1}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z)] \\ \quad \cdot \boldsymbol{R}_k^{-1}[\nabla_{\boldsymbol{\theta}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z) - \nabla_{\boldsymbol{\theta}} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]'\} \\ \boldsymbol{D}_k^{33} = E\{[\nabla_{\boldsymbol{\theta}} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]\boldsymbol{Q}_k^{-1}[\nabla_{\boldsymbol{\theta}} f'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)]'\} \\ \quad + E[\nabla_{\boldsymbol{\theta}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z) - \nabla_{\boldsymbol{\theta}} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]\boldsymbol{R}_k^{-1} \\ \quad \cdot [\nabla_{\boldsymbol{\theta}} l'_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{\theta}_z) - \nabla_{\boldsymbol{\theta}} l'_k(\boldsymbol{x}_k, \boldsymbol{\theta}_z)\boldsymbol{\Psi}'_k]'\} \end{cases} \tag{25}$$

where $\boldsymbol{\theta} = [\boldsymbol{\theta}'_x, \theta'_z]'$

Proof See Appendix 4.

Remark 2 Suppose that the non-linear system (Equations (20) and (22)) does not depend on any non-random parameters, that is, $\boldsymbol{\theta} \in \varnothing$. By applying Corollary 3, we can obtain the PCRLB for non-linear systems with autocorrelated measurement noises, which is the same as Corollary 3.5 in [33].

Corollary 4 *If the non-linear system (Equations (20) and (22)) is further reduced to a linear system*

$$\begin{cases} x_{k+1} = F_k x_k + F_k^\theta \theta_x + w_k \\ y_k = L_k x_k + L_k^\theta \theta_z + e_k \\ e_k = \Psi_{k-1} e_{k-1} + \xi_{k-1} \end{cases} \tag{26}$$

then Equation (25) in Theorem 2 will be further simplified to

$$\begin{cases} D_k^{11} = F_k' Q_k^{-1} F_k + L_k' \Psi_k' R_k^{-1} \Psi_k L_k \\ D_k^{12} = -F_k' Q_k^{-1} - L_k' \Psi_k' R_k^{-1} L_{k+1} = (D_k^{21})' \\ D_k^{13} = F_k' Q_k^{-1} [F_k^\theta, 0] - L_k' \Psi_k' R_k^{-1} ([0, L_{k+1}^\theta] - \Psi_k [0, L_k^\theta]) \\ D_k^{22} = Q_k^{-1} + L_{k+1}' R_k^{-1} L_{k+1} \\ D_k^{23} = -Q_k^{-1} [F_k^\theta, 0] + L_{k+1}' R_k^{-1} ([0, L_{k+1}^\theta] - \Psi_k [0, L_k^\theta]) \\ D_k^{33} = [F_k^\theta, 0]' Q_k^{-1} [F_k^\theta, 0] + ([0, L_{k+1}^\theta] - \Psi_k [0, L_k^\theta])' R_k^{-1} \\ \qquad \cdot ([0, L_{k+1}^\theta] - \Psi_k [0, L_k^\theta]) \end{cases} \tag{27}$$

where $\mathbf{0}$'s are zero matrices with appropriate dimensions.

4.2 | JCRLB for non-linear systems with noises cross-correlated at one time step apart

Consider the following non-linear system:

$$x_{k+1} = f_k(x_k, \theta_x) + w_k \tag{28}$$

$$y_k = l_k(x_k, \theta_z) + e_k \tag{29}$$

where θ_x and θ_z are unknown deterministic parameter vectors, $\langle w_k \rangle$ and $\langle e_k \rangle$ are zero-mean white noises cross-correlated at one time step apart [30] and independent of the initial state x_0 as well, satisfying

$$\text{cov}(w_k) = Q_k$$
$$\text{cov}(e_k) = E_k$$
$$E[w_k e_j'] = U_k \delta_{k,j-1}$$

where $\delta_{k,j-1}$ is the Kronecker delta function.

As in [35], the following TASD measurement equation can be obtained

$$\begin{aligned} z_k &= l_k(x_k, \theta_z) + e_k + G_k(x_k - f_{k-1}(x_{k-1}, \theta_x) - w_{k-1}) \\ &= l_k(x_k, \theta_z) + G_k(x_k - f_{k-1}(x_{k-1}, \theta_x)) + e_k - G_k w_{k-1} \\ &= h_k(x_k, x_{k-1}, \theta) + v_k \end{aligned} \tag{30}$$

where

$$h_k(x_k, x_{k-1}, \theta) = l_k(x_k, \theta_z) + G_k(x_k - f_{k-1}(x_{k-1}, \theta_x))$$
$$v_k = e_k - G_k w_{k-1}$$
$$G_k = U_{k-1}' Q_{k-1}^{-1}$$
$$\theta = [\theta_x', \theta_z']'$$

and the pseudo measurement noise $\langle v_k \rangle$ and $\langle w_{k-1} \rangle$ are mutually independent white noises, and the mean and covariance of $\langle v_k \rangle$ are zero and $R_k = E_k - U_{k-1}' Q_{k-1}^{-1} U_{k-1}$, respectively.

Applying Theorem 1 to the equivalent TASD system (Equations (28) and (30)), we can obtain the recursive JCRLB for non-linear systems with cross-correlated process and measurement noises at one time step apart.

Next we discuss some specific forms of the JFIM for non-linear systems with noises cross-correlated at one time step apart when the noises are Gaussian and the systems are linear.

Theorem 3 *For the non-linear system (Equations (28) and (29)), if $w_k \sim \mathcal{N}(0, Q_k)$ and $e_k \sim \mathcal{N}(0, E_k)$, then Equation (6) in Theorem 1 will be simplified to*

$$\begin{cases} D_k^{11} = E\{[\nabla_{x_k} f_k'(x_k, \theta_x)] Q_k^{-1} [\nabla_{x_k} f_k'(x_k, \theta_x)]'\} \\ \qquad + E\{[\nabla_{x_k} f_k'(x_k, \theta_x) G_{k+1}'] R_{k+1}^{-1} [\nabla_{x_k} f_k'(x_k, \theta_x) G_{k+1}']'\} \\ D_k^{12} = -E\{[\nabla_{x_k} f_k'(x_k, \theta_x)]\} Q_k^{-1} - E[\nabla_{x_k} f_k'(x_k, \theta_x) G_{k+1}'] \\ \qquad \cdot R_{k+1}^{-1} [\nabla_{x_{k+1}} l_{k+1}'(x_{k+1}, \theta_z) + G_{k+1}']'\} = (D_k^{21})' \\ D_k^{13} = E\{[\nabla_{x_k} f_k'(x_k, \theta_x)] Q_k^{-1} [\nabla_\theta f_k'(x_k, \theta_x)]'\} \\ \qquad - E\{[\nabla_{x_k} f_k'(x_k, \theta_x) G_{k+1}'] R_{k+1}^{-1} \\ \qquad \cdot [\nabla_\theta l_{k+1}'(x_{k+1}, \theta_z) - \nabla_\theta f_k'(x_k, \theta_x) G_{k+1}']'\} = (D_k^{31})' \\ D_k^{22} = Q_k^{-1} + E\{[\nabla_{x_{k+1}} l_{k+1}'(x_{k+1}, \theta_z) + G_{k+1}'] \\ \qquad \cdot R_{k+1}^{-1} [\nabla_{x_{k+1}} l_{k+1}'(x_{k+1}, \theta_z) + G_{k+1}']'\} \\ D_k^{23} = -Q_k^{-1} E\{[\nabla_\theta f_k'(x_k, \theta_x)]'\} \\ \qquad + E\{[\nabla_{x_{k+1}} l_{k+1}'(x_{k+1}, \theta_z) + G_{k+1}'] R_{k+1}^{-1} \\ \qquad \cdot [\nabla_\theta l_{k+1}'(x_{k+1}, \theta_z) - \nabla_\theta f_k'(x_k, \theta_x) G_{k+1}']'\} = (D_k^{32})' \\ D_k^{33} = E\{[\nabla_\theta f_k'(x_k, \theta_x)] Q_k^{-1} [\nabla_\theta f_k'(x_k, \theta_x)]'\} \\ \qquad + E\{[\nabla_\theta l_{k+1}'(x_{k+1}, \theta_z) - \nabla_\theta f_k'(x_k, \theta_x) G_{k+1}'] \\ \qquad \cdot R_{k+1}^{-1} [\nabla_\theta l_{k+1}'(x_{k+1}, \theta_z) - \nabla_\theta f_k'(x_k, \theta_x) G_{k+1}']'\} \end{cases} \tag{31}$$

Proof See Appendix 5.

Remark 3 Suppose that the non-linear system (Equations (28) and (29)) does not depend on any non-random parameters, that is, $\boldsymbol{\theta} \in \emptyset$. By applying Corollary 3, we can obtain the PCRLB for non-linear systems with noises cross-correlated at one time step apart, which is the same as Theorem 4 in [34].

