Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

Binary-Encoding-Based Quantized Kalman Filter: An Approximate MMSE Approach

Qinyuan Liu, Yao Nie, Zidong Wang, Hongli Dong, and Changjun Jiang

***Abstract*—In this paper, the Kalman filter design problem is investigated for linear discrete-time systems under binary encoding schemes. Under such a scheme, the local information is quantized into a bit string by the remote sensor based on a probabilistic quantizer, and then the bit string is transmitted via memoryless binary symmetric channels (BSCs). Due to the communication link noises, the bit flipping occurs in a random manner, and thus, the transmission of the bit string would suffer from specific bit-error rates. With the received bits, a recursive binary-encoding-based quantized Kalman filter is established in the approximate minimum mean-square error (MMSE) sense, which relies on the Gaussian approximation of the conditional probability density function at each iteration. Furthermore, the proposed estimator is shown to be in a Kalman-like type through performance analysis, which exhibits computational complexity comparable to the conventional Kalman filter. Subsequently, a posterior Cramér-Rao lower bound is derived for the proposed binary-encoding-based quantized Kalman filter. The effectiveness of the proposed estimator is demonstrated through numerical results.**

***Index terms*—** Networked systems, Kalman filter, probabilistic quantizer, binary encoding scheme, iterative Bayesian estimate, minimum mean-square error.

I. Introduction

Benefiting from the tremendous advances in digital communication, information sensing, and electronic technologies, the rapid development of networked systems has been witnessed in the past few decades [5], [10]–[12], [45]. Networked systems have shown a surge of potential applications in various fields such as industrial automation [9], smart grids [4], [7], and transportation systems [18], owing to their significant advantages in reliability, flexibility, and adaptability [49]. Among these applications, one of the fundamental challenges is the development of appropriate remote estimation schemes to recover the system's internal state based on a series of measurement signals transmitted through digital networks, e.g., [6], [13], [35]. Consequently, considerable research attention has been devoted to this topic under various network environments, including packet losses, signal fading, communication delays, and so on [15], [20]–[22], [39], [41], [43], [44], [48].

It is worth noting that the majority of modern automatic systems are designed to store, process, and communicate information in digital form. Due to constrained bandwidth in communication channels, analog measurements collected by sensors need to be quantized into a finite number of bits before transmission to the remote estimation center. If not appropriately handled, the quantization effects could significantly degrade system performance and even lead to divergence. As a result, there has been a great deal of research interest focused on the problem of signal quantization. Generally, quantizers can be classified into different types based on their structures, with examples including linear [33], logarithmic [14], [23], and Lloyd-Max quantizers [34]. Also, some other types of quantizers have been developed by optimizing the quantized error between the input and output [19], [32].

Most existing quantizers adopt a deterministic truncation principle, which introduces unknown quantization errors and leads to biased quantized signals that are often difficult to handle. To address this challenge, probabilistic quantization has been explored in more recent years, e.g., [2], [25], [26]. By implementing truncation functions in a probabilistic manner, probabilistic quantizers ensure that the quantization error has a zero mean, thereby simplifying subsequent system analysis to a great extent. In fact, remote estimation with quantized measurements has emerged as a cutting-edge research topic in both academia and industry. In the quantization process, the truncation error resulting from the quantization process introduces nonlinearities even if the target plant is linear, and this makes quantization-based estimation an essentially nonlinear estimation problem.

To address the quantization-induced nonlinearities, researchers have treated the quantization error as norm-bounded uncertainties in [16], where a robust estimator has been designed to provide a suboptimal estimate using a min-max approach. Various nonlinear filtering techniques have been employed for estimation with quantized measurements to achieve a more accurate minimum mean-square error (MMSE) estimate. Examples of such techniques can be found in [17], [29], [31], [36] and references therein. In [36], the distributed estimation has been investigated with single-bit quantization, and a Gaussian approximation approach has been employed to derive an approximate MMSE estimate. The results have been further extended in [29] to the distributed quantized Kalman filter with multiple quantization bits, where the trade-offs between bandwidth requirements (represented by the

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB4501704, in part by the National Science Foundation of China under Grants 62222312 and 62473285, in part by the Shanghai Science and Technology Innovation Action Plan Project of China under Grant 22511100700, in part by Fundamental Research Funds for the Central Universities, in part by the Royal Society of the U.K., and in part by the Alexander von Humboldt Foundation of Germany.
(Corresponding author: Hongli Dong)

Q. Liu, Y. Nie, and C. Jiang are with the School of Computer Science and Technology, Tongji University, Shanghai 201804, China, with the Key Laboratory of Embedded System and Service Computing, Ministry of Education, Shanghai 201804, China, and also with the Shanghai Artificial Intelligence Laboratory, Shanghai, China. Email: liuqy@tongji.edu.cn)

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. Email: Zidong.Wang@brunel.ac.uk.

H. Dong is with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China, and also with the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Northeast Petroleum University, Daqing 163318, China. Email: shiningdhl@vip.126.com.

Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

number of quantization bits) and system performance have been thoroughly discussed.

In the existing literature concerning quantization-based state estimation, an implicit assumption is that the encoded bit string of the quantization output can be transmitted *accurately*. However, due to channel noises, distortions, signal interference, and errors in transmitter-receiver clock synchronization, bit flipping frequently occurs during transmission, resulting in inevitable communication errors. To accurately capture both quantization and network characteristics, a more practical communication model known as binary encoding schemes (BESs) has been widely used in the field of communication [24], [27], [47]. Under such a scheme, the sensor information is quantized and encoded into a bit string, which is then transmitted via memoryless binary symmetric channels (BSCs), and recovered at the estimation center for further processing. BSCs have a wide range of applications, including eavesdropped communication, radar signal encoding, telemetry, and voice communication [1], [3], [28]. To date, various studies have focused on the estimation problem involving BESs. For instance, in [24], the distributed parameter estimation has been investigated for wireless sensor networks under BESs, and a closed-form bound has been derived on the mean-square error at the estimation center. A common feature of these results is that the estimation performance has been evaluated by establishing an upper bound on the error covariance matrix (ECM), which may be conservative in certain situations. Therefore, our primary motivation in this paper is to develop a more accurate MMSE estimate that eliminates unnecessary conservatism.

Based on the above discussion, the aim of this paper is to develop the approximate MMSE estimate for linear discrete-time systems under BESs. Because of the complicated transmission errors resulting from stochastic quantization and random bit flipping, the analytical expression of the MMSE estimate cannot be directly obtained in practical applications. It is, therefore, of particular significance to deal with this challenge by developing an approximate MMSE estimator that considers both accuracy and computational feasibility. This paper will employ an iterative Bayesian estimate approach with Gaussian approximation on the conditional probability density function (PDF) at each iteration to derive a binary-encoding-based quantized Kalman filter. Subsequently, in-depth analysis is carried out on the system performance, including the influence of the quantization effects, the computational complexity, and the lower bound on the ECM of the proposed binary-encoding-based quantized Kalman filter.

The main contributions of this paper are highlighted in the following aspects: *i) to the best of our knowledge, this paper is one of the first few attempts to design approximate MMSE estimators for dynamical systems under binary encoding schemes within an iterative Bayesian framework; ii) the proposed binary-encoding-based quantized Kalman filter exhibits a Kalman-like type, demonstrating computational complexity comparable to the conventional Kalman filter; and iii) the corresponding theoretical performance bounds, i.e., the posterior Cramér-Rao lower bound (PCRLB), are derived in order to evaluate the performance of the proposed estimator.*

The organization of our paper is as follows. Section II provides system descriptions including the dynamic model and the BESs. Section III presents the derivation of the binary-encoding-based quantized Kalman filter under BESs in the approximate MMSE sense. In Section IV, the computational complexity and the PCRLB of the proposed estimator are discussed. Section V presents the numerical simulation results. Finally, in Section VI, concluding remarks are provided for this work.

Notations: The matrix inequality $\mathbf{A} \succeq \mathbf{B}$ ($\mathbf{A} \succ \mathbf{B}$) represents that $\mathbf{A} - \mathbf{B}$ is a positive semi-definite (definite) matrix. $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{-1}$ and $\mathrm{tr}\{\mathbf{A}\}$ denote the transpose, the inverse and the trace of the matrix $\mathbf{A}$, respectively. $\mathrm{diag}\{x_1, x_2, \ldots, x_N\}$ denotes a diagonal matrix with entries $\{x_1, x_2, \ldots, x_N\}$ on the diagonal. $\{x_i\}_{i=1}^{N}$ represents the set $\{x_1, x_2, \ldots, x_N\}$. The Gaussian probability density function (PDF) with mean $\mathbb{E}\{\mathbf{x}\} = \mu$ and covariance matrix $\mathrm{Cov}\{\mathbf{x}\} = \Sigma$ is represented as $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mu, \Sigma)$ and $\varphi(t) = \int_{-\infty}^{t} \mathcal{N}(x; 0, 1)\mathrm{d}x$ denotes the standard Gaussian cumulative distribution function (CDF). 1_A means a indicator function, which is equal to 1 when the event A occurs and 0 otherwise.

II. System Description

Consider a linear discrete time-varying system described by the following state-space model

$$\mathbf{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{w}_n, \tag{1}$$

where $\mathbf{x}_n \in \mathbb{R}^N$ is the system state and $\mathbf{w}_n \in \mathbb{R}^N$ is the sequence of process noises which is assumed to be Gaussian with zero mean and covariance $\Sigma_{\mathbf{w}} \succeq 0$. $\mathbf{A}_n \in \mathbb{R}^{N \times N}$ is a known system matrix. The sensor measurement is described by

$$\mathbf{y}_n = \mathbf{C}_n \mathbf{x}_n + \mathbf{v}_n, \tag{2}$$

where $\mathbf{y}_n \in \mathbb{R}^M$ stands for the measurement output and $\mathbf{v}_n \in \mathbb{R}^M$ is the sequence of measurement noises which is assumed to be Gaussian with zero mean and covariance $\Sigma_{\mathbf{v}} \succeq 0$. $\mathbf{C}_n \in \mathbb{R}^{M \times N}$ is a known measurement matrix.

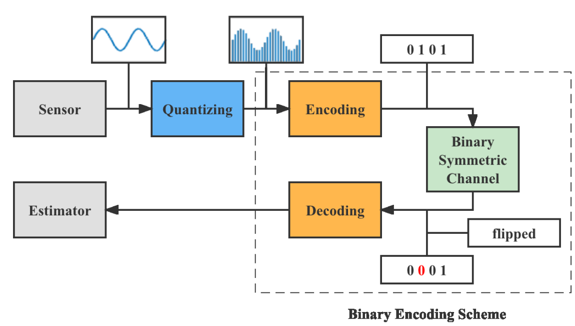


Fig. 1. Remote state estimation under binary encoding schemes.

We consider the remote state estimation problem as shown in Fig. 1, where the BESs are utilized to transmit the sensor information. To be specific, a probabilistic quantizer is first employed to convert the transmitted signals into a finite codebook as follows

$$\mathcal{Q}(\cdot) \triangleq \begin{bmatrix} \mathcal{Q}_1(\cdot) & \mathcal{Q}_2(\cdot) & \cdots & \mathcal{Q}_M(\cdot) \end{bmatrix}^{\mathrm{T}},$$

where $\mathcal{Q}_k(\cdot)$ is a quantizer for scalar input. The static quantization strategy is utilized with one quantizer per component

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

for vector signals, that is, for arbitrary vector $\mathbf{b}_n \in \mathbb{R}^M$, the map of the quantizer is given by

$$\mathcal{Q}(\mathbf{b}_n) = \begin{bmatrix} \mathcal{Q}_1(b_{n,1}) & \mathcal{Q}_2(b_{n,2}) & \cdots & \mathcal{Q}_M(b_{n,M}) \end{bmatrix}^{\mathrm{T}},$$

where $b_{n,k}$ is the kth component of $\mathbf{b}_n$.