Corollary 5 *If the non-linear system (Equations 28 and 29) is further reduced to a linear system,*

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{F}_k^{\theta} \boldsymbol{\theta}_x + \boldsymbol{w}_k \tag{32}$$

$$\boldsymbol{y}_k = \boldsymbol{L}_k \boldsymbol{x}_k + \boldsymbol{L}_k^{\theta} \boldsymbol{\theta}_z + \boldsymbol{e}_k \tag{33}$$

then Equation (31) in Theorem 3 will be further simplified to

$$\begin{cases} \boldsymbol{D}_k^{11} = \boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} \boldsymbol{F}_k + \boldsymbol{F}_k' \boldsymbol{G}_{k+1}' \boldsymbol{R}_{k+1}^{-1} \boldsymbol{G}_{k+1} \boldsymbol{F}_k \\ \boldsymbol{D}_k^{12} = -\boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} - \boldsymbol{F}_k' \boldsymbol{G}_{k+1}' \boldsymbol{R}_{k+1}^{-1} (\boldsymbol{L}_{k+1}' + \boldsymbol{G}_{k+1}')' = (\boldsymbol{D}_k^{21})' \\ \boldsymbol{D}_k^{13} = \boldsymbol{F}_k' \boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] - \boldsymbol{F}_k' \boldsymbol{G}_{k+1}' \boldsymbol{R}_{k+1}^{-1} \\ \qquad \cdot ([\boldsymbol{0}, \boldsymbol{L}_{k+1}^{\theta}] - \boldsymbol{G}_{k+1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}]) = (\boldsymbol{D}_k^{31})' \\ \boldsymbol{D}_k^{22} = \boldsymbol{Q}_k^{-1} + (\boldsymbol{L}_{k+1}' + \boldsymbol{G}_{k+1}') \boldsymbol{R}_{k+1}^{-1} (\boldsymbol{L}_{k+1}' + \boldsymbol{G}_{k+1}')' \\ \boldsymbol{D}_k^{23} = -\boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] + (\boldsymbol{L}_{k+1}' + \boldsymbol{G}_{k+1}') \boldsymbol{R}_{k+1}^{-1} \\ \qquad \cdot ([\boldsymbol{0}, \boldsymbol{L}_{k+1}^{\theta}] - \boldsymbol{G}_{k+1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}]) = (\boldsymbol{D}_k^{32})' \\ \boldsymbol{D}_k^{33} = [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}]' \boldsymbol{Q}_k^{-1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}] + ([\boldsymbol{0}, \boldsymbol{L}_{k+1}^{\theta}] - \boldsymbol{G}_{k+1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}])' \boldsymbol{R}_{k+1}^{-1} \\ \qquad \cdot ([\boldsymbol{0}, \boldsymbol{L}_{k+1}^{\theta}] - \boldsymbol{G}_{k+1} [\boldsymbol{F}_k^{\theta}, \boldsymbol{0}]) \end{cases} \tag{34}$$

where $\boldsymbol{0}$'s are zero matrices with appropriate dimensions.

5 | ILLUSTRATIVE EXAMPLES

In this section, numerical examples in the radar target tracking are provided to demonstrate the effectiveness of the proposed recursive JCRLB for JSPE of non-linear parametric systems with TASD measurements. However, to the best of our knowledge, there does not exist any unbiased joint estimator for this problem. So we design the following numerical examples to show how hard it is to jointly estimate the target motion state and radar measurement biases.

Consider a single target with coordinated turn motion in the two-dimensional plane [14, 16, 29, 33].

$$\boldsymbol{X}_{k+1} = \begin{bmatrix} 1 & \dfrac{\sin \omega T}{\omega} & 0 & \dfrac{\cos \omega T - 1}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \dfrac{1 - \cos \omega T}{\omega} & 1 & \dfrac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix} \boldsymbol{X}_k + \boldsymbol{w}_k \tag{35}$$

where $\boldsymbol{X}_k = [x_k, \dot{x}_k, y_k, \dot{y}_k]'$ is the state vector, $T = 1$ s is the sampling interval, $\omega = 2^{\circ} \mathrm{s}^{-1}$ is the turning rate, $\boldsymbol{w}_k$ is a zero-mean white Gaussian process noise with covariance $\boldsymbol{Q}_k = \mathrm{cov}(\boldsymbol{w}_k) = \mathrm{diag}(q\boldsymbol{M}, q\boldsymbol{M})$, $q = 0.01$ m^2 s^{-3} and

$$\boldsymbol{M} = \begin{bmatrix} \dfrac{T^3}{3} & \dfrac{T^2}{2} \\ \dfrac{T^2}{2} & T \end{bmatrix} \tag{36}$$

A two-dimensional radar located at the origin is used to track the motion of the target

$$\boldsymbol{z}_{k+1} = \begin{bmatrix} r_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} \sqrt{x_{k+1}^2 + y_{k+1}^2} \\ \tan^{-1}\left(\dfrac{y_{k+1}}{x_{k+1}}\right) \end{bmatrix} + \begin{bmatrix} \Delta_r \\ \Delta_\theta \end{bmatrix} + \boldsymbol{e}_{k+1} \tag{37}$$

where $\boldsymbol{z}_{k+1}$ is the radar measurement vector consisting of the range measurement r_{k+1} and the bearing measurement θ_{k+1}, Δ_r and Δ_θ are range and bearing measurement biases, which are unknown non-random parameters with ground truth 20 m and 5 mrad, respectively, and $\langle \boldsymbol{e}_{k+1} \rangle$ is the measurement noise.

5.1 | Example 1: autocorrelated measurement noises

In this example, we assume that the measurement noise $\langle \boldsymbol{e}_{k+1} \rangle$ in Equation (37) is autocorrelated and governed by

$$\boldsymbol{e}_{k+1} = 0.5 \boldsymbol{I} \boldsymbol{e}_k + \boldsymbol{\xi}_k \tag{38}$$

where $\boldsymbol{I}$ is a 2×2 identity matrix, $\langle \boldsymbol{\xi}_k \rangle$ is a zero-mean white Gaussian driven noise with covariance $\boldsymbol{R}_k = \mathrm{diag}(\sigma_r^2(\boldsymbol{\xi}), \sigma_\theta^2(\boldsymbol{\xi}))$. Moreover, $\langle \boldsymbol{w}_k \rangle$ and $\langle \boldsymbol{\xi}_k \rangle$ are mutually independent, and both of them are independent of the initial state $\boldsymbol{X}_0$. The distribution of $\boldsymbol{X}_0$ is $\boldsymbol{X}_0 \sim \mathcal{N}(\bar{\boldsymbol{X}}_0, \boldsymbol{P}_0)$ with

$$\bar{\boldsymbol{X}}_0 = [2000\ \mathrm{m}, 10\ \mathrm{ms}^{-1}, 4000\ \mathrm{m}, 10\ \mathrm{ms}^{-1}]'$$
$$\boldsymbol{P}_0 = \mathrm{diag}(500^2\ \mathrm{m}^2, 200^2\ \mathrm{m}^2\mathrm{s}^{-2}, 500^2\ \mathrm{m}^2, 15^2\ \mathrm{m}^2\mathrm{s}^{-2})$$

We consider three different settings for $\sigma_r(\boldsymbol{\xi})$ and $\sigma_\theta(\boldsymbol{\xi})$ in Table 1. Experiments over 1000 Monte Carlo runs are used to show how the proposed JCRLB can be used to evaluate the

TABLE 1 Settings of $\sigma_r(\boldsymbol{\xi})$ and $\sigma_\theta(\boldsymbol{\xi})$ in Example 1

Case 1	$\sigma_r(\boldsymbol{\xi}) = 10$ m	$\sigma_\theta(\boldsymbol{\xi}) = 3$ mrad
Case 2	$\sigma_r(\boldsymbol{\xi}) = 15$ m	$\sigma_\theta(\boldsymbol{\xi}) = 5$ mrad
Case 3	$\sigma_r(\boldsymbol{\xi}) = 20$ m	$\sigma_\theta(\boldsymbol{\xi}) = 7$ mrad

estimation performance in this radar target tracking example with autocorrelated measurement noises.

According to Equation (24), the equivalent TASD measurement equation for this example is

$$z^*_{k+1} = \begin{bmatrix} \sqrt{x^2_{k+1} + y^2_{k+1}} - 0.5\sqrt{x^2_k + y^2_k} \\ \tan^{-1}\left(\frac{y_{k+1}}{x_{k+1}}\right) - 0.5\tan^{-1}\left(\frac{y_k}{x_k}\right) \end{bmatrix} + 0.5\begin{bmatrix} \Delta_r \\ \Delta_\theta \end{bmatrix} + \boldsymbol{\xi}_k \tag{39}$$

Figures 1–4 show that the JCRLBs of position, velocity, range bias Δ_r and bearing bias Δ_θ increase as the covariance of the pseudo measurement noise $\langle \boldsymbol{\xi}_k \rangle$ increases. From Equation (39), we know that the uncertainty of the pseudo measurement $\boldsymbol{z}^*_{k+1}$ is only dependent on the uncertainty of the white noise $\langle \boldsymbol{\xi}_k \rangle$ owing to its independence of state $\boldsymbol{x}_{k+1}$ and $\boldsymbol{x}_k$ and that the uncertainty of $\boldsymbol{x}_{k+1}$ and $\boldsymbol{x}_k$ is fixed. So if the covariance $\boldsymbol{R}_k$ increases, then the uncertainty of the pseudo measurement noise increases and the pseudo measurement $\boldsymbol{z}^*_{k+1}$ is more uncertain. Thus the JCRLB of the non-linear system gets larger when $\boldsymbol{R}_k$ gets larger. That is, the larger the pseudo measurement noise level is, the more difficult it is to jointly estimate the target motion state and radar measurement biases.

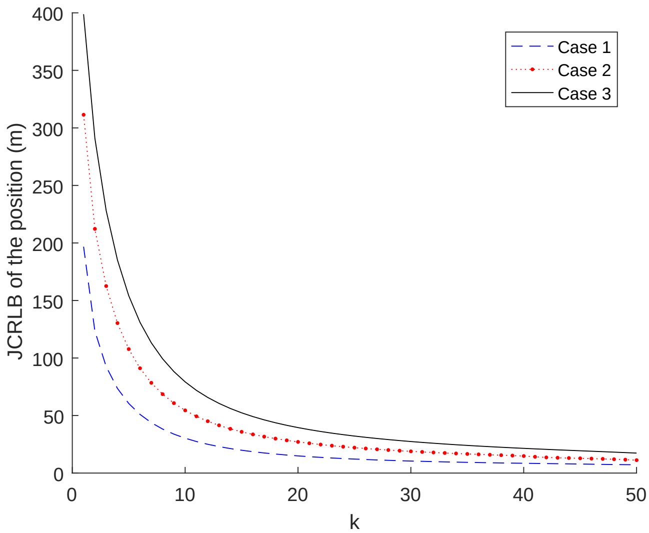

FIGURE 1 Joint Cramér-Rao lower bound for position under different pseudo measurement noise level in Example 1

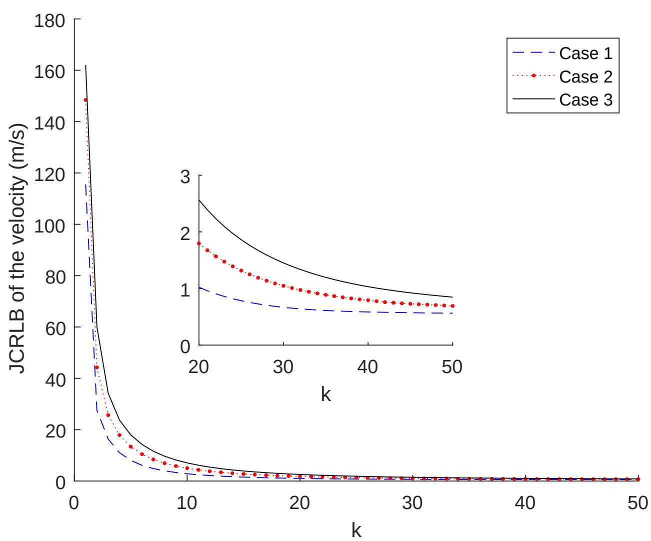

FIGURE 2 Joint Cramér-Rao lower bound for velocity under different pseudo measurement noise level in Example 1

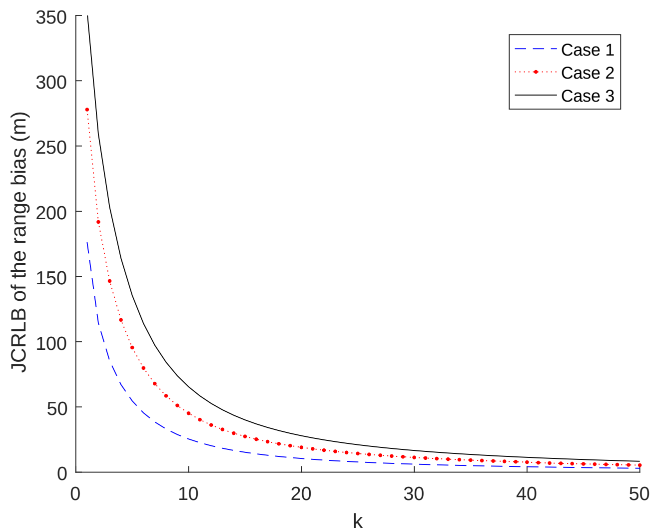

FIGURE 3 Joint Cramér-Rao lower bound for range bias under different pseudo measurement noise level in Example 1

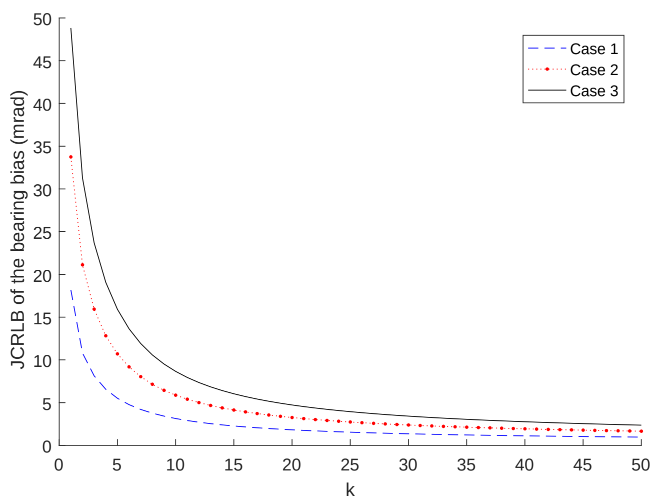

FIGURE 4 Joint Cramér-Rao lower bound for bearing bias under different pseudo measurement noise level in Example 1

5.2 | Example 2: cross-correlated process and measurement noises at one time step apart

In this example, we assume that the measurement noise $\boldsymbol{e}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{E}_k)$ in Equation (37) is cross-correlated with the process noise $\langle \boldsymbol{w}_k \rangle$ in Equation (35) at one time step apart. The cross-correlation between $\boldsymbol{w}_k$ and $\boldsymbol{e}_{k+1}$ is $E[\boldsymbol{w}_k \boldsymbol{e}'_{k+1}] = \boldsymbol{U}_k = \begin{bmatrix} 0.5 & 0.5 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0 \end{bmatrix}'$. The measurement noise covariance is $\boldsymbol{E}_k = \mathrm{diag}(\sigma_r^2(\boldsymbol{e}), \sigma_\theta^2(\boldsymbol{e}))$. Also, the distribution of $\boldsymbol{X}_0$ is $\boldsymbol{X}_0 \sim \mathcal{N}(\bar{\boldsymbol{X}}_0, \boldsymbol{P}_0)$ with

$$\bar{\boldsymbol{X}}_0 = [2000 \text{ m}, 10 \text{ ms}^{-1}, 4000 \text{ m}, 10 \text{ ms}^{-1}]'$$
$$\boldsymbol{P}_0 = \mathrm{diag}(500^2 \text{ m}^2, 100^2 \text{ m}^2\text{s}^{-2}, 500^2 \text{ m}^2, 15^2 \text{ m}^2\text{s}^{-2})$$

We consider three different settings for $\sigma_r(\boldsymbol{e})$ and $\sigma_\theta(\boldsymbol{e})$ in Table 2. Experiments over 1000 Monte Carlo runs are used to show how the proposed JCRLB can be used to evaluate the estimation performance in this radar target tracking example with cross-correlated process and measurement noises at one time step apart.

According to Equation (30), we can obtain the covariance of the equivalent TASD measurement noise for this example as

$$\boldsymbol{R}_{k+1} = \boldsymbol{E}_{k+1} - \boldsymbol{U}'_k \boldsymbol{Q}_k^{-1} \boldsymbol{U}_k \tag{40}$$

TABLE 2 Settings of $\sigma_r(\boldsymbol{e})$ and $\sigma_\theta(\boldsymbol{e})$ in Example 2

Case 1	$\sigma_r(\boldsymbol{e}) = 15$ m	$\sigma_\theta(\boldsymbol{e}) = 3$ mrad
Case 2	$\sigma_r(\boldsymbol{e}) = 20$ m	$\sigma_\theta(\boldsymbol{e}) = 5$ mrad
Case 3	$\sigma_r(\boldsymbol{e}) = 25$ m	$\sigma_\theta(\boldsymbol{e}) = 7$ mrad

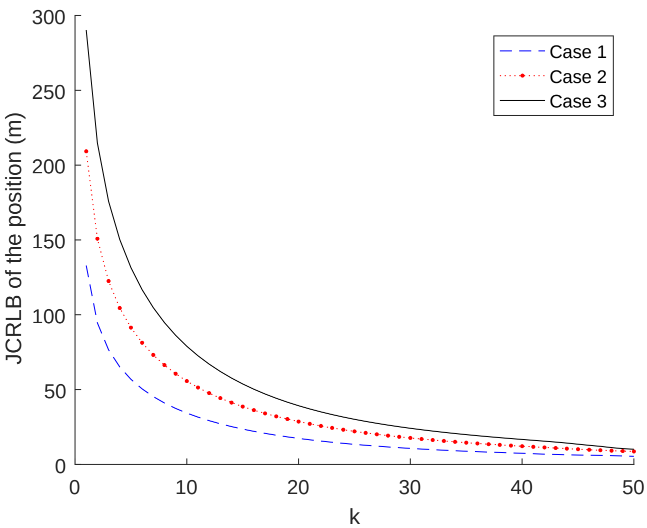

FIGURE 5 Joint Cramér-Rao lower bound for position under different measurement noise level in Example 2

Figures 5–8 show that the JCRLBs of position, velocity, range bias Δ_r and bearing bias Δ_θ increase as the covariance of measurement noise $\langle \boldsymbol{e}_k \rangle$ increases. From Equations (30) and (40), we know that if the covariance $\boldsymbol{E}_{k+1}$ of the measurement noise $\langle \boldsymbol{e}_k \rangle$ increases while $\boldsymbol{U}_k$ and $\boldsymbol{Q}_k$ are fixed, then the covariance $\boldsymbol{R}_{k+1}$ of the equivalent TASD measurement noise $\langle \boldsymbol{v}_k \rangle$ increases and the equivalent TASD measurement is more uncertain. Thus the JCRLB of the non-linear system gets larger when $\boldsymbol{E}_{k+1}$ gets larger. In other words, the larger the measurement noise level gets, the more difficult it is to estimate the target motion state and radar measurement biases.