For brevity of analysis, we consider that $b_{n,k}$ is within $[-W, W]$, where W is a positive scalar. Let L denote the number of bits for the encoder output. The quantization range is uniformly divided into $2^L - 1$ segments and, correspondingly, we have 2^L points denoted as follows

$$\mathcal{U} \triangleq \{\tau_1, \tau_2, \dots, \tau_{2^L}\}.$$

The interval length is $\Delta = \tau_{i+1} - \tau_i$, for $i = 1, 2, \dots, 2^L - 1$, and thus it is apparent that $\Delta = \frac{2W}{2^L - 1}$.

When $\tau_i \leq b_{n,k} \leq \tau_{i+1}$, $b_{n,k}$ is quantized according to the following probabilistic manner

$$\begin{cases} \mathbb{P}\{\mathcal{Q}_k(b_{n,k}) = \tau_i\} = 1 - r_{n,k} \\ \mathbb{P}\{\mathcal{Q}_k(b_{n,k}) = \tau_{i+1}\} = r_{n,k} \end{cases} \tag{3}$$

with $r_{n,k} \triangleq (b_{n,k} - \tau_i)/\Delta$ and $0 \leq r_{n,k} \leq 1$. On the basis of binary bits $h_{n,k}^{(l)}$, the quantization output $\mathcal{Q}_k(b_{n,k})$ can be further represented by

$$\mathcal{Q}_k(b_{n,k}) = -W + \sum_{l=1}^{L} h_{n,k}^{(l)} 2^{l-1} \Delta$$

and, therefore, $\mathcal{Q}_k(b_{n,k})$ can be encoded into the following binary bit string

$$\mathcal{D}_{n,k} \triangleq \{h_{n,k}^{(1)}, h_{n,k}^{(2)}, \dots, h_{n,k}^{(L)}\}, \quad h_{n,k}^{(l)} \in \{0, 1\}$$

Remark 1. *The rational $r_{n,k}$ represents the probability that the signal $b_{n,k}$ is quantized to the right bound τ_{i+1} of ith interval. As $b_{n,k}$ approaches the right boundary τ_{i+1}, the distance between $b_{n,k}$ and τ_i increases, resulting in a larger value of $r_{n,k}$. Consequently, the probability that $b_{n,k}$ is quantized to τ_{i+1} becomes higher, and vice versa. (See Fig. 2)*

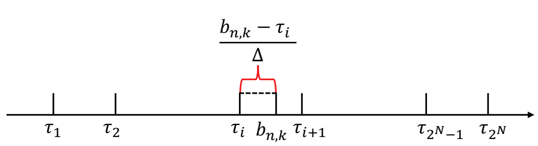


Fig. 2. The stochastic quantization scheme.

Next, $\mathcal{D}_{n,k}$ will be transmitted by a memoryless BSC with crossover probability p, where the received binary bit string can be presented as follows:

$$\mathcal{D}'_{n,k} \triangleq \{\hbar_{n,k}^{(1)}, \hbar_{n,k}^{(2)}, \dots, \hbar_{n,k}^{(L)}\}, \quad \hbar_{n,k}^{(l)} \in \{0, 1\}.$$

Here, $\hbar_{n,k}^{(l)} = \theta_{n,k}^{(l)}(1 - h_{n,k}^{(l)}) + (1 - \theta_{n,k}^{(l)}) h_{n,k}^{(l)}$ with

$$\theta_{n,k}^{(l)} = \begin{cases} 1, \text{ if the } l\text{th bit is flipped} \\ 0, \text{ if the } l\text{th bit is not flipped} \end{cases} \tag{4}$$

and $\mathbb{P}\{\theta_{n,k}^{(l)} = 1\} = p$. With the bit string $\mathcal{D}'_{n,k}$, the signals are decoded as follows

$$m_{n,k} = -W + \sum_{l=1}^{L} \hbar_{n,k}^{(l)} 2^{l-1} \Delta.$$

Moreover, the signals received at the estimator during the interval $[0, n]$ are denoted as follow

$$\mathcal{M}_n \triangleq \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n\}$$

where $\mathbf{m}_k = \begin{bmatrix} m_{k,1}, m_{k,2}, \cdots, m_{k,M} \end{bmatrix}^{\mathrm{T}}$.

Throughout the paper, the normalized innovation sequence

$$\tilde{\mathbf{y}}_n \triangleq \mathbf{S}_{n|n-1}^{-\frac{1}{2}} (\mathbf{y}_n - \hat{\mathbf{y}}_{n|n-1}) \tag{5}$$

is taken as the quantizer input, which will be further transmitted to the remote estimator using BESs. Here, $\hat{\mathbf{y}}_{n|n-1}$ is the one-step prediction of the measurement $\mathbf{y}_n$, i.e., $\hat{\mathbf{y}}_{n|n-1} \triangleq \mathbb{E}\{\mathbf{y}_n | \mathcal{M}_{n-1}\}$ and $\mathbf{S}_{n|n-1} \triangleq \mathrm{Cov}\{\mathbf{y}_n | \mathcal{M}_{n-1}\}$ is the corresponding measurement prediction ECM.

Before proceeding, the following assumptions are made.

Assumption 1. *The initial state $\mathbf{x}_0$ obeys Gaussian distribution with mean $\mu_0 \in \mathbb{R}^N$ and covariance matrix $\Sigma_0 \in \mathbb{R}^{N \times N}$, i.e. $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0; \mu_0, \Sigma_0)$.*

Assumption 2. *The process noise $\mathbf{w}_n$ and measurement noise $\mathbf{v}_n$ are mutually independent, i.e.*

$$\mathbb{E}\left\{ \begin{bmatrix} \mathbf{w}_n \\ \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \mathbf{w}_n^{\mathrm{T}}, \mathbf{v}_n^{\mathrm{T}} \end{bmatrix} \right\} = \begin{bmatrix} \Sigma_{\mathbf{w}} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\mathbf{v}} \end{bmatrix}$$

Assumption 3. *The indicators $\theta_{n,k}^{(l)}$ defined in* (4) *are white and mutually independent.*

III. Binary-Encoding-Based Quantized Kalman Filter

In this section, a binary-encoding-based quantized Kalman filter in the approximate MMSE sense will be developed based on an iterative Bayesian approach.

Firstly, we consider a posterior MMSE estimate of $\mathbf{x}_n$ with the channel outputs $\mathcal{M}_n$ through BSCs at instant n as follows:

$$\hat{\mathbf{x}}_n \triangleq \mathbb{E}\{\mathbf{x}_n \mid \mathcal{M}_n\} = \int_{\mathbb{R}^N} \mathbf{x}_n p(\mathbf{x}_n \mid \mathcal{M}_n) \, \mathrm{d}\mathbf{x}_n, \tag{6}$$

and the corresponding ECM is defined as

$$\mathbf{P}_n \triangleq \mathbb{E}\{(\mathbf{x}_n - \hat{\mathbf{x}}_n)(\mathbf{x}_n - \hat{\mathbf{x}}_n)^{\mathrm{T}} | \mathcal{M}_n\}. \tag{7}$$

The posterior distribution $p(\mathbf{x}_n | \mathcal{M}_n)$ in (6) can be calculated by using the prediction-correction steps [P1]-[C1] outlined as follows.

[P1] **Prediction step:** Given the prior PDF $p(\mathbf{x}_{n-1} | \mathcal{M}_{n-1})$ at instant n, we calculate the distribution $p(\mathbf{x}_n | \mathcal{M}_{n-1})$ by the law of total probability

$$p(\mathbf{x}_n \mid \mathcal{M}_{n-1}) = \int_{\mathbb{R}^N} p(\mathbf{x}_n \mid \mathbf{x}_{n-1}, \mathcal{M}_{n-1}) \times p(\mathbf{x}_{n-1} \mid \mathcal{M}_{n-1}) \, \mathrm{d}\mathbf{x}_{n-1}.$$

As $\mathbf{x}_n$ depends completely on $\mathbf{x}_{n-1}$ and $\mathbf{w}_{n-1}$, the condition $\mathcal{M}_{n-1}$ is redundant. Consequently, we have

$$p(\mathbf{x}_n | \mathbf{x}_{n-1}, \mathcal{M}_{n-1}) = \mathcal{N}(\mathbf{x}_n; \mathbf{A}_{n-1}\mathbf{x}_{n-1}, \Sigma_{\mathbf{w}}).$$

[C1] **Correction step:** When the channel outputs $\mathcal{M}_n$ are available at instant n, the posterior PDF $p(\mathbf{x}_n \mid \mathcal{M}_n)$ can be derived based on the following Bayesian rule

$$p(\mathbf{x}_n \mid \mathcal{M}_n) = \frac{\mathbb{P}\{\mathbf{m}_n \mid \mathcal{M}_{n-1}, \mathbf{x}_n\}}{\mathbb{P}\{\mathbf{m}_n \mid \mathcal{M}_{n-1}\}} p(\mathbf{x}_n \mid \mathcal{M}_{n-1})$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

where the probability $\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1},\mathbf{x}_n\}$ and $\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1}\}$ can be obtained with the knowledge of the quantization rule $\mathcal{Q}_k(\cdot)$ and the statistics of the random bit flipping $\theta_{n,k}^{(l)}$. In fact,

$$\frac{\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1},\mathbf{x}_n\}}{\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1}\}}$$

can be viewed as a correction coefficient that extracts the newly arrival information to update the prediction.

Noticing that $\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1},\mathbf{x}_n\}=\mathbb{P}\{\mathbf{m}_n|\mathbf{x}_n\}$, we have

$$\begin{aligned}\mathbb{P}\{\mathbf{m}_n \mid \mathcal{M}_{n-1},\mathbf{x}_n\} &= \int_{\mathbb{R}^M}\mathbb{P}\left\{\tilde{\mathbf{y}}_n,\mathbf{m}_n \mid \mathbf{x}_n\right\}\ \mathrm{d}\tilde{\mathbf{y}}_n\\ &= \int_{\mathbb{R}^M}\mathbb{P}\left\{\mathbf{m}_n \mid \tilde{\mathbf{y}}_n,\mathbf{x}_n\right\}p(\tilde{\mathbf{y}}_n \mid \mathbf{x}_n)\ \mathrm{d}\tilde{\mathbf{y}}_n\end{aligned}$$

where $\mathbb{P}\{\mathbf{m}_n|\tilde{\mathbf{y}}_n,\mathbf{x}_n\} = \mathbb{P}\{\mathbf{m}_n|\tilde{\mathbf{y}}_n\}$ and $p(\tilde{\mathbf{y}}_n|\mathbf{x}_n)$ obeys the Gaussian distribution. Moreover, the probability $\mathbb{P}\{\mathbf{m}_n|\mathcal{M}_{n-1}\}$ can be derived as follows:

$$\mathbb{P}\left\{\mathbf{m}_n \mid \mathcal{M}_{n-1}\right\} = \int_{\mathbb{R}^M}\mathbb{P}\left\{\mathbf{m}_n \mid \tilde{\mathbf{y}}_n\right\}p(\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1})\ \mathrm{d}\tilde{\mathbf{y}}_n$$

where

$$p(\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}) = \int_{\mathbb{R}^N}p(\tilde{\mathbf{y}}_n \mid \mathbf{x}_n)p(\mathbf{x}_n \mid \mathcal{M}_{n-1})\ \mathrm{d}\mathbf{x}_n$$

The computation of the MMSE estimate using [P1]-[C1] involves integrating several nonlinear functions at each iteration, making it impossible to obtain a closed-form solution. Utilizing numerical integrations is a feasible approach, but the significant computational burden limits its practical applicability. An alternative workaround is to approximate the conditional PDF at the previous step using a Gaussian distribution, i.e.,

$$p(\mathbf{x}_{n-1}|\mathcal{M}_{n-1}) \sim \mathcal{N}(\mathbf{x}_n;\hat{\mathbf{x}}_{n-1},\mathbf{P}_{n-1}). \tag{8}$$

This method simplifies coping with the analytically intractable PDF $p(\mathbf{x}_n|\mathcal{M}_n)$ by tracking its expectation and covariance matrix.