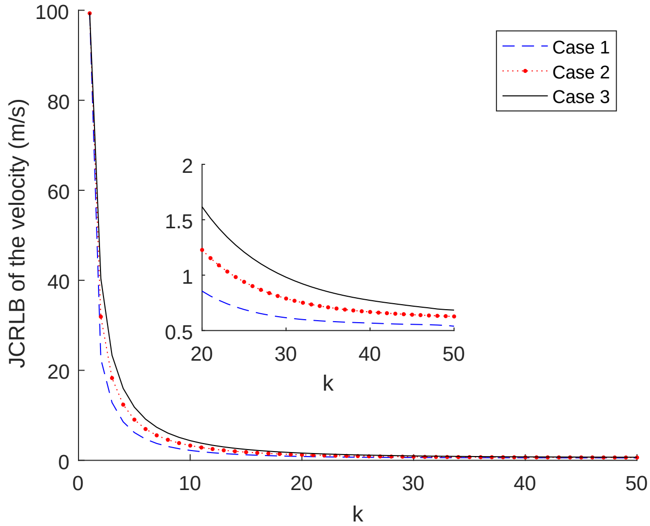

FIGURE 6 Joint Cramér-Rao lower bound for velocity under different measurement noise level in Example 2

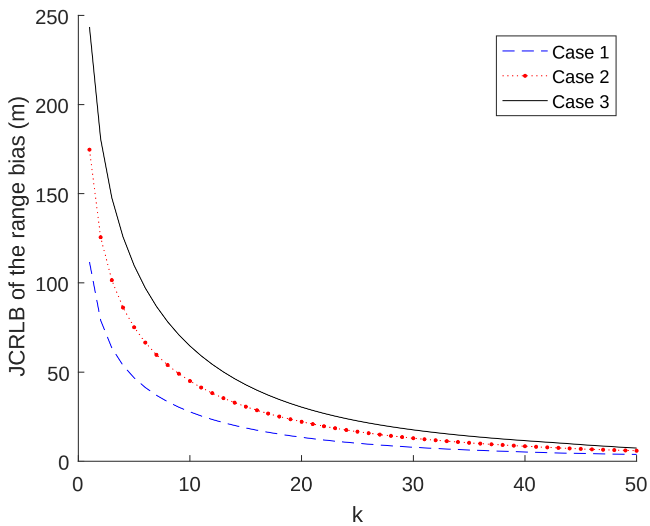

FIGURE 7 Joint Cramér-Rao lower bound for range bias under different measurement noise level in Example 2

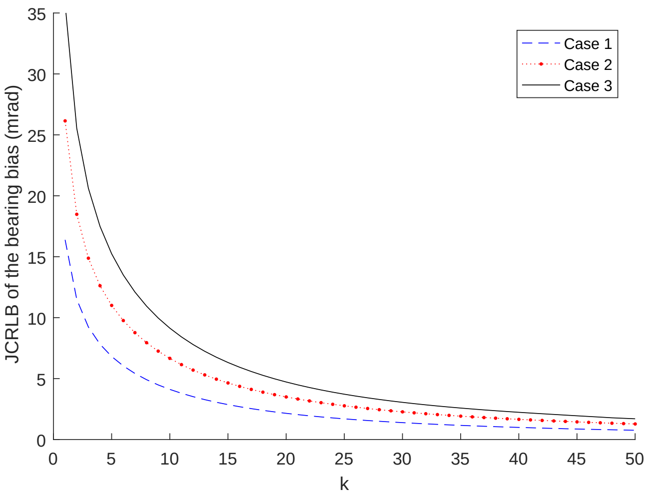

FIGURE 8 Joint Cramér-Rao lower bound for bearing bias under different measurement noise level in Example 2

6 | CONCLUSIONS

We have developed a recursive JCRLB for JSPE of dynamic nonlinear parametric systems with TASD measurements, that is, the current measurement depends on both the current state and the most recent previous state. Its connections with the PCRLB for the systems with TASD measurements and the HCRLB for regular parametric systems, in which the current measurement only depends on the current state, have been studied as well. It is found that both the PCRLB for TASD systems and the HCRLB for regular parametric systems are special cases of the JCRLB for the TASD parametric systems. The main difference between them is due to the use of different likelihood functions. Meanwhile, specific and explicit forms of JCRLBs for commonly encountered Gaussian noise case and linear Gaussian case have been presented. The recursive JCRLBs for two typical forms of the systems with TASD measurements, that is, systems with autocorrelated measurement noises or cross-correlated process and measurement noises at one time step apart, have also been investigated. Moreover, simplified and explicit forms of the recursive JCRLBs have been obtained for these two types of systems for the Gaussian noise case and the linear Gaussian case as well.

ACKNOWLEDGEMENTS

This work is supported in part by National Key Research and Development Plan 2017YFB1301101, and National Natural Science Foundation of China through grants 61673317 and 61673313.

ORCID

Xianqing Li https://orcid.org/0000-0001-9107-8767
Zhansheng Duan https://orcid.org/0000-0001-7366-5984

REFERENCES

1. Ding, F., Chen, T.W.: Performance bounds of forgetting factor least-squares algorithms for time-varying systems with finite measurement data. IEEE Trans. Circ. Syst. I-Reg. Pap. 52(3), 555–566 (2005)
2. Ding, F., Chen, T.W.: Performance analysis of multi-innovation gradient type identification methods. Automatica. 43(1), 1–14 (2007)
3. Ding, F., Xu, L., Zhu, Q.M.: Performance analysis of the generalised projection identification for time-varying systems. IET Control Theory & Appl. 10(18), 2506–2514 (2016)
4. Ding, F., Liu, X.P., Liu, G.J.: Multi-innovation least-squares identification for system modeling. IEEE Trans. Syst. Man. Cybern. B Cybern. 40(3), 767–778 (2010)
5. Ding, J., et al.: Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data. IEEE Trans. Automat. Contr. 56(11), 2677–2683 (2011)
6. Liu, Y.J., Ding, F., Shi, Y.: An efficient hierarchical identification method for general dual-rate sampled-data systems. Automatica. 50(3), 962–970 (2014)
7. Ding, F., Liu, G.J., Liu, X.P.: Parameter estimation with scarce measurements. Automatica. 47(8), 1646–1655 (2011)
8. Aidala, V.J.: Kalman filter behavior in bearings-Only tracking applications. IEEE Trans. Aero. Electron. Syst. 15(1), 29–39 (1979)
9. Duan, Z.S., Li, X.Y.: A new nonlinear state estimator using the fusion of multiple extended Kalman filters. In: Proceedings of 18th International Conference on Information Fusion, pp. 90–97, Washington, DC (2015)
10. Julier, S.J., Uhlmann, J.K., Durrant-Whyte, H.F.: A new method for the nonlinear transformation of means and covariances in filters and estimators. IEEE Trans. Automat. Contr. 45(3), 472–482 (2000)
11. Julier, S.J., Uhlmann, J.K.: Unscented filtering and nonlinear estimation. Proc. IEEE. 92(3), 401–422 (2004)
12. Ito, K., Xiong, K.: Gaussian filters for nonlinear filtering problems. IEEE Trans. Automat. Contr. 45(5), 910–927 (2000)
13. Arasaratnam, I., Haykin, S., Elliott, R.J.: Discrete-time nonlinear filtering algorithms using Gauss–Hermite quadrature. Proc. IEEE. 95(5), 953–977 (2007)
14. Arasaratnam, I., Haykin, S.: Cubature Kalman filters. IEEE Trans. Automat. Contr. 54(6), 1254–1269 (2009)
15. Chen, J.G., et al.: Extended target probability hypothesis density filter based on cubature Kalman filter. IET Radar, Sonar Navig. 9(3), 324–332 (2015)
16. Jia, B., Xin, M., Cheng, Y.: High-degree cubature Kalman filter. Automatica. 49(2), 510–518 (2013)
17. Arulampalam, M.S., et al.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. 50(3), 174–188 (2002)
18. Chen, J., Ma, L.: Particle filtering with correlated measurement and process noise at the same time. IET Radar, Sonar Navig. 5(7), 726–730 (2011)
19. Zhang, X., Ding, F., Yang, E.F.: State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors. Int. J. Adapt. Contr. Signal Process. 33(7), 1157–1173 (2019)
20. Zhang, X., et al.: Highly computationally efficient state filter based on the delta operator. Int. J. Adapt. Contr. Signal Process. 33(6), 875–889 (2019)
21. Bar-Shalom, Y., Li, X.R., Kirubarajan, T.: Estimation with applications to tracking and navigation. John Wiley & Sons, Inc, New York (2001)
22. Wen, W.R., Chang, D.C.: Maneuvering target tracking with colored noise. IEEE Trans. Aero. Electron. Syst. 32(4), 1311–1320 (1996)
23. Wang, X.X., et al.: Nonlinear Gaussian smoothers with colored measurement noise. IEEE Trans. Automat. Contr. 60(3), 870–876 (2015)
24. Petovello, M.G., et al.: Consideration of time-correlated errors in a Kalman filter applicable to GNSS. J. Geodes. 83, 51–56 (2009)
25. Wang, K.D., Li, Y., Rizos, C.: Practical approaches to Kalman filtering with time-correlated measurement errors. IEEE Trans. Aero. Electron. Syst. 48(2), 1669–1681 (2012)
26. Cui, B.B., et al.: Robust cubature Kalman filter for GNSS/INS with missing observations and colored measurement noise. ISA Trans. 72, 138–146 (2018)
27. Mahmoudi, A., Karimi, M., Amindavar, H.: Parameter estimation of autoregressive signals in presence of colored AR(1) noise as a quadratic eigenvalue problem. Signal Process. 92(4), 1151–1156 (2012)