Remark 2. *In this paper, the distribution of $\mathbf{x}_n$ conditioned on the quantized measurements $\mathcal{M}_n$ under binary encoding schemes is approximated by a Gaussian distribution. The rationality of the Gaussian approximation lies in the Bayesian Gaussian mixture model (GMM) theory. Theoretically, the weighted Gaussian sum can fit arbitrary distribution according to Wiener's tauberian theorem [38]. In this paper, the posterior distribution $p(\mathbf{x}_n|\mathcal{M}_n)$ can be approximated by GMM as follows:*

$$p(\mathbf{x}_n|\mathcal{M}_n) \approx \sum_{i=1}^{K} h_i\mathcal{N}(\hat{\mathbf{x}}_n^i,\mathbf{P}_n^i), \tag{9}$$

where K is the number of Gaussian clusters. For the sake of derivation convenience and computational resource saving, we take a special case of Gaussian mixture model where $K=1$. In this case, the values of $\hat{\mathbf{x}}_{n-1}^1$ and $\mathbf{P}_{n-1}^1$ are taken as the expectation and the covariance of the conditional distribution, $\hat{\mathbf{x}}_{n-1}$ and $\mathbf{P}_{n-1}$, respectively.

Obviously, according to assumption (8), $\hat{\mathbf{y}}_{n|n-1}$ and $\mathbf{S}_{n|n-1}$ can be calculated based on their definitions as follows:

$$\begin{aligned}\hat{\mathbf{y}}_{n|n-1} &= \mathbb{E}\left\{\mathbf{C}_n\mathbf{x}_n+\mathbf{v}_n \mid \mathcal{M}_{n-1}\right\} = \mathbf{C}_n\mathbf{A}_{n-1}\hat{\mathbf{x}}_{n-1},\\ \mathbf{S}_{n|n-1} &= \mathrm{Cov}\left\{\mathbf{C}_n\mathbf{x}_n+\mathbf{v}_n \mid \mathcal{M}_{n-1}\right\}\\ &= \mathbf{C}_n(\mathbf{A}_{n-1}\mathbf{P}_{n-1}\mathbf{A}_{n-1}^{\mathrm{T}}+\Sigma_{\mathbf{w}})\mathbf{C}_n^{\mathrm{T}}+\Sigma_{\mathbf{v}}.\end{aligned} \tag{10}$$

Noting $\tilde{\mathbf{y}}_n \triangleq \mathbf{S}_{n|n-1}^{-\frac{1}{2}}(\mathbf{y}_n-\hat{\mathbf{y}}_{n|n-1})$, we know from the Gaussian approximation that $\tilde{\mathbf{y}}_n \triangleq [\tilde{y}_{n,1},\tilde{y}_{n,2},\cdots,\tilde{y}_{n,M}]^{\mathrm{T}}$ is an M-component standard Gaussian distribution.

For sake of brevity, we utilize a set of binary variables $\{\theta_{j\to i}^{(l)}\}_{l=0}^{L}$ with $\theta_{j\to i}^{(l)} \in \{0,1\}$ to specify the random flipping $\{\theta_{n,k}^{(l)}\}_{l=0}^{L}$ for the case where the channel input τ_j is decoded into τ_i after transmission. For example, suppose that the channel input is τ_j represented by

$$\tau_j = -W+\sum_{l=1}^{L} h_{n,k}^{(l)}2^{l-1}\Delta$$

and the received message is τ_i given by

$$\tau_i = -W+\sum_{l=1}^{L} \hbar_{n,k}^{(l)}2^{l-1}\Delta.$$

Apparently, the relationship

$$\hbar_{n,k}^{(l)} = \theta_{j\to i}^{(l)}(1-h_{n,k}^{(l)})+(1-\theta_{j\to i}^{(l)})h_{n,k}^{(l)}$$

holds, and we can see that $\theta_{ji}=\sum_{l=1}^{L}\theta_{j\to i}^{(l)}$ counts the number of the flipped bits. Furthermore, we denote

$$\begin{aligned}&h(\tau) = \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2})+\tau\varphi(\tau),\\ &f(\tau) = -\varphi(\tau), \quad g(\tau) = \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2}).\end{aligned} \tag{11}$$

With the Gaussian approximation, a binary-encoding-based quantized Kalman filter (BQKF) can now be developed to provide the approximate MMSE estimate as follows.

Theorem 1. *If $p(\mathbf{x}_{n-1}|\mathcal{M}_{n-1}) \sim \mathcal{N}(\mathbf{x}_n;\hat{\mathbf{x}}_{n-1},\mathbf{P}_{n-1})$, then the BQKF can be derived through the following prediction-correction steps*

[P2] **Prediction step:**

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \mathbf{A}_{n-1}\hat{\mathbf{x}}_{n-1},\\ \mathbf{P}_{n|n-1} &= \mathbf{A}_{n-1}\mathbf{P}_{n-1}\mathbf{A}_{n-1}^{\mathrm{T}}+\Sigma_{\mathbf{w}},\end{aligned} \tag{12}$$

[C2] **Correction step:**

$$\begin{aligned}\mathbf{K}_n &= \mathbf{P}_{n|n-1}\mathbf{C}_n^{\mathrm{T}}\mathbf{S}_{n|n-1}^{-\frac{1}{2}},\\ \hat{\mathbf{x}}_n &= \hat{\mathbf{x}}_{n|n-1}+\mathbf{K}_n\mathbf{g}_n,\\ \mathbf{P}_n &= \mathbf{P}_{n|n-1}-\mathbf{K}_n\mathbf{G}_n\mathbf{K}_n^{\mathrm{T}},\end{aligned} \tag{13}$$

where

$$\mathbf{g}_n \triangleq [\alpha_{m_{n,1}},\alpha_{m_{n,1}},\cdots,\alpha_{m_{n,M}}]^{\mathrm{T}}$$

and

$$\mathbf{G}_n \triangleq \mathrm{diag}\{\beta_{m_{n,1}},\beta_{m_{n,2}},\ldots,\beta_{m_{n,M}}\}$$

are coefficient vector and matrix which correct the prediction and ECM, respectively. Moreover, α_i and β_i are given as

$$\begin{aligned}\alpha_i &= \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\frac{F(\tau_j)}{H(\tau_j)},\\ \beta_i &= \alpha_i^2-\sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\frac{G(\tau_j)}{H(\tau_j)},\end{aligned}$$

where

$$\begin{aligned}H(\tau) &= 1_{\{\tau\neq W\}}\left[h(\tau+\Delta)-h(\tau)+f(\tau)\Delta\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[h(\tau-\Delta)-h(\tau)-f(\tau)\Delta\right],\\ F(\tau) &= 1_{\{\tau\neq W\}}\left[f(\tau+\Delta)-f(\tau)+g(\tau)\Delta\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[f(\tau-\Delta)-f(\tau)-g(\tau)\Delta\right],\\ G(\tau) &= 1_{\{\tau\neq W\}}\left[g(\tau+\Delta)-(1-\Delta\tau)g(\tau)\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[g(\tau-\Delta)-(1+\Delta\tau)g(\tau)\right].\end{aligned}$$

Proof: See Appendix A. ■

The proposed BQKF in the above ftheorem, composed of the prediction step [P2] and the correction step [C2] like Kalman filter (KF), is an approximate MMSE estimate for linear discrete-time systems under BESs. To evaluate the difference between BQKF and standard KF, we define the mean ECM corrections for these two filters. For standard KF, it is well acknowledged that $\mathbf{P}_n = \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{K}_n^{\mathrm{T}}$, and therefore the mean ECM correction (given the ideal information $\mathcal{M}'_{n-1}$) is defined as follows:

$$\Delta\mathbf{P}_n^{\mathrm{KF}} = \mathbb{E}\left\{\mathbf{P}_{n|n-1} - \mathbf{P}_n \mid \mathcal{M}'_{n-1}\right\} = \mathbf{K}_n\mathbf{K}_n^{\mathrm{T}} \tag{14}$$

As for BQKF, according to (13), the expectation of the mean ECM correction (given $\mathcal{M}_{n-1}$) can be obtained as follows:

$$\begin{aligned}\Delta\mathbf{P}_n^{\mathrm{BQKF}} &= \mathbb{E}\left\{\mathbf{P}_{n|n-1} - \mathbf{P}_n \mid \mathcal{M}_{n-1}\right\}\\ &= \mathbf{K}_n\mathbb{E}\left\{\mathbf{G}_n \mid \mathcal{M}_{n-1}\right\}\mathbf{K}_n^{\mathrm{T}}\end{aligned} \tag{15}$$

Comparing (14) with (15), we find that the difference between the mean ECM corrections of BQKF and KF simply lies in the divergence between the identity matrix $\mathbf{I}$ with the term $\bar{\mathbf{G}}_n \triangleq \mathbb{E}\{\mathbf{G}_n|\mathcal{M}_{n-1}\}$, where the latter one indicates the influence of randomness brought by BESs and stochastic quantizers on the ECM correction. Then, we have the following result.

Corollary 1. *For BQKF with prediction-correction steps [P2]-[C2], the following inequalities hold:*

$$\mathbf{0} \preceq \bar{\mathbf{G}}_n \preceq \mathbf{I},$$

where $\bar{\mathbf{G}}_n = \mathbb{E}\{\mathbf{G}_n|\mathcal{M}_{n-1}\}$ is the expectation of $\mathbf{G}_n$ conditioned on $\mathcal{M}_{n-1}$.

Proof: Noting that $\bar{\mathbf{G}}_n$ is a diagonal matrix, we denote the kth element of $\bar{\mathbf{G}}_n$ as $\bar{\beta}_{n,k}$. We have shown in Appendix A that $\beta_i = 1 - \mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_n\}$. Since $\mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_n\} \geq 0$, it is trivial to see $\beta_i \leq 1$. Thus, we have $\bar{\beta}_{n,k} = \mathbb{E}\{\beta_i|\mathcal{M}_{n-1}\} \leq 1$. On the other hand, according to [37], the following property holds:

$$\mathrm{Cov}\{X\} = \mathbb{E}\{\mathrm{Cov}\{X|Y\}\} + \mathrm{Cov}\{\mathbb{E}\{X|Y\}\}$$

and, accordingly, one has

$$\mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}\} \geq \mathbb{E}_{\mathbf{m}_n}\{\mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_n\}\}.$$

From the definition of $\tilde{\mathbf{y}}_n$, we know $\mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}\} = 1$. Therefore, $\bar{\beta}_{n,k} = 1 - \mathbb{E}_{\mathbf{m}_n}\{\mathrm{Var}\{\tilde{y}_{n,k}|\mathcal{M}_n\}\} \geq 0$. To this end, we can draw the conclusion that $0 \leq \bar{\beta}_{n,k} \leq 1$ hold for $k = 1, 2, \ldots, M$. Since $\bar{\mathbf{G}}_n$ is diagonal, it is not difficult to observe that $\mathbf{0} \preceq \bar{\mathbf{G}}_n \preceq \mathbf{I}$, which ends the proof. ■

From the above corollary, it can be observed that the ECM correction of the proposed BQKF is lower than that of the KF since $\bar{\mathbf{G}}_n \preceq \mathbf{I}$. This difference arises primarily from the penalty induced by quantization and transmission errors. Moreover, as $\bar{\mathbf{G}}_n \succeq \mathbf{0}$, we can anticipate a decrease in the (approximate) ECM after the correction step, which suggests that the received normalized innovation $\mathbf{m}_n$ still provides valuable information even after quantization and bit flipping.

In the following analysis, we will consider a special case where the flipping probability $p = 0$, meaning that no bit flipping occurs during transmission. In this scenario, the BQKF can be simplified into a more concise form as follows.

Proposition 1. *If $p(\mathbf{x}_{n-1}|\mathcal{M}_{n-1}) \sim \mathcal{N}(\mathbf{x}_n; \hat{\mathbf{x}}_{n-1}, \mathbf{P}_{n-1})$, and the flipping probability $p = 0$, then the BQKF in [P2]-[C2] reduces to the following prediction-correction steps*

[P3] **Prediction step:**

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1} &= \mathbf{A}_{n-1}\hat{\mathbf{x}}_{n-1|n-1}\\ \mathbf{P}_{n|n-1} &= \mathbf{A}_{n-1}\mathbf{P}_{n-1|n-1}\mathbf{A}_{n-1}^{\mathrm{T}} + \Sigma_{\mathbf{w}}\end{aligned}$$

[C3] **Correction step:**

$$\begin{aligned}\hat{\mathbf{x}}_n &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n\mathbf{g}_n\\ \mathbf{P}_n &= \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{G}_n\mathbf{K}_n^{\mathrm{T}}\end{aligned}$$

where $\mathbf{K}_n$, $\mathbf{g}_n$ and $\mathbf{G}_n$ are defined in Theorem 1. Moreover, α_i and β_i can be calculated as follow

$$\alpha_i = \frac{F(\tau_i)}{H(\tau_i)}, \quad \beta_i = \alpha_i^2 - \frac{G(\tau_i)}{H(\tau_i)}$$

Proof: The above proposition can be easily proven by setting $p = 0$. Therefore, the proof is omitted here for the sake of brevity. ■

The following corollary can be deduced from Proposition 1.

Corollary 2. *For BQKF with prediction-correction steps [P3]-[C3], we have $\mathbf{G}_n \to \mathbf{I}$ when $\Delta \to 0$.*

Proof: Consider the kth entry on the diagonal of $\mathbf{G}_n$. Since $\beta_i = \alpha_i^2 - G(\tau_i)/H(\tau_i)$, we consider the term $\lim_{\Delta\to 0}\alpha_i$ at first. For $i = 2, \ldots, 2^L - 1$, $F(\tau_i) = f(\tau_i - \Delta) + f(\tau_i + \Delta) - 2f(\tau_i)$, $G(\tau_i) = g(\tau_i - \Delta) + g(\tau_i + \Delta) - 2g(\tau_i)$ and $H(\tau_i) = h(\tau_i - \Delta) + h(\tau_i + \Delta) - 2h(\tau_i)$. When Δ tends to 0, both the numerator and the denominator of $\alpha_i = F(\tau_i)/H(\tau_i)$ tend to 0. Moreover, the partial derivatives of $F(\tau_i)$ and $H(\tau_i)$ with respect to Δ can be calculated as follows:

$$\begin{aligned}\frac{\partial F}{\partial \Delta} &= \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i-\Delta)^2}{2}) - \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i+\Delta)^2}{2})\\ \frac{\partial H}{\partial \Delta} &= \varphi(\tau_i+\Delta) - \varphi(\tau_i-\Delta)\end{aligned} \tag{16}$$

and

$$\begin{aligned}\frac{\partial^2 F}{\partial \Delta^2} &= \tfrac{\tau_i-\Delta}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i-\Delta)^2}{2}) + \tfrac{\tau_i+\Delta}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i+\Delta)^2}{2})\\ \frac{\partial^2 H}{\partial \Delta^2} &= \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i-\Delta)^2}{2}) + \tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{(\tau_i+\Delta)^2}{2})\end{aligned} \tag{17}$$

Applying L'Hospital's rule yields

$$\lim_{\Delta\to 0}\alpha_i = \frac{\sqrt{2\pi}}{\sqrt{2\pi}}\frac{2\tau_i\exp(-\frac{\tau_i^2}{2})}{2\exp(-\frac{\tau_i^2}{2})} = \tau_i. \tag{18}$$

Moreover, applying L'Hospital's rule on $G(\tau_i)/H(\tau_i)$, we have

$$\lim_{\Delta\to 0}\frac{G}{H}=\frac{\sqrt{2\pi}}{\sqrt{2\pi}}\frac{2\tau_i^2\exp(-\frac{\tau_i^2}{2})-2\exp(-\frac{\tau_i^2}{2})}{2\exp(-\frac{\tau_i^2}{2})}=\tau_i^2-1 \quad (19)$$

Therefore, $\lim_{\Delta\to 0}\beta_i=\lim_{\Delta\to 0}\alpha_i^2-G(\tau_i)/H(\tau_i)=\tau_i^2-(\tau_i^2-1)=1$. The same conclusion can be proven in a similar manner for the cases of $i=1$ or $i=2^L$. Therefore, the proof is omitted for brevity. Finally, we have $\lim_{\Delta\to 0}\mathbf{G}_n=\mathbf{I}$. ■

According to Corollary 2, it is evident that selecting a smaller interval length for the quantizer will decrease the quantization errors and thereby provide more effective information for the estimator. Furthermore, as Δ tends to zero, the quantization errors vanish, resulting in $\mathbf{G}_n\to\mathbf{I}$ and the BQKF degenerates to the traditional KF.

Remark 3. *In practical applications, the selection of an appropriate quantization interval, denoted as W, is crucial. Since $\tilde{y}_{n,k}$ follows a standard Gaussian distribution, a feasible choice is $W=3$ as it corresponds to a probability of $\mathbb{P}\{-W\le\tilde{y}_{n,k}\le W\}=99.73\%$, encompassing the majority of the distribution. Next, the length of the bit stream, denoted as L, needs to be determined based on the network capacity. Subsequently, the thresholds $\mathcal{U}=\{\tau_1,\tau_2,\ldots,\tau_{2^L}\}$ are set. From* (11)*, it can be observed that $f(\tau_i)$, $g(\tau_i)$ and $h(\tau_i)$ are solely dependent on $\{\tau_i\}_{i=1}^{2^L}$. Consequently, the values of α_i and β_i can be computed and stored in a coefficient table prior to implementation. This precomputation simplifies the online calculations, improving computational efficiency.*

IV. Performance analysis

In this section, we will address the computational complexity of the proposed BQKF and provide the posterior Cram'er-Rao lower bound for performance evaluation.

A. Computational Complexity

To analyze the computational complexity of the proposed estimator, it is important to consider the computational complexities of various operations involved.

The computational complexity of multiplying an $m\times n$ and an $n\times p$ matrix is $O(nmp)$ (for $m\neq n\neq p$), adding two $m\times n$ matrices is $O(mn)$, and the transpose operation is $O(n^2)$, respectively [30]. Additionally, the computational complexity of Cholesky decomposition of an $n\times n$ matrix is $O(n^3)$.

Regarding the computation of $\hat{\mathbf{x}}_n$, it involves the following three steps.

1) Calculation of the prediction and the corresponding ECM, with a computational complexity of $O(N^3)$.
2) Calculation of the gain $\mathbf{K}_n$ defined in (13) with a computational complexity of $\max\{O(N^2M),O(NM^2),O(N^3)\}$.
3) Calculation of the estimate and the corresponding ECM, with a computational complexity of $\max\{O(N^2M),O(NM^2)\}$.

Therefore, the total computational complexity of the proposed BQKF at each instant is $O(\max\{N^2M,NM^2,N^3\})=O(N\cdot\max\{N,M\}^2)$. This complexity is consistent with that of the traditional KF.

B. Posterior Cramér-Rao Lower Bound

The posterior Cramér-Rao lower bound (PCRLB) provides a lower bound on the estimation ECM of the proposed BQKF. Specifically, the actual estimation ECM of the BQKF is bounded below by the inverse of the Fisher information matrix (FIM) $\mathbf{J}_n$, i.e.

$$\mathbb{E}\{(\hat{\mathbf{x}}_n-\mathbf{x}_n)(\hat{\mathbf{x}}_n-\mathbf{x}_n)^{\mathrm{T}}\}\succeq\mathbf{J}_n^{-1} \quad (20)$$

Without loss of generality, we will calculate the FIM for the dynamical systems with the assumption of $\Sigma_{\mathbf{v}}=\sigma_v^2\mathbf{I}$. In fact, even in cases where $\Sigma_{\mathbf{v}}$ is not a diagonal matrix, we can perform a diagonalization by transforming the measurement equation.

For convenience of presentation, let $\mathbf{C}_n^{(k,:)}$ be the kth row of the matrix $\mathbf{C}_n$, and denote

$$\begin{aligned}&\psi(\tau)\triangleq\tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2})+\tau\varphi(\tau),\\&\dot{\psi}(\tau)\triangleq\varphi(\tau),\quad\ddot{\psi}(\tau)\triangleq\tfrac{1}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2}).\end{aligned} \quad (21)$$

Now, we have the following proposition.