28. Zheng, W.X.: Estimation of the parameters of autoregressive signals from colored noise-corrupted measurements. IEEE Signal Process. Lett. 7(7), 201–204 (2000)
29. Huang, Y.L., et al.: Design of Gaussian approximate filter and smoother for nonlinear systems with correlated noises at one epoch apart. Circ. Syst. Signal Process. 35(11), 3981–4008 (2016)
30. Saha, S., Gustafsson, F.: Particle filtering with dependent noise processes. IEEE Trans. Signal Process. 60(9), 4497–4508 (2012)
31. Gustafsson, F., Saha, S.: Particle filtering with dependent noise. In: Proceedings of 13th International Conference on Information Fusion, pp. 26–29, Edinburgh (2010)
32. Tichavsky, P., Muravchik, C.H., Nehorai, A.: Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. IEEE Trans. Signal Process. 46(5), 1386–1396 (1998)
33. Wang, Z.G., Shen, X.J., Zhu, Y.M.: Posterior Cramér-Rao bounds for nonlinear dynamic system with colored noises. J. Syst. Sci. Complex. 32(6), 1526–1543 (2019)
34. Fritsche, C., Saha, S., Gustafsson, F.: Bayesian Cramér-Rao bound for nonlinear filtering with dependent noise processes. In: Proceedings of 16th International Conference on Information Fusion, pp. 797–804, Istanbul (2013)
35. Huang, Y.L., Zhang, Y.G.: A new conditional posterior Cramér-Rao lower bound for a class of nonlinear systems. Int. J. Syst. Sci. 47(13), 3206–3218 (2016)
36. Geng, H., et al.: Joint estimation of target state and ionospheric height bias in over-the-horizon radar target tracking. IET Radar, Sonar Navig. 10(7), 1153–1167 (2016)
37. Li, S.H., Feng, X.X., Pan, F.: Joint state and parameter estimation for hidden wide-sense stationary ARMA processes under unknown noise. In: Proceedings of 20th International Conference on Information Fusion, pp. 1228–1235, Xi'an (2017)
38. Fortunati, S., et al.: On the application of the expectation-maximisation algorithm to the relative sensor registration problem. IET Radar, Sonar Navig. 7(2), 191–203 (2013)
39. Nikoofard, A., et al.: State and parameter estimation of a drift-flux model for underbalanced drilling operations. IEEE Trans. Contr. Syst. Technol. 25(6), 2000–2009 (2017)
40. Wielitzka, M., Dagen, M., Ortmaier, T.: Joint unscented Kalman filter for state and parameter estimation in vehicle dynamics. In: Proceedings of IEEE International Conference on Control Applications (CCA), pp. 1945–1950, Sydney (2015)
41. Manganiello, P., et al.: Dual-Kalman-filter-based identification and real-time optimization of PV systems. IEEE Trans. Ind. Electron. 62(11), 7266–7275 (2015)
42. Chen, H.M.: Adaptive cubature Kalman filter for nonlinear state and parameter estimation. In: Proceedings of 15th International Conference on Information Fusion, pp. 1413–1420, Singapore (2012)
43. Zia, A., et al.: An EM algorithm for nonlinear state estimation with model uncertainties. IEEE Trans. Signal Process. 56(3), 921–936 (2008)
44. Li, M.Z., et al.: Joint registration and multi-target tracking based on labelled random finite set and expectation maximisation. IET Radar, Sonar Navig. 12(3), 312–322 (2018)
45. Yang, J.X., Geller, B., Bay, S.: Bayesian and hybrid Cramér-Rao bounds for the carrier recovery under dynamic phase uncertain channels. IEEE Trans. Signal Process. 59(2), 667–680 (2011)
46. Bay, S., et al.: On the hybrid Cramér-Rao bound and its application to dynamical phase estimation. IEEE Signal Process. Lett. 15, 453–456 (2008)
47. Fortunati, S., et al.: Least squares estimation and Cramér-Rao type lower bounds for relative sensor registration process. IEEE Trans. Signal Process. 59(3), 1075–1087 (2011)
48. Gini, F., Reggiannini, R.: On the use of Cramér-Rao-like bounds in the presence of random nuisance parameters. IEEE Trans. Commun. 48(12), 2120–2126 (2000)
49. Trees, H.V., Bell, K.: Bayesian bounds for parameter estimation and nonlinear filtering/tracking. Wiley-IEEE, Hoboken (2007)
50. Messer, H.: The hybrid Cramér-Rao lower bound-from practice to theory. In: Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop, pp. 304–307, Waltham (2006)
51. Noam, Y., Messer, H.: Notes on the tightness of the hybrid Cramér-Rao lower bound. IEEE Trans. Signal Process. 57(6), 2074–2084 (2009)
52. Xu, L.F., Li, X.R.: Hybrid Cramér-Rao lower bound on tracking ground moving extended target. In: Proceedings of 12th International Conference on Information Fusion, pp. 1037–1044, Seattle (2009)
53. Ren, C., et al.: Recursive hybrid Cramér-Rao bound for discrete-time Markovian dynamic systems. IEEE Signal Process Lett. 22(10), 1543–4547 (2015)
54. Galy, J., et al.: Recursive hybrid CRB for Markovian systems with time-variant measurement parameters. In: Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), pp. 473–476, Cancun (2015)
55. Horn, R.A., Johnson, C.R.: Matrix analysis. Cambridge University Press, New York (2012)

How to cite this article: Li X, Duan Z, Hanebeck UD. Recursive joint Cramér-Rao lower bound for parametric systems with two-adjacent-states dependent measurements. *IET Signal Process*. 2021;15:221–237. https://doi.org/10.1049/sil2.12025

APPENDIX

Appendix 1: Proof of Theorem 1

Decompose $\boldsymbol{\chi}^k$ as $\boldsymbol{\chi}^k = [(\boldsymbol{x}^{k-1})', (\boldsymbol{x}_k)', \boldsymbol{\theta}']'$ and $\boldsymbol{J}^k$ as

$$\boldsymbol{J}^k = -E_{p_k}\begin{bmatrix} \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}}\ln p_k & \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}_k}\ln p_k & \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{\theta}}\ln p_k \\ \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}^{k-1}}\ln p_k & \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p_k & \Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}}\ln p_k \\ \Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}^{k-1}}\ln p_k & \Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}_k}\ln p_k & \Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}}\ln p_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}_k^{11} & \boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{13} \\ \boldsymbol{J}_k^{21} & \boldsymbol{J}_k^{22} & \boldsymbol{J}_k^{23} \\ \boldsymbol{J}_k^{31} & \boldsymbol{J}_k^{32} & \boldsymbol{J}_k^{33} \end{bmatrix}$$

The matrix inversion formula [55] is

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{C} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{D}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{E}^{-1} \\ -\boldsymbol{E}^{-1}\boldsymbol{B}'\boldsymbol{A}^{-1} & \boldsymbol{E}^{-1} \end{bmatrix} \tag{41}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are submatrices and $\boldsymbol{D} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}'$, $\boldsymbol{E} = \boldsymbol{C} - \boldsymbol{B}'\boldsymbol{A}^{-1}\boldsymbol{B}$.