Proposition 2. *With the initial value $\mathbf{J}_0=\Sigma_0^{-1}$, the FIM of the proposed BQKF can be recursively calculated as follows:*

$$\mathbf{J}_n=(\Sigma_{\mathbf{w}}+\mathbf{A}_{n-1}\mathbf{J}_{n-1}^{-1}\mathbf{A}_{n-1}^{\mathrm{T}})^{-1}+\mathbf{C}_n^{\mathrm{T}}\bar{\Xi}_n^{-1}\mathbf{C}_n \quad (22)$$

where $\bar{\Xi}_n=\mathrm{diag}\{\rho_{n,1},\rho_{n,2},\ldots,\rho_{n,M}\}$ is a diagonal matrix with the diagonal entry as

$$\rho_{n,k}=\sigma_v^2\ \mathbb{E}\left\{\left(\frac{\dot{\Upsilon}_{n,k}}{\Upsilon_{n,k}}\right)^2-\frac{\ddot{\Upsilon}_{n,k}}{\Upsilon_{n,k}}\right\}^{-1}, \quad (23)$$

for $k=1,2,\ldots,M$, which can be approximately computed by Monte Carlo sampling:

$$\rho_{n,k}\approx\frac{\sigma_v^2}{T_{\mathrm{MC}}}\sum_{t=1}^{T_{\mathrm{MC}}}\left[\left(\frac{\dot{\Upsilon}_{n,k}^{(t)}}{\Upsilon_{n,k}^{(t)}}\right)^2-\frac{\ddot{\Upsilon}_{n,k}^{(t)}}{\Upsilon_{n,k}^{(t)}}\right]^{-1}. \quad (24)$$

The superscript "(t)" implies that the value is obtained from the t-th sampling, and T_{MC} denotes the total number of sampling. Moreover,

$$\mu_{n,k}^{(t)}=\frac{\mathbf{C}_n^{(k,:)}(\mathbf{x}_n^{(t)}-\hat{\mathbf{x}}_{n|n-1}^{(t)})}{\sqrt{\mathbf{C}_n^{(k,:)}\mathbf{P}_{n|n-1}^{(t)}(\mathbf{C}_n^{(k,:)})^{\mathrm{T}}+\sigma_v^2}}$$

$$\sigma_{n,k}^{(t)}=\frac{\sigma_v}{\sqrt{\mathbf{C}_n^{(k,:)}\mathbf{P}_{n|n-1}^{(t)}(\mathbf{C}_n^{(k,:)})^{\mathrm{T}}+\sigma_v^2}}$$

$$\psi_{n,k}^{(t)}(\tau)=\psi\Big(\frac{\tau-\mu_{n,k}^{(t)}}{\sigma_{n,k}^{(t)}}\Big)$$

$$\dot{\psi}_{n,k}^{(t)}(\tau)=\dot{\psi}\Big(\frac{\tau-\mu_{n,k}^{(t)}}{\sigma_{n,k}^{(t)}}\Big)$$

$$\ddot{\psi}_{n,k}^{(t)}(\tau)=\ddot{\psi}\Big(\frac{\tau-\mu_{n,k}^{(t)}}{\sigma_{n,k}^{(t)}}\Big)$$

$$\begin{aligned}\Psi_{n,k}^{(t)}(\tau) &= 1_{\{\tau\neq W\}}\left[\psi_{n,k}^{(t)}(\tau+\Delta)-\psi_{n,k}^{(t)}(\tau)-\tfrac{\Delta}{\sigma_{n,k}^{(t)}}\dot{\psi}_{n,k}^{(t)}(\tau)\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[\psi_{n,k}^{(t)}(\tau-\Delta)-\psi_{n,k}^{(t)}(\tau)+\tfrac{\Delta}{\sigma_{n,k}^{(t)}}\dot{\psi}_{n,k}^{(t)}(\tau)\right]\\ \dot{\Psi}_{n,k}^{(t)}(\tau) &= 1_{\{\tau\neq W\}}\left[\dot{\psi}_{n,k}^{(t)}(\tau+\Delta)-\dot{\psi}_{n,k}^{(t)}(\tau)-\tfrac{\Delta}{\sigma_{n,k}^{(t)}}\ddot{\psi}_{n,k}^{(t)}(\tau)\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[\dot{\psi}_{n,k}^{(t)}(\tau-\Delta)-\dot{\psi}_{n,k}^{(t)}(\tau)+\tfrac{\Delta}{\sigma_{n,k}^{(t)}}\ddot{\psi}_{n,k}^{(t)}(\tau)\right]\\ \ddot{\Psi}_{n,k}^{(t)}(\tau) &= 1_{\{\tau\neq W\}}\left[\ddot{\psi}_{n,k}^{(t)}(\tau+\Delta)-\left(1-\Delta\tfrac{\tau-\mu_{n,k}^{(t)}}{(\sigma_{n,k}^{(t)})^2}\right)\ddot{\psi}_{n,k}^{(t)}(\tau)\right]\\ &\quad+1_{\{\tau\neq -W\}}\left[\ddot{\psi}_{n,k}^{(t)}(\tau-\Delta)-\left(1+\Delta\tfrac{\tau-\mu_{n,k}^{(t)}}{(\sigma_{n,k}^{(t)})^2}\right)\ddot{\psi}_{n,k}^{(t)}(\tau)\right]\end{aligned}$$

$$\begin{aligned}\Upsilon_{n,k}^{(t)} &= \sum_{j=1}^{2^L} p^{\theta_{ji}^{(t)}}(1-p)^{L-\theta_{ji}^{(t)}}\Psi_{n,k}^{(t)}(\tau_j)\\ \dot{\Upsilon}_{n,k}^{(t)} &= \sum_{j=1}^{2^L} p^{\theta_{ji}^{(t)}}(1-p)^{L-\theta_{ji}^{(t)}}\dot{\Psi}_{n,k}^{(t)}(\tau_j)\\ \ddot{\Upsilon}_{n,k}^{(t)} &= \sum_{j=1}^{2^L} p^{\theta_{ji}^{(t)}}(1-p)^{L-\theta_{ji}^{(t)}}\ddot{\Psi}_{n,k}^{(t)}(\tau_j)\end{aligned}$$

Proof: See Appendix B. ■

If the normalized innovation can be transmitted accurately, i.e., $\mathbf{m}_n = \tilde{\mathbf{y}}_n$, the BQKF simplifies to a clairvoyant KF. In this case, the second term of $\mathbf{J}_n$ in (22) becomes $\mathbf{C}_n^{\mathrm{T}}\Sigma_{\mathbf{v}}^{-1}\mathbf{C}_n$ [42]. A closer examination of $\mathbf{C}_n^{\mathrm{T}}\Sigma_{\mathbf{v}}^{-1}\mathbf{C}_n$ and $\mathbf{C}_n^{\mathrm{T}}\Xi_n^{-1}\mathbf{C}_n$ reveals that the main difference arises from the effects of stochastic quantization and random bit flipping during transmission.

C. Stability Analysis

The above subsection gives a theoretical lower bound of the ECM. Next, we will further elucidate the relationship between the quantization error Δ and filter performance, and investigate the stability of the proposed filter. For convenience, we consider a time-invariant system with fixed parameters $\mathbf{A}_n = \mathbf{A}$ and $\mathbf{C}_n = \mathbf{C}$.

It is easy to verify that the prediction ECM in (12) and (13) can be rewritten as follows:

$$\begin{aligned}\mathbf{P}_{n+1|n} &= \mathbf{A}_n\mathbf{P}_{n|n-1}\mathbf{A}_n^\top + \Sigma_{\mathbf{w}}\\ &\quad- \mathbf{A}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^\top(\mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-\frac{1}{2}}\\ &\quad\times \mathbf{G}_n(\mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-\frac{1}{2}}\mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{A}_n^\top.\end{aligned}$$

We define the modified algebraic Riccati equation for the BQKF as follows:

$$\begin{aligned}r(\mathbf{X}) &= \mathbf{A}_n\mathbf{X}\mathbf{A}_n^\top + \Sigma_{\mathbf{w}} - \mathbf{A}_n\mathbf{X}\mathbf{C}_n^\top(\mathbf{C}_n\mathbf{X}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-\frac{1}{2}}\\ &\quad\times \mathbf{G}_n(\mathbf{C}_n\mathbf{X}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-\frac{1}{2}}\mathbf{C}_n\mathbf{X}\mathbf{A}_n^\top.\end{aligned}$$

Since the values of β_i are finite, it is not difficult to obtain the minimum value $\beta_{\mathrm{m}} \triangleq \min\{\beta_i:\ i=1,2,\ldots,2^L\}$ satisfying

$$\mathbf{G}_n \succeq \beta_{\mathrm{m}}\mathbf{I}.$$

Noting that both $\mathbf{G}_n$ and $\beta_{\mathrm{m}}\mathbf{I}$ are diagonal matrices, the following inequality holds:

$$\begin{aligned}\mathbf{P}_{n+1|n} &\preceq \mathbf{A}_n\mathbf{P}_{n|n-1}\mathbf{A}_n^\top + \Sigma_{\mathbf{w}} - \beta_{\mathrm{m}}\mathbf{A}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^\top\\ &\quad\times(\mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-1}\mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{A}_n^\top.\end{aligned}$$

We define a new algebraic Riccati equation

$$\begin{aligned}r_\beta(\mathbf{X}) &= \mathbf{A}_n\mathbf{X}\mathbf{A}_n^\top + \Sigma_{\mathbf{w}}\\ &\quad- \beta\mathbf{A}_n\mathbf{X}\mathbf{C}_n^\top(\mathbf{C}_n\mathbf{X}\mathbf{C}_n^\top+\Sigma_{\mathbf{v}})^{-1}\mathbf{C}_n\mathbf{X}\mathbf{A}_n^\top\end{aligned}$$

where $0 \le \beta \le 1$, and a new covariance sequence $\mathbf{P}'_{n+1|n} = r_{\beta_m}(\mathbf{P}'_{n|n-1})$. In fact, we have

$$\mathbb{E}\{r(\mathbf{P}_{n|n-1})\} \preceq \mathbb{E}\{r_{\beta_m}(\mathbf{P}'_{n|n-1})\}$$

If the sequence $\mathbb{E}\{\mathbf{P}'_{n+1|n}\}$ is bounded, the considered sequence $\mathbb{E}\{\mathbf{P}_{n+1|n}\}$ is also bounded as $\mathbf{G}_n \succeq \beta_{\mathrm{m}}\mathbf{I}$. According to Theorems 2-4 in [40], we have the following theorem.

Theorem 2. *If $(\mathbf{A}, \Sigma_{\mathbf{w}}^{\frac{1}{2}})$ is controllable, $(\mathbf{A},\mathbf{C})$ is observable, and $\mathbf{A}$ is unstable, then there exists a $\beta^* \in [0,1)$ such that for each β satisfying $\beta^* < \beta \le 1$, the following inequality holds:*

$$\lim_{n\to\infty}\mathbb{E}\{\mathbf{P}_{n|n-1}\} \preceq \bar{\mathbf{M}},$$

where $\bar{\mathbf{M}}$ is the solution of the algebraic equation $\bar{\mathbf{M}} = r_\beta(\bar{\mathbf{M}})$, and β^* satisfies

$$\underline{\beta} \le \beta^* \le \bar{\beta}.$$

The upper and lower bounds are

$$\begin{aligned}\underline{\beta} &= 1-\frac{1}{\delta^2},\\ \bar{\beta} &= \arg\inf_{\beta}\left[\exists(\mathbf{K},\mathbf{X}) \mid \mathbf{X}\succeq\phi(\mathbf{K},\mathbf{X})\right],\end{aligned}$$

respectively, where $\delta = \max_i|\sigma_i|$ and σ_i are the eigenvalues of $\mathbf{A}$. The operator $\phi(\mathbf{K},\mathbf{X}) = (1-\beta)(\mathbf{A}\mathbf{X}\mathbf{A}^\top+\Sigma_{\mathbf{w}}) + \beta(\mathbf{F}\mathbf{X}\mathbf{F}^\top+\mathbf{V})$, where $\mathbf{F}=\mathbf{A}+\mathbf{K}\mathbf{C}$ and $\mathbf{V}=\mathbf{K}\Sigma_{\mathbf{v}}\mathbf{K}^\top+\Sigma_{\mathbf{w}}$.

According to this theorem, we have the following corollary.

Corollary 3. *If $(\mathbf{A}, \Sigma_{\mathbf{w}}^{\frac{1}{2}})$ is controllable, $(\mathbf{A},\mathbf{C})$ is observable, $\mathbf{A}$ is unstable, and the minimum value $\beta_{\mathrm{m}} = \min\{\beta_i:\ i=1,2,\ldots,2^L\}$ satisfies*

$$\beta_{\mathrm{m}} > \bar{\beta},$$

then $\lim_{n\to\infty}\mathbb{E}\{\mathbf{P}_{n|n-1}\} \preceq \bar{\mathbf{M}}_1$ where $\bar{\mathbf{M}}_1$ is the solution of the algebraic equation $\bar{\mathbf{M}}_1 = r_{\beta_m}(\bar{\mathbf{M}}_1)$.

In light of the above corollary, we can see that as long as the minimum value of the coefficients $\{\beta_i\}_{i=1}^{2^L}$ exceeds the threshold $\bar{\beta}$, the limit of the approximate ECM's expectation can be bounded within a certain value, thereby ensuring the feasibility of the proposed filter. Since the values of $\{\beta_i\}_{i=1}^{2^L}$ depend on the selection of the interval length W and bit length L, this theorem reveals the influence of the quantization error ($\Delta = \frac{2W}{2^L-1}$) on the filter performance.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

V. An Illustrative Example

In this section, we provide a numerical example to demonstrate the performance of the proposed BQKF.

We consider the problem of tracking a moving target in a planar space. The state vector of the target is denoted as $\mathbf{x}_n = [s_n^x, s_n^y, v_n^x, v_n^y]^{\mathrm{T}}$, where s_n^x and s_n^y represent the position of the target along the x and y axes, and v_n^x and v_n^y represent the velocities along the x and y axes, respectively. We assume a constant velocity scenario and define the state-space equation as follows:

$$\mathbf{x}_{n+1} = \begin{bmatrix} 1 & 0 & \mathrm{d}t & 0 \\ 0 & 1 & 0 & \mathrm{d}t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_n + \begin{bmatrix} \frac{1}{2}\mathrm{d}t^2 & 0 \\ 0 & \frac{1}{2}\mathrm{d}t^2 \\ \mathrm{d}t & 0 \\ 0 & \mathrm{d}t \end{bmatrix} \mathbf{w}_n$$

where $\mathrm{d}t = 1$ denotes the sampling period. The sensor is capable of measuring the target's position along the x and y axes. Therefore, the measurement matrix can be expressed as:

$$\mathbf{C}_n = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

The process and measurement noises, i.e., $\mathbf{w}_n$ and $\mathbf{v}_n$, are zero-mean Gaussian white noise sequences with covariances

$$\Sigma_{\mathbf{w}} = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix}, \quad \Sigma_{\mathbf{v}} = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix}$$

The initial value $\mathbf{x}_0$ of the state obeys Gaussian distribution with mean $\mu_0 = [100, 2, 200, 20]^{\mathrm{T}}$ and covariance $\Sigma_0 = \mathbf{I}$. We choose the range $W = 2$ and the bit number $L = 2/3/4$, and therefore the length of the interval is $\Delta = 1.33/0.57/0.26$. The flipping probability p is assumed to be 0.01.

The simulation results are illustrated in Figs. 3-7. Fig. 3 shows the trajectories of the target plant as well as the corresponding estimates obtained using the BQKF ($L = 3$). To evaluate the performance of the proposed estimator, we conduct $T_{\mathrm{MC}} = 1000$ independent simulations. Define the empirical mean-square error (MSE) as follows:

$$\mathrm{EMSE}_n = \frac{1}{T_{\mathrm{MC}}} \sum_{t=1}^{T_{\mathrm{MC}}} \| \hat{\mathbf{x}}_n^{(t)} - \mathbf{x}_n^{(t)} \|^2 \tag{25}$$

where the superscript "(t)" implies that the value is obtained in the t-th run. Moreover, the analytical MSE is obtained from the trace of the ECM defined as

$$\mathrm{AMSE}_n = \frac{1}{T_{\mathrm{MC}}} \sum_{t=1}^{T_{\mathrm{MC}}} \mathrm{tr}\left\{ \mathbf{P}_n^{(t)} \right\} \tag{26}$$

When the quantization bit length is set to $L = 3$, A comparison of the empirical MSE, the analytical MSE and the trace of the PCRLB is shown in Fig. 4. It can be observed that they exhibit the same trend over time, and eventually, the empirical MSE becomes nearly equal to the analytical MSE. The PCRLB consistently serves as a lower bound of both the empirical and analytical MSEs. Furthermore, we compare the empirical and analytical MSEs of the conventional KF and the proposed BQKF with different bit stream lengths ($L = 2, 3, 4$) in Figs. 5-6. It is evident that as the length of the bit stream increases, both the empirical and analytical MSEs decrease and approach the MSE of the KF. Moreover, Fig. 7 illustrates the variation of the PCRLB for the proposed estimator as the quantization bit length changes from 2 to 4. The PCRLB decreases sequentially and remains lower bounded by the analytical MSE of the KF. Furthermore, when the quantization bit length is set to $L = 4$, the performance of the BQKF is comparable to that of the KF.

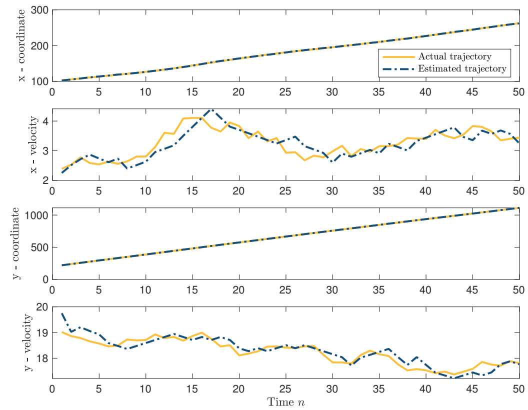


Fig. 3. The trajectories of the target plant and the corresponding estimates using KF and BQKF ($L = 3$).

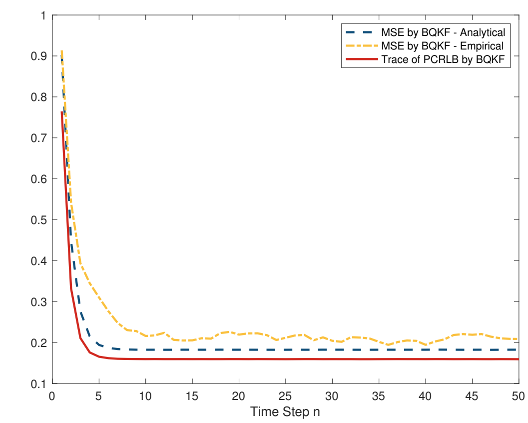


Fig. 4. The empirical MSE, analytical MSE, and the trace of the PCRLB for the proposed BQKF ($L = 3$).

VI. Conclusions

In this paper, we have investigated the problem of remote estimation for a class of linear discrete-time systems under binary encoding schemes (BESs). BESs involve quantizing the information from the remote sensor into a bit string in a probabilistic manner and transmitting it through noisy channels with random bit flipping. Due to the nonlinearities arising from quantization and transmission, calculating the exact MMSE estimate is computationally burdensome. To address this issue, we have developed the BQKF using the Gaussian approximation approach, which serves as an approximate MMSE estimator. We have also discussed the computational complexity of the proposed BQKF, which is comparable to that of the conventional Kalman filter. Furthermore, we have established the posterior Cram'er-Rao lower bound to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

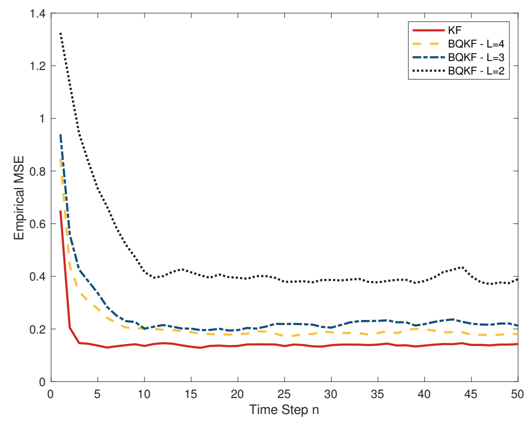


Fig. 5. The empirical MSE of KF and the empirical MSE of the proposed BQKF with different bits $L = 2, 3, 4$.

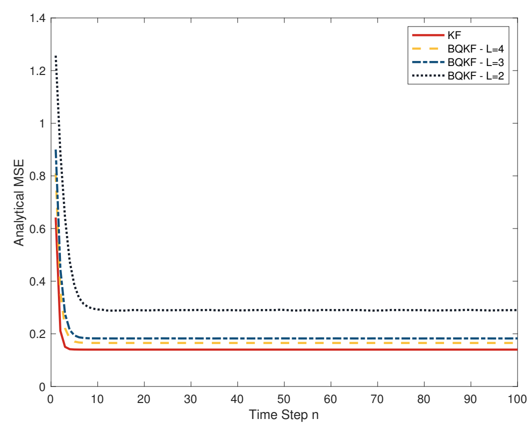


Fig. 6. The analytical MSE of KF and the analytical MSE of the proposed BQKF with different bits $L = 2, 3, 4$.

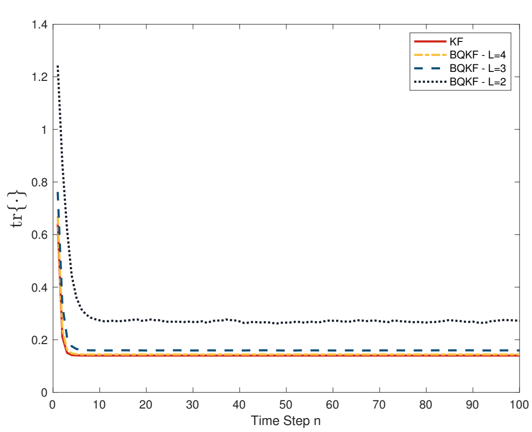


Fig. 7. The analytical MSE of KF and the trace of the PCRLB of the proposed BQKF with different bits $L = 2, 3, 4$.

assess the performance of the BQKF. Finally, numerical results have been presented to demonstrate the effectiveness of the proposed BQKF. A potential future work is to consider the state-estimation under BESs for more complicated systems, such as nonlinear systems, distributed sensor networks, or mixed analog-to-digital converter (ADC) systems with both analog and quantized data.

Appendix A
Proof of Theorem 1

Before the proof, two useful lemmas are presented.