Since $\boldsymbol{J}_k^{-1}$ is equal to the $(n+m) \times (n+m)$ right-lower block of $(\boldsymbol{J}^k)^{-1}$, and by using Equation (41) we have

$$\boldsymbol{J}_k = \begin{bmatrix} \boldsymbol{J}_k^{x,x} & \boldsymbol{J}_k^{x,\theta} \\ \boldsymbol{J}_k^{\theta,x} & \boldsymbol{J}_k^{\theta,\theta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}_k^{22} - \boldsymbol{J}_k^{21}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{23} - \boldsymbol{J}_k^{21}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{13} \\ \boldsymbol{J}_k^{32} - \boldsymbol{J}_k^{31}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{12} & \boldsymbol{J}_k^{33} - \boldsymbol{J}_k^{31}(\boldsymbol{J}_k^{11})^{-1}\boldsymbol{J}_k^{13} \end{bmatrix} \tag{42}$$

Similarly, decompose χ^{k+1} as $\chi^{k+1} = [(\boldsymbol{x}^{k-1})', (\boldsymbol{x}_k)', (\boldsymbol{x}_{k+1})', \boldsymbol{\theta}']'$ and $\boldsymbol{J}^{k+1}$ as

$$\boldsymbol{J}^{k+1} = -E_{p_{k+1}} \begin{bmatrix} \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}_k} \ln p_{k+1} & \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}_{k+1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{\theta}} \ln p_{k+1} \\ \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}^{k-1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \ln p_{k+1} & \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_{k+1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} \ln p_{k+1} \\ \Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}^{k-1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_k} \ln p_{k+1} & \Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{x}_{k+1}} \ln p_{k+1} & \Delta_{\boldsymbol{x}_{k+1}}^{\boldsymbol{\theta}} \ln p_{k+1} \\ \Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}^{k-1}} \ln p_{k+1} & \Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}_k} \ln p_{k+1} & \Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}_{k+1}} \ln p_{k+1} & \Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p_{k+1} \end{bmatrix}$$

where p_{k+1} is defined in Equation (3). Therefore, we have

$$\ln p_{k+1} = \ln p_k + \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta}) + \ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta})$$

Hence, we have

$$\boldsymbol{J}^{k+1} = \begin{bmatrix} E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_k) & E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}_k} \ln p_k) & \boldsymbol{0} & E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{\theta}} \ln p_k) \\ E_{p_{k+1}}(-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}^{k-1}} \ln p_k) & E_{p_{k+1}}(-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \ln p_k) + \boldsymbol{D}_k^{11} & \boldsymbol{D}_k^{12} & E_{p_{k+1}}(-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} \ln p_k) + \boldsymbol{D}_k^{13} \\ \boldsymbol{0} & \boldsymbol{D}_k^{21} & \boldsymbol{D}_k^{22} & \boldsymbol{D}_k^{23} \\ E_{p_{k+1}}(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}^{k-1}} \ln p_k) & E_{p_{k+1}}(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{x}_k} \ln p_k) + \boldsymbol{D}_k^{31} & \boldsymbol{D}_k^{32} & E_{p_{k+1}}(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p_k) + \boldsymbol{D}_k^{33} \end{bmatrix}$$

where $\boldsymbol{D}_k^{ij}$, $i = 1, 2, 3, j = 1, 2, 3$ are defined in Equation (6). Then,

$$\begin{aligned} E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_k) &= -\int p_{k+1} \cdot (\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_k) d_{\boldsymbol{x}^{k+1}} d_{\boldsymbol{z}^{k+1}} \\ &= -\int p_k \cdot p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta}) \cdot p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta}) \\ &\quad \cdot (\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_k) d_{\boldsymbol{x}^{k+1}} d_{\boldsymbol{z}^{k+1}} \\ &= -\int p_k (\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}^{k-1}} \ln p_k) d_{\boldsymbol{x}^k} d_{\boldsymbol{z}^k} \\ &= \boldsymbol{J}_k^{11} \end{aligned}$$

Similarly,

$$E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{x}_k} \ln p_k) = \boldsymbol{J}_k^{12}$$

$$E_{p_{k+1}}(-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k} \ln p_k) = \boldsymbol{J}_k^{22}$$

$$E_{p_{k+1}}(-\Delta_{\boldsymbol{x}^{k-1}}^{\boldsymbol{\theta}} \ln p_k) = \boldsymbol{J}_k^{13}$$

$$E_{p_{k+1}}(-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{\theta}} \ln p_k) = \boldsymbol{J}_k^{23}$$

$$E_{p_{k+1}}(-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \ln p_k) = \boldsymbol{J}_k^{33}$$

Hence,

$$\boldsymbol{J}^{k+1} = \left[\begin{array}{cc|cc} \boldsymbol{J}_k^{11} & \boldsymbol{J}_k^{12} & \boldsymbol{0} & \boldsymbol{J}_k^{13} \\ \boldsymbol{J}_k^{21} & \boldsymbol{J}_k^{22} + \boldsymbol{D}_k^{11} & \boldsymbol{D}_k^{12} & \boldsymbol{J}_k^{23} + \boldsymbol{D}_k^{13} \\ \hline \boldsymbol{0} & \boldsymbol{D}_k^{21} & \boldsymbol{D}_k^{22} & \boldsymbol{D}_k^{23} \\ \boldsymbol{J}_k^{31} & \boldsymbol{J}_k^{32} + \boldsymbol{D}_k^{31} & \boldsymbol{D}_k^{32} & \boldsymbol{J}_k^{33} + \boldsymbol{D}_k^{33} \end{array}\right] = \left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{C} & \boldsymbol{D} \end{array}\right]$$

Then, using the matrix inversion formula (Equation (41)), the information matrix $\boldsymbol{J}_{k+1}$ about $\boldsymbol{x}_{k+1}$ and $\boldsymbol{\theta}$ is

$$\boldsymbol{J}_{k+1} = \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B} = \begin{bmatrix} \boldsymbol{J}_{k+1}^{x,x} & \boldsymbol{J}_{k+1}^{x,\theta} \\ \boldsymbol{J}_{k+1}^{\theta,x} & \boldsymbol{J}_{k+1}^{\theta,\theta} \end{bmatrix} \tag{43}$$

Substituting Equation (42) into Equation (43), the recursive form is obtained as Equations (5) and (6).

Appendix 2: Proof of Corollary 1

From the assumptions that the noises are additive Gaussian noises, we have

$$\begin{aligned} \ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k, \boldsymbol{\theta}_x) = c_1 - \frac{1}{2}(\boldsymbol{x}_{k+1} - f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x))' \\ \cdot \boldsymbol{Q}_k^{-1}(\boldsymbol{x}_{k+1} - f_k(\boldsymbol{x}_k, \boldsymbol{\theta}_x)) \end{aligned} \tag{44}$$

$$\begin{aligned} &\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta}_z) \\ &= c_2 - \frac{1}{2}[\boldsymbol{z}_{k+1} - h_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta}_z)]' \\ &\quad \cdot \boldsymbol{R}_{k+1}^{-1}[\boldsymbol{z}_{k+1} - h_{k+1}(\boldsymbol{x}_{k+1}, \boldsymbol{x}_k, \boldsymbol{\theta}_z)] \end{aligned} \tag{45}$$

where c_1 and c_2 are constants.

Therefore, for calculating the $\boldsymbol{D}_k^{11,a}$, the partial derivatives of $\ln p(\boldsymbol{x}_{k+1} | \boldsymbol{x}_k, \boldsymbol{\theta}_x)$ are

$$-\nabla_{\boldsymbol{x}_k}\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}_x)$$

$$=\nabla_{\boldsymbol{x}_k}\left[\frac{1}{2}(\boldsymbol{x}_{k+1}-f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))'\boldsymbol{Q}_k^{-}(\boldsymbol{x}_{k+1}-f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))\right]$$

$$=\nabla_{\boldsymbol{x}_k}\frac{1}{2}\left[\boldsymbol{x}'_{k+1}\boldsymbol{Q}_k^{-1}\boldsymbol{x}_{k+1}-\boldsymbol{x}'_{k+1}\boldsymbol{Q}_k^{-1}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\right.$$

$$\left.-f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\boldsymbol{x}_{k+1}+f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))\right]$$

$$=\frac{1}{2}[-2\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\boldsymbol{x}_{k+1}$$

$$+2\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]$$

$$=\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}(f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)-\boldsymbol{x}_{k+1}) \tag{46}$$

$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}_x)$$
$$=-\nabla_{\boldsymbol{x}_k}\nabla'_{\boldsymbol{x}_k}\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}_x)$$
$$=\nabla_{\boldsymbol{x}_k}[(f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)-\boldsymbol{x}'_{k+1})\boldsymbol{Q}_k^{-1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]$$
$$=\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)$$
$$+\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x) \tag{47}$$
$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\boldsymbol{x}_{k+1}$$
$$=\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)$$
$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}(\boldsymbol{x}_{k+1}-f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))$$

By substituting Equation (47) into Equation (6), we have

$$\boldsymbol{D}_k^{11,a}=E_{p_{k+1}}[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}_x)]$$
$$=E_{p_{k+1}}\left[\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\right.$$
$$\left.-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}(\boldsymbol{x}_{k+1}-f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))\right]$$
$$=E_{p_{k+1}}[\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{Q}_k^{-1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]$$

in which the following identity has been used

$$E_{p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k,\boldsymbol{\theta}_x)}[\boldsymbol{x}_{k+1}-f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]=0$$

For calculating the $\boldsymbol{D}_k^{11,b}$, the partial derivatives of $\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$ is

$$-\nabla_{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$

$$=\nabla_{\boldsymbol{x}_k}\left[\frac{1}{2}(\boldsymbol{z}_{k+1}-h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z))'\right.$$

$$\left.\cdot\boldsymbol{R}_{k+1}^{-1}(\boldsymbol{z}_{k+1}-h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z))\right]$$

$$=\nabla_{\boldsymbol{x}_k}\frac{1}{2}\left[\boldsymbol{z}'_{k+1}\boldsymbol{R}_{k+1}^{-1}\boldsymbol{z}_{k+1}-\boldsymbol{z}'_{k+1}\boldsymbol{R}_{k+1}^{-1}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\right.$$

$$-h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}\boldsymbol{z}_{k+1}$$

$$\left.+h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\right]$$

$$=\frac{1}{2}\left[-2\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}\boldsymbol{z}_{k+1}\right.$$

$$\left.+2\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\right]$$

$$=\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}(h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)-\boldsymbol{z}_{k+1}) \tag{48}$$