Lemma 1. *If m events $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_n$ are mutually exclusive with each other, the following holds:*

$$\mathbb{P}\left\{\mathcal{I} \mid \cup_{i=1}^{n} \mathcal{I}_i\right\} = \sum_{i=1}^{n} \frac{\mathbb{P}\left\{\mathcal{I}_i\right\}}{\sum_{j=1}^{n} \mathbb{P}\left\{\mathcal{I}_j\right\}} \mathbb{P}\left\{\mathcal{I} \mid \mathcal{I}_i\right\}$$

Proof: It is easy to see that

$$\mathbb{P}\left\{\mathcal{I} \mid \cup_{i=1}^{n} \mathcal{I}_i\right\} = \frac{\mathbb{P}\left\{\cup_{i=1}^{n} \mathcal{I}_i \mid \mathcal{I}\right\} \mathbb{P}\left\{\mathcal{I}\right\}}{\mathbb{P}\left\{\cup_{i=1}^{n} \mathcal{I}_i\right\}}$$

According to the mutual exclusivity between events $\{\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_n\}$, we have $\mathbb{P}\{\cup_{i=1}^{n} \mathcal{I}_i\} = \sum_{i=1}^{n} \mathbb{P}\{\mathcal{I}_i\}$ and $\mathbb{P}\{\cup_{i=1}^{n} \mathcal{I}_i | \mathcal{I}\} = \sum_{i=1}^{n} \mathbb{P}\{\mathcal{I}_i | \mathcal{I}\}$. Therefore,

$$\begin{aligned}
\mathbb{P}\left\{\mathcal{I} \mid \cup_{i=1}^{n} \mathcal{I}_i\right\} &= \sum_{i=1}^{n} \frac{\mathbb{P}\left\{\mathcal{I}_i \mid \mathcal{I}\right\} \mathbb{P}\left\{\mathcal{I}\right\}}{\sum_{j=1}^{n} \mathbb{P}\left\{\mathcal{I}_j\right\}} \\
&= \sum_{i=1}^{n} \frac{\mathbb{P}\left\{\mathcal{I}_i\right\}}{\sum_{j=1}^{n} \mathbb{P}\left\{\mathcal{I}_j\right\}} \frac{\mathbb{P}\left\{\mathcal{I}_i \mid \mathcal{I}\right\} \mathbb{P}\left\{\mathcal{I}\right\}}{\mathbb{P}\left\{\mathcal{I}_i\right\}} \\
&= \sum_{i=1}^{n} \frac{\mathbb{P}\left\{\mathcal{I}_i\right\}}{\sum_{j=1}^{n} \mathbb{P}\left\{\mathcal{I}_j\right\}} \mathbb{P}\left\{\mathcal{I} \mid \mathcal{I}_i\right\}
\end{aligned}$$

■

Lemma 2 ([8])**.** *If ζ_1, ζ_2 are σ-algebras with $\zeta_1 \subset \zeta_2$ and $\mathbb{E}\{|\chi|\} < \infty$, then*

$$\mathbb{E}\left\{\mathcal{I} \mid \zeta_1\right\} = \mathbb{E}\left\{\mathbb{E}\left\{\mathcal{I} \mid \zeta_2\right\} \mid \zeta_1\right\} = \mathbb{E}\left\{\mathbb{E}\left\{\mathcal{I} \mid \zeta_1\right\} \mid \zeta_2\right\}$$

Now, we are in position to derive the approximate MMSE estimate $\hat{\mathbf{x}}_n = \mathbb{E}\{\mathbf{x}_n | \mathcal{M}_n\}$. According to Lemma 2, it is not difficult to see that

$$\begin{aligned}
\hat{\mathbf{x}}_n &= \mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_n\right\} = \mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}, \mathbf{m}_n\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\right\} \mid \mathcal{M}_{n-1}, \mathbf{m}_n\right\}
\end{aligned} \tag{27}$$

To begin with, we will evaluate the conditional expectation $\mathbb{E}\{\mathbf{x}_n | \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\}$ in (27). Given $p(\mathbf{x}_{n-1} | \mathcal{M}_{n-1}) \sim \mathcal{N}(\mathbf{x}_n; \hat{\mathbf{x}}_{n-1}, \mathbf{P}_{n-1})$, the joint distribution of random variables $\mathbf{x}_n$ and $\tilde{\mathbf{y}}_n$ conditioned on $\mathcal{M}_{n-1}$ is as follows:

$$\begin{aligned}
p(\mathbf{x}_n, \tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}) \sim \mathcal{N}\Bigg(&\mathbf{x}_n, \tilde{\mathbf{y}}_n; \begin{bmatrix} \mathbb{E}\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\} \\ \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\} \end{bmatrix}, \\
&\begin{bmatrix} \mathrm{Cov}\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\} & \mathrm{Cov}\{\mathbf{x}_n, \tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\} \\ \mathrm{Cov}\{\tilde{\mathbf{y}}_n, \mathbf{x}_n \mid \mathcal{M}_{n-1}\} & \mathrm{Cov}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\} \end{bmatrix}\Bigg)
\end{aligned}$$

The expectation and the covariance of $p(\mathbf{x}_n | \mathcal{M}_{n-1})$ can be calculated by

$$\begin{aligned}
\hat{\mathbf{x}}_{n|n-1} &\triangleq \mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\right\} \\
&= \mathbb{E}\left\{\mathbf{A}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_{n-1} \mid \mathcal{M}_{n-1}\right\} \\
&= \mathbf{A}_{n-1}\hat{\mathbf{x}}_{n-1}
\end{aligned}$$

and

$$\begin{aligned}\mathbf{P}_{n|n-1} &\triangleq \mathrm{Cov}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\right\}\\ &= \mathrm{Cov}\left\{\mathbf{A}_{n-1}\mathbf{x}_{n-1} + \mathbf{w}_{n-1} \mid \mathcal{M}_{n-1}\right\}\\ &= \mathbf{A}_{n-1}\mathbf{P}_{n-1}\mathbf{A}_{n-1}^{\mathrm{T}} + \Sigma_{\mathbf{w}}\end{aligned}$$

For the PDF $p(\tilde{\mathbf{y}}_n|\mathcal{M}_{n-1})$, its expectation can be calculated by

$$\mathbb{E}\left\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\} = \mathbf{0}$$

Moreover, since $\mathrm{Cov}\{\mathbf{y}_n|\mathcal{M}_{n-1}\} = \mathbf{C}_n\mathbf{P}_{n|n-1}\mathbf{C}_n^{\mathrm{T}} + \Sigma_{\mathbf{v}}$, one can derive that

$$\mathrm{Cov}\left\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\} = \mathbf{I}$$

In addition, one has

$$\begin{aligned}&\mathrm{Cov}\left\{\mathbf{x}_n, \tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\}\\ &= \mathbb{E}\left\{(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})[\mathbf{S}_{n|n-1}^{-\frac{1}{2}}(\mathbf{y}_n - \hat{\mathbf{y}}_{n|n-1})]^{\mathrm{T}} \mid \mathcal{M}_{n-1}\right\}\\ &= \mathbb{E}\left\{(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})^{\mathrm{T}} \mid \mathcal{M}_{n-1}\right\}\mathbf{C}_n^{\mathrm{T}}\mathbf{S}_{n|n-1}^{-\frac{1}{2}}\\ &\quad + \mathbb{E}\left\{(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1})\mathbf{v}_n^{\mathrm{T}} \mid \mathcal{M}_{n-1}\right\}\mathbf{S}_{n|n-1}^{-\frac{1}{2}}\\ &= \mathbf{P}_{n|n-1}\mathbf{C}_n^{\mathrm{T}}\mathbf{S}_{n|n-1}^{-\frac{1}{2}}\end{aligned}$$

Noticing that the covariance matrices are symmetric, one has $\mathrm{Cov}\{\tilde{\mathbf{y}}_n, \mathbf{x}_n|\mathcal{M}_{n-1}\} = \mathbf{S}_{n|n-1}^{-\frac{1}{2}}\mathbf{C}_n\mathbf{P}_{n|n-1}$. Therefore, $p(\mathbf{x}_n, \tilde{\mathbf{y}}_n|\mathcal{M}_{n-1})$ obeys the following distribution

$$\mathcal{N}\left(\mathbf{x}_n, \tilde{\mathbf{y}}_n; \begin{bmatrix}\hat{\mathbf{x}}_{n|n-1}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{P}_{n|n-1} & \mathbf{K}_n\\ \mathbf{K}_n^{\mathrm{T}} & \mathbf{I}\end{bmatrix}\right) \tag{28}$$

where $\mathbf{K}_n^{\mathrm{T}} = \mathbf{S}_{n|n-1}^{-\frac{1}{2}}\mathbf{C}_n\mathbf{P}_{n|n-1}$. As a consequence, we have

$$\begin{aligned}\hat{\mathbf{x}}_n^o &\triangleq \mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\right\}\\ &= \mathbb{E}\left\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\right\} + \mathrm{Cov}\left\{\mathbf{x}_n, \tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\}\\ &\quad \times \mathrm{Cov}\left\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\}^{-1}\left(\tilde{\mathbf{y}}_n - \mathbb{E}\left\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\right\}\right)\\ &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n\tilde{\mathbf{y}}_n\end{aligned} \tag{29}$$

From (27), we know that $\hat{\mathbf{x}}_n$ can be derived by evaluating the expectation $\mathbb{E}\{\hat{\mathbf{x}}_n^o|\mathcal{M}_{n-1}, \mathbf{m}_n\}$ as follows:

$$\begin{aligned}\hat{\mathbf{x}}_n &= \mathbb{E}\left\{\hat{\mathbf{x}}_n^o \mid \mathcal{M}_{n-1}, \mathbf{m}_n\right\}\\ &= \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n\mathbb{E}\left\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}, \mathbf{m}_n\right\}\end{aligned} \tag{30}$$

Next, we proceed to compute $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, m_{n,k}\}$. Let $\mathcal{F}_j^i$ denote the event where the channel input $\mathcal{Q}_k(\tilde{y}_{n,k}) = \tau_j$ (abbreviated as $\mathcal{Q}_{n,k} = \tau_j$ in the following content) is changed to $m_{n,k} = \tau_i$ after transmission due to the bit flipping. To be specific, the received signal $m_{n,k} = \tau_i$ can be generated from 2^L different cases as follows:

$$\left\{\mathcal{F}_1^i, \mathcal{F}_2^i, \ldots, \mathcal{F}_{2^L}^i\right\} \tag{31}$$

According to Lemma 1, we have

$$\begin{aligned}&\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, m_{n,k} = \tau_i\right\}\\ &= \sum_{j=1}^{2^L}\frac{\mathbb{P}\{\mathcal{F}_j^i \mid \mathcal{M}_{n-1}\}}{\sum_{s=1}^{2^L}\mathbb{P}\{\mathcal{F}_s^i \mid \mathcal{M}_{n-1}\}}\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = \tau_j\right\}\end{aligned} \tag{32}$$

Since $\mathcal{M}_{n-1}$ is redundant to $\mathcal{F}_j^i$, it is clear to see that

$$\mathbb{P}\{\mathcal{F}_j^i|\mathcal{M}_{n-1}\} = \mathbb{P}\{\mathcal{F}_j^i\}.$$

Furthermore, we call one bit crossover as a Bernoulli trial. Since $\mathcal{F}_j^i$ can be expressed into a bit string with length L and among each bit the crossover is independent, we can regard $\mathcal{F}_j^i$ as L-Bernoulli trials. Suppose that

$$\tau_i = -W + \sum_{l=1}^{L}\hbar_{n,k}^{(l)}2^{l-1}\Delta$$

and

$$\tau_j = -W + \sum_{l=1}^{L}h_{n,k}^{(l)}2^{l-1}\Delta.$$

with $\hbar_{n,k}^{(l)} = \theta_{j\to i}^{(l)}(1 - h_{n,k}^{(l)}) + (1 - \theta_{j\to i}^{(l)})h_{n,k}^{(l)}$, where $\theta_{j\to i}^{(l)} = 1$ if the lth bit of τ_i is flipped and 0 otherwise. As a result, the probability of the event $\mathcal{F}_j^i$ is as follows:

$$\mathbb{P}\left\{\mathcal{F}_j^i\right\} = p^{\theta_{ij}}(1-p)^{L-\theta_{ij}}, \tag{33}$$

with $\theta_{ij} = \sum_{l=1}^{L}\theta_{j\to i}^{(l)}$. Apparently, the sum of all the probabilities in L-Bernoulli trials is equal to one, i.e., $\sum_{l=1}^{2^L}\mathbb{P}\{\mathcal{F}_l^i\} = 1$. Substituting (33) into (32) yields

$$\begin{aligned}&\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, m_{n,k} = \tau_i\right\}\\ &= \sum_{j=1}^{2^L}p^{\theta_{ij}}(1-p)^{L-\theta_{ij}}\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = \tau_j\right\}\end{aligned} \tag{34}$$