$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$=-\nabla_{\boldsymbol{x}_k}\nabla'_{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$=\nabla_{\boldsymbol{x}_k}[(h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)-\boldsymbol{z}'_{k+1})$$
$$\times\boldsymbol{R}_{k+1}^{-1}\nabla_{\boldsymbol{x}_k}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]$$
$$=\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}\nabla_{\boldsymbol{x}_k}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$+\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}\boldsymbol{z}_{k+1}$$
$$=\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\boldsymbol{R}_{k+1}^{-1}\nabla_{\boldsymbol{x}_k}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)(\boldsymbol{z}_{k+1}-h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)) \tag{49}$$

By substituting Equation (49) into Equation (6), we have

$$\boldsymbol{D}_k^{11,b}=E_{p_{k+1}}[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]$$
$$=E_{p_{k+1}}\left[\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)\right.$$
$$\times\boldsymbol{R}_{k+1}^{-1}\nabla_{\boldsymbol{x}_k}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$\left.-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)(\boldsymbol{z}_{k+1}-h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z))\right]$$
$$=E_{p_{k+1}}[\nabla_{\boldsymbol{x}_k}h'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)$$
$$\times\boldsymbol{R}_{k+1}^{-1}\nabla_{\boldsymbol{x}_k}h_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta}_z)]$$

where the following identity has been used

$$E_{p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1},\mathbf{x}_k,\boldsymbol{\theta}_z)}[\mathbf{z}_{k+1} - h_{k+1}(\mathbf{x}_{k+1},\mathbf{x}_k,\boldsymbol{\theta}_z)] = 0$$

Similarly, $\mathbf{D}_k^{12}$, $\mathbf{D}_k^{13}$, $\mathbf{D}_k^{22}$, $\mathbf{D}_k^{23}$ and $\mathbf{D}_k^{33}$ can be obtained.

Appendix 3: Proof of Corollary 2

From the assumptions that the noises are additive Gaussian noises, we can get

$$\ln p(\mathbf{x}_{k+1}|\mathbf{x}_k,\boldsymbol{\theta}_x) = c_3 - \frac{1}{2}(\mathbf{x}_{k+1} - \mathbf{F}_k\mathbf{x}_k - \mathbf{F}_k^{\theta}\boldsymbol{\theta}_x)' \cdot \mathbf{Q}_k^{-1}(\mathbf{x}_{k+1} - \mathbf{F}_k\mathbf{x}_k - \mathbf{F}_k^{\theta}\boldsymbol{\theta}_x) \quad (50)$$

$$\ln p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1},\mathbf{x}_k,\boldsymbol{\theta}_z) = c_4 - \frac{1}{2}[\mathbf{z}_{k+1} - \mathbf{H}_{k+1}\mathbf{x}_{k+1} - \mathbf{C}_k\mathbf{x}_k - \mathbf{H}_{k+1}^{\theta}\boldsymbol{\theta}_z]' \cdot \mathbf{R}_k^{-1}[\mathbf{z}_{k+1} - \mathbf{H}_{k+1}\mathbf{x}_{k+1} - \mathbf{C}_k\mathbf{x}_k - \mathbf{H}_{k+1}^{\theta}\boldsymbol{\theta}_z] \quad (51)$$

where c_3 and c_4 are constants and

$$\begin{aligned}\mathbf{D}_k^{12,a} &= E_{p_{k+1}}[-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k+1}}\ln p(\mathbf{x}_{k+1}|\mathbf{x}_k,\boldsymbol{\theta}_x)]\\ &= E_{p_{k+1}}\Big\{\frac{1}{2}\nabla_{\mathbf{x}_k}\nabla_{\mathbf{x}_{k+1}}\big(\mathbf{x}_{k+1}'\mathbf{Q}_k^{-1}\mathbf{x}_{k+1} - \mathbf{x}_{k+1}'\mathbf{Q}_k^{-1}\mathbf{F}_k\mathbf{x}_k\\ &\quad - \mathbf{x}_{k+1}'\mathbf{Q}_k^{-1}\mathbf{F}_k^{\theta}\boldsymbol{\theta}_x - \mathbf{x}_k'\mathbf{F}_k'\mathbf{Q}_k^{-1}\mathbf{x}_{k+1}\\ &\quad + \mathbf{x}_k'\mathbf{F}_k'\mathbf{Q}_k^{-1}\mathbf{F}_k\mathbf{x}_k - (\mathbf{F}_k^{\theta}\boldsymbol{\theta}_x)'\mathbf{Q}_k^{-1}\mathbf{x}_{k+1}\big)'\Big\}\\ &= E_{p_{k+1}}\Big\{\frac{1}{2}\nabla_{\mathbf{x}_k}[2\mathbf{Q}_k^{-1}\mathbf{x}_{k+1} - \mathbf{Q}_k^{-1}\mathbf{F}_k\mathbf{x}_k\\ &\quad - \mathbf{Q}_k^{-1}\mathbf{F}_k\mathbf{x}_k - \mathbf{Q}_k^{-1}\mathbf{F}_k^{\theta}\boldsymbol{\theta}_x]'\Big\}\\ &= E_{p_{k+1}}\Big\{\frac{1}{2}\nabla_{\mathbf{x}_k}[-2\mathbf{x}_k'\mathbf{F}_k'\mathbf{Q}_k^{-1}]\Big\}\\ &= -E_{p_{k+1}}\{\mathbf{F}_k'\mathbf{Q}_k^{-1}\}\\ &= -\mathbf{F}_k'\mathbf{Q}_k^{-1}\end{aligned}$$

$$\begin{aligned}\mathbf{D}_k^{12,b} &= E_{p_{k+1}}[-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k+1}}\ln p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1},\mathbf{x}_k)]\\ &= E_{p_{k+1}}\Big\{\nabla_{\mathbf{x}_k}\nabla_{\mathbf{x}_{k+1}}'\Big[\frac{1}{2}\big(\mathbf{x}_{k+1}'\mathbf{H}_{k+1}'\mathbf{R}_{k+1}^{-1}\mathbf{C}_k\mathbf{x}_k\\ &\quad + \mathbf{x}_k'\mathbf{C}_k'\mathbf{R}_{k+1}^{-1}\mathbf{H}_{k+1}\mathbf{x}_{k+1}\big]\Big\}\\ &= E_{p_{k+1}}\Big\{\nabla_{\mathbf{x}_k}\frac{1}{2}[\mathbf{H}_{k+1}'\mathbf{R}_{k+1}^{-1}\mathbf{C}_k\mathbf{x}_k + \mathbf{H}_{k+1}'\mathbf{R}_{k+1}^{-1}\mathbf{C}_k\mathbf{x}_k]'\Big\}\\ &= E_{p_{k+1}}\{\nabla_{\mathbf{x}_k}[\mathbf{x}_k'\mathbf{C}_k'\mathbf{R}_{k+1}^{-1}\mathbf{H}_{k+1}]\}\\ &= E_{p_{k+1}}\{\mathbf{C}_k'\mathbf{R}_{k+1}^{-1}\mathbf{H}_{k+1}\}\\ &= \mathbf{C}_k'\mathbf{R}_{k+1}^{-1}\mathbf{H}_{k+1}\end{aligned}$$

Similarly, $\mathbf{D}_k^{11}$, $\mathbf{D}_k^{13}$, $\mathbf{D}_k^{22}$, $\mathbf{D}_k^{23}$ and $\mathbf{D}_k^{33}$ can be obtained.

Appendix 4: Proof of Theorem 2

From the assumptions that the noises are additive Gaussian noises, we can get

$$\ln p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1},\mathbf{x}_k,\boldsymbol{\theta}_z) = c_5 - \frac{1}{2}[\mathbf{z}_{k+1} - l_{k+1}(\mathbf{x}_{k+1},\boldsymbol{\theta}_z) + \boldsymbol{\Psi}_k l_k(\mathbf{x}_k,\boldsymbol{\theta}_z)]' \cdot \mathbf{R}_k^{-1}[\mathbf{z}_{k+1} - l_{k+1}(\mathbf{x}_{k+1},\boldsymbol{\theta}_z) + \boldsymbol{\Psi}_k l_k(\mathbf{x}_k,\boldsymbol{\theta}_z)] \quad (52)$$

where c_5 is a constant.

The proof of $\mathbf{D}_k^{11,a}$ can refer to the proof of Corollary 1. By substituting Equation (47) into Equation (6), we have

$$\begin{aligned}\mathbf{D}_k^{11,a} &= E_{p_{k+1}}[-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}\ln p(\mathbf{x}_{k+1}|\mathbf{x}_k,\boldsymbol{\theta}_x)]\\ &= E_{p_{k+1}}\{[\nabla_{\mathbf{x}_k} f_k'(\mathbf{x}_k,\boldsymbol{\theta}_x)]\mathbf{Q}_k^{-1}[\nabla_{\mathbf{x}_k} f_k'(\mathbf{x}_k,\boldsymbol{\theta}_x)]'\}\end{aligned}$$

For calculating the $\mathbf{D}_k^{11,b}$, the partial derivatives of $\ln p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1},\mathbf{x}_k,\boldsymbol{\theta}_z)$ is

$$
\begin{aligned}
&-\nabla_{x_k}\ln p(z_{k+1}|x_{k+1},x_k,\theta_z)\\
&=\nabla_{x_k}\left\{\frac{1}{2}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)+\Psi_k l_k(x_k,\theta_z)]'\right.\\
&\quad\left.\cdot R_k^{-1}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)+\Psi_k l_k(x_k,\theta_z)]\right\}\\
&=\nabla_{x_k}\frac{1}{2}[z'_{k+1}R_k^{-1}\Psi_k l_k(x_k,\theta_z)-l'_{k+1}(x_{k+1},\theta_z)\\
&\quad\times R_k^{-1}\Psi_k l_k(x_k,\theta_z)\\
&\quad+l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}z_{k+1}-l'_k(x_k,\theta_z)\\
&\quad\times\Psi'_k R_k^{-1}l_{k+1}(x_{k+1},\theta_z)\\
&\quad+l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}\Psi_k l_k(x_k,\theta_z)]\\
&=\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}z_{k+1}-\nabla_{x_k}l'_k(x_k,\theta_z)\\
&\quad\times\Psi'_k R_k^{-1}l_{k+1}(x_{k+1},\theta_z)
\end{aligned}
\tag{53}
$$

$$
\begin{aligned}
&\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}\Psi_k l_k(x_k,\theta_z)\\
&=\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}(z_{k+1}-l_{k+1}(x_{k+1},\theta_z)\\
&\quad+\Psi_k l_k(x_k,\theta_z))
\end{aligned}
$$

$$
\begin{aligned}
&-\Delta_{x_k}^{x_k}\ln p(z_{k+1}|x_{k+1},x_k,\theta_z)\\
&=-\nabla_{x_k}\nabla'_{x_k}\ln p(z_{k+1}|x_{k+1},x_k,\theta_z)\\
&=\nabla_{x_k}\{[z'_{k+1}-l'_{k+1}(x_{k+1},\theta_z)+l'_k(x_k,\theta_z)\Psi'_k]\\
&\quad\cdot R_k^{-1}\Psi_k\nabla_{x_k}l_k(x_k,\theta_z)\}\\
&=\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}\Psi_k\nabla_{x_k}l_k(x_k,\theta_z)\\
&\quad+\Delta_{x_k}^{x_k}l'_k(x_k,\theta_z)\Psi'_k\\
&\quad\cdot R_k^{-1}(z_{k+1}-l_{k+1}(x_{k+1},\theta_z)+\Psi_k l_k(x_k,\theta_z))
\end{aligned}
\tag{54}
$$

By substituting Equation (49) into Equation (6), we have

$$
\begin{aligned}
D_k^{11,b}&=E_{p_{k+1}}[-\Delta_{x_k}^{x_k}\ln p(z_{k+1}|x_{k+1},x_k,\theta_z)]\\
&=E_{p_{k+1}}[\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k R_k^{-1}\Psi_k\nabla_{x_k}l_k(x_k,\theta_z)]\\
&=E_{p_{k+1}}\{[\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k]R_k^{-1}[\nabla_{x_k}l'_k(x_k,\theta_z)\Psi'_k]'\}
\end{aligned}
$$

where the following identity has been used

$$
E_{p(z_{k+1}|x_{k+1},x_k,\theta_z)}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)+\Psi_k l_k(x_k,\theta_z)]=0
$$

Similarly, D_k^{12}, D_k^{13}, D_k^{22}, D_k^{23} and D_k^{33} can be obtained.

Appendix 5: Proof of Theorem 3

From the reconstructed pseudo-measurement Equation (30), we can get the likelihood PDF:

$$
\begin{aligned}
&\ln p(z_{k+1}|x_{k+1},x_k,\theta)\\
&=c_6-\frac{1}{2}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)\\
&\quad-G_{k+1}(x_{k+1}-f_k(x_k,\theta_x))]'\\
&\quad\cdot R_{k+1}^{-1}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)\\
&\quad-G_{k+1}(x_{k+1}-f_k(x_k,\theta_x))]
\end{aligned}
\tag{55}
$$

where c_6 is a constant.

The proof of $D_k^{11,a}$ can refer to the proof of Corollary 1. By substituting Equation (47) into Equation (6), we have

$$
\begin{aligned}
D_k^{11,a}&=E_{p_{k+1}}[-\Delta_{x_k}^{x_k}\ln p(x_{k+1}|x_k,\theta_x)]\\
&=E_{p_{k+1}}\{[\nabla_{x_k}f'_k(x_k,\theta_x)]Q_k^{-1}[\nabla_{x_k}f'_k(x_k,\theta_x)]'\}
\end{aligned}
$$

When calculating the $D_k^{11,b}$, the partial derivatives of $\ln p(z_{k+1}|x_{k+1},x_k,\theta)$ is

$$
\begin{aligned}
&-\nabla_{x_k}\ln p(z_{k+1}|x_{k+1},x_k,\theta)\\
&=\nabla_{x_k}\left\{\frac{1}{2}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)-G_{k+1}(x_{k+1}-f_k(x_k,\theta_x))]'\right.\\
&\quad\left.\cdot R_{k+1}^{-1}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)-G_{k+1}(x_{k+1}-f_k(x_k,\theta_x))]\right\}\\
&=\nabla_{x_k}\frac{1}{2}[z'_{k+1}R_{k+1}^{-1}G_{k+1}f_k(x_k,\theta_x)+f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}z_{k+1}\\
&\quad-l'_{k+1}(x_{k+1},\theta_z)R_{k+1}^{-1}G_{k+1}f_k(x_k,\theta_x)\\
&\quad-x'_{k+1}G'_{k+1}R_{k+1}^{-1}G_{k+1}f_k(x_k,\theta_x)\\
&\quad-f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}l_{k+1}(x_{k+1},\theta_z)\\
&\quad-f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}G_{k+1}x_{k+1}\\
&\quad+f'_k(x_k,\theta_x)G'_{k+1}R_k^{-1}G_{k+1}f_k(x_k,\theta_x)]\\
&=\nabla_{x_k}f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}z_{k+1}\\
&\quad-\nabla_{x_k}f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}G_{k+1}x_{k+1}\\
&\quad-\nabla_{x_k}f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}l_{k+1}(x_{k+1},\theta_z)\\
&\quad+\nabla_{x_k}f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}G_{k+1}f_k(x_k,\theta_x)\\
&=\nabla_{x_k}f'_k(x_k,\theta_x)G'_{k+1}R_{k+1}^{-1}[z_{k+1}-l_{k+1}(x_{k+1},\theta_z)\\
&\quad-G_{k+1}(x_{k+1}-f_k(x_k,\theta_x))]
\end{aligned}
\tag{56}
$$

$$
\begin{aligned}
&-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})\\
&\quad= -\nabla_{\boldsymbol{x}_k}\nabla'_{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})\\
&\quad= \nabla_{\boldsymbol{x}_k}\{[\boldsymbol{z}'_{k+1} - l'_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{\theta}_z) - (\boldsymbol{x}'_{k+1} - f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_{\boldsymbol{x}}))\boldsymbol{G}'_{k+1}]\\
&\qquad \cdot\boldsymbol{R}_{k+1}^{-1}\boldsymbol{G}_{k+1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\}\\
&\quad= \nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{G}'_{k+1}\boldsymbol{R}_{k+1}^{-1}\boldsymbol{G}_{k+1}\\
&\qquad\times\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x) + \Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\\
&\qquad\cdot\boldsymbol{G}'_{k+1}\boldsymbol{R}_{k+1}^{-1}[\boldsymbol{z}_{k+1} - l_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{\theta}_z)\\
&\qquad - \boldsymbol{G}_{k+1}(\boldsymbol{x}_{k+1} - f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))]
\end{aligned}
\tag{57}
$$

By substituting Equation (57) into Equation (6), we have

$$
\begin{aligned}
\boldsymbol{D}_k^{11,b} &= E_{p_{k+1}}[-\Delta_{\boldsymbol{x}_k}^{\boldsymbol{x}_k}\ln p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})]\\
&= E_{p_{k+1}}[\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_{\boldsymbol{x}})\boldsymbol{G}'_{k+1}\boldsymbol{R}_{k+1}^{-1}\boldsymbol{G}_{k+1}\nabla_{\boldsymbol{x}_k}f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)]\\
&= E_{p_{k+1}}\{[\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{G}'_{k+1}]\boldsymbol{R}_k^{-1}[\nabla_{\boldsymbol{x}_k}f'_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x)\boldsymbol{G}'_{k+1}]'\}
\end{aligned}
$$

where the following identity has been used:

$$
\begin{aligned}
E_{p(\boldsymbol{z}_{k+1}|\boldsymbol{x}_{k+1},\boldsymbol{x}_k,\boldsymbol{\theta})}[\boldsymbol{z}_{k+1} - l_{k+1}(\boldsymbol{x}_{k+1},\boldsymbol{\theta}_z)\\
-\boldsymbol{G}_{k+1}(\boldsymbol{x}_{k+1} - f_k(\boldsymbol{x}_k,\boldsymbol{\theta}_x))] = 0
\end{aligned}
$$

Similarly, $\boldsymbol{D}_k^{12}$, $\boldsymbol{D}_k^{13}$, $\boldsymbol{D}_k^{22}$, $\boldsymbol{D}_k^{23}$ and $\boldsymbol{D}_k^{33}$ can be obtained.