Moreover, the information set $\{\mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = \tau_i\}$ is equal to the union of the sets $\{\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i\}$ and $\{\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i\}$, for $i = 2, \ldots, 2^L - 1$, where $\mathcal{U}_i$ denotes the interval $[\tau_i, \tau_{i+1}]$ and $\mathcal{L}_n^i$ denotes the case that $\tilde{y}_{n,k} \in \mathcal{U}_i$ is quantized to τ_i (the left boundary of $\mathcal{U}_i$), and correspondingly $\mathcal{R}_n^i$ denotes the case that $\tilde{y}_{n,k} \in \mathcal{U}_{i-1}$ is quantized to τ_i (the right boundary of $\mathcal{U}_{i-1}$). According to Lemma 1, we have

$$\begin{aligned}&\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = i\right\}\\ &= \frac{p_{\mathcal{R}_n^i}}{p_{\mathcal{L}_n^i} + p_{\mathcal{R}_n^i}}\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i\right\}\\ &\quad + \frac{p_{\mathcal{L}_n^i}}{p_{\mathcal{L}_n^i} + p_{\mathcal{R}_n^i}}\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i\right\}\end{aligned} \tag{35}$$

where $p_{\mathcal{R}_n^i}$ is the shorthand of $\mathbb{P}\{\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i\}$, and $p_{\mathcal{L}_n^i}$ is the shorthand of $\mathbb{P}\{\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i\}$.

In the following calculations, we will determine the probabilities $p_{\mathcal{L}_n^i}$ and $p_{\mathcal{R}_n^i}$. Based on Bayes' theorem, $p_{\mathcal{L}_n^i}$ can be decomposed into the following three components:

$$\begin{aligned}p_{\mathcal{L}_n^i} &= \mathbb{P}\left\{\mathcal{L}_n^i \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\right\}\\ &\quad \times \mathbb{P}\left\{\tilde{y}_{n,k} \in \mathcal{U}_i \mid \mathcal{M}_{n-1}\right\}\mathbb{P}\left\{\mathcal{M}_{n-1}\right\}\end{aligned}$$

Noting that $p(\tilde{y}_{n,k}|\mathcal{M}_{n-1}) \sim \mathcal{N}(\tilde{y}_{n,k}; 0, 1)$, it is easy to see

$$\mathbb{P}\left\{\tilde{y}_{n,k} \in \mathcal{U}_i \mid \mathcal{M}_{n-1}\right\} = \varphi(\tau_{i+1}) - \varphi(\tau_i).$$

The conditional probability of $\mathcal{L}_n^i$ is

$$\begin{aligned}\mathbb{P}\left\{\mathcal{L}_n^i \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\right\} &= \int_{\mathbb{R}}p(\tilde{y}_{n,k} \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1})\\ &\quad \times \mathbb{P}\left\{\mathcal{L}_n^i \mid \tilde{y}_{n,k}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\right\}\,\mathrm{d}\tilde{y}_{n,k}\end{aligned}$$

Given the fact that $\tilde{y}_{n,k} \in \mathcal{U}_i$, $\mathcal{M}_{n-1}$ is redundant, we have

$$\mathbb{P}\{\mathcal{L}_n^i \mid \tilde{y}_{n,k}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\} = (\tau_{i+1} - \tilde{y}_{n,k})/\Delta$$

Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

Moreover, one has

$$
\begin{aligned}
&p(\tilde{y}_{n,k} \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}) \\
&= \begin{cases} \frac{\mathcal{N}(\tilde{y}_{n,k};0,1)}{\varphi(\tau_{i+1})-\varphi(\tau_i)}, & \text{if } \tilde{y}_{n,k} \in \mathcal{U}_i \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

As a consequence, it can be derived that

$$
\begin{aligned}
&\mathbb{P}\{\mathcal{L}_n^i | \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\} \\
&= \int_{\mathcal{U}_i} \frac{\mathcal{N}(\tilde{y}_{n,k};0,1)}{\varphi(\tau_{i+1})-\varphi(\tau_i)} \frac{\tau_{i+1}-\tilde{y}_{n,k}}{\Delta} \, \mathrm{d}\tilde{y}_{n,k}
\end{aligned}
$$

For brevity of presentation, denote

$$
\begin{aligned}
\bar{p}_{\mathcal{L}_n^i} &\triangleq \mathbb{P}\{\tilde{y}_{n,k} \in \mathcal{U}_i \mid \mathcal{M}_{n-1}\} \mathbb{P}\left\{\mathcal{L}_n^i \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{M}_{n-1}\right\} \\
&= \int_{\mathcal{U}_i} \frac{\tau_{i+1}-\tilde{y}_{n,k}}{\Delta} \mathcal{N}(\tilde{y}_{n,k};0,1) \, \mathrm{d}\tilde{y}_{n,k} \\
&= \frac{\exp(-\frac{\tau_{i+1}^2}{2}) - \exp(-\frac{\tau_i^2}{2})}{\sqrt{2\pi}\Delta} + \frac{\tau_{i+1}[\varphi(\tau_{i+1})-\varphi(\tau_i)]}{\Delta}
\end{aligned}
$$

Therefore, the probability $p_{\mathcal{L}_n^i}$ can be calculated as

$$
p_{\mathcal{L}_n^i} = \bar{p}_{\mathcal{L}_n^i} \mathbb{P}\{\mathcal{M}_{n-1}\}
$$

Similarly, we have

$$
\bar{p}_{\mathcal{R}_n^i} \triangleq \frac{\exp(-\frac{\tau_{i-1}^2}{2}) - \exp(-\frac{\tau_i^2}{2})}{\sqrt{2\pi}\Delta} + \frac{\tau_{i-1}[\varphi(\tau_{i-1})-\varphi(\tau_i)]}{\Delta}
$$

and

$$
p_{\mathcal{R}_n^i} = \bar{p}_{\mathcal{R}_n^i} \mathbb{P}\{\mathcal{M}_{n-1}\}
$$

Apparently, $\bar{p}_{\mathcal{L}_n^i} + \bar{p}_{\mathcal{R}_n^i}$ can be simplified as $H(\tau_i)/\Delta$, where

$$
H(\tau) = h(\tau-\Delta) + h(\tau+\Delta) - 2h(\tau)
$$

and

$$
h(\tau) = \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{\tau^2}{2}) + \tau\varphi(\tau) \tag{36}
$$

Now, let us calculate the conditional expectations $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i\}$ and $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i\}$. It can be derived that

$$
\begin{aligned}
&p(\tilde{y}_{n,k} \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i, \mathcal{M}_{n-1}) \\
&= \begin{cases} \frac{p(\tilde{y}_{n,k}|\mathcal{M}_{n-1})\mathbb{P}\{\mathcal{L}_n^i|\tilde{y}_{n,k}\}}{\mathbb{P}\{\tilde{y}_{n,k}\in\mathcal{U}_i, \mathcal{L}_n^i|\mathcal{M}_{n-1}\}} & , \text{ if } \tilde{y}_{n,k} \in \mathcal{U}_i \\ 0 & , \text{ otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{\bar{p}_{\mathcal{L}_n^i}} \frac{\tau_{i+1}-\tilde{y}_{n,k}}{\Delta} \mathcal{N}(\tilde{y}_{n,k};0,1) & , \text{ if } \tilde{y}_{n,k} \in \mathcal{U}_i \\ 0 & , \text{ otherwise} \end{cases}
\end{aligned}
$$

which yields

$$
\begin{aligned}
&\mathbb{E}\left\{\tilde{y}_{n,k} \mid \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i, \mathcal{M}_{n-1}\right\} \\
&= \frac{1}{\bar{p}_{\mathcal{L}_n^i}\Delta} \int_{\mathcal{U}_i} \tilde{y}_{n,k}(\tau_{i+1}-\tilde{y}_{n,k}) \mathcal{N}(\tilde{y}_{n,k};0,1) \, \mathrm{d}\tilde{y}_{n,k} \\
&= \frac{1}{\bar{p}_{\mathcal{L}_n^i}\Delta} \left[\varphi(\tau_i) - \varphi(\tau_{i+1}) + \frac{\Delta}{\sqrt{2\pi}} \exp(-\tfrac{\tau_i^2}{2})\right]
\end{aligned} \tag{37}
$$

Similarly, we have

$$
\begin{aligned}
&\mathbb{E}\left\{\tilde{y}_{n,k} \mid \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i, \mathcal{M}_{n-1}\right\} \\
&= \frac{1}{\bar{p}_{\mathcal{R}_n^i}\Delta} \left[\varphi(\tau_i) - \varphi(\tau_{i-1}) - \frac{\Delta}{\sqrt{2\pi}} \exp(-\tfrac{\tau_i^2}{2})\right]
\end{aligned} \tag{38}
$$

Substituting (37) and (38) into (35) yields that

$$
\mathbb{E}\left\{\tilde{y}_{n,k} \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = i\right\} = \frac{F(\tau_i)}{H(\tau_i)} \tag{39}
$$

where $F(\tau) = f(\tau-\Delta) + f(\tau+\Delta) - 2f(\tau)$ and $f(\tau) = -\varphi(\tau)$. In particular, $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, \mathcal{Q}_{n,k} = i\}$ is reduced to $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_i, \mathcal{L}_n^i\}$ for $i = 1$ and $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, \tilde{y}_{n,k} \in \mathcal{U}_{i-1}, \mathcal{R}_n^i\}$ for $i = 2^L$. Consequently, one has

$$
F(\tau) = \begin{cases} \varphi(\tau) - \varphi(\tau+\Delta) + \frac{\Delta}{\sqrt{2\pi}} \exp(-\frac{\tau^2}{2}), & \tau = -W \\ \varphi(\tau) - \varphi(\tau-\Delta) - \frac{\Delta}{\sqrt{2\pi}} \exp(-\frac{\tau^2}{2}), & \tau = W \end{cases}
$$

and

$$
H(\tau) = \begin{cases} \frac{1}{\sqrt{2\pi}}[\exp(-\frac{(\tau+\Delta)^2}{2}) - \exp(-\frac{\tau^2}{2})] & \\ \quad +(\tau+\Delta)[\varphi(\tau+\Delta) - \varphi(\tau)], & \tau = -W \\ \frac{1}{\sqrt{2\pi}}[\exp(-\frac{(\tau-\Delta)^2}{2}) - \exp(-\frac{\tau^2}{2})] & \\ \quad +(\tau-\Delta)[\varphi(\tau-\Delta) - \varphi(\tau)], & \tau = W \end{cases}
$$

Denote $\alpha_i \triangleq \mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_{n-1}, m_{n,k} = \tau_i\}$. When the received signal is τ_i at instant n, according to (34) and (39), we have $\alpha_i = \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}} F(\tau_j)/H(\tau_j)$, and hence (30) can be rearranged as

$$
\hat{\mathbf{x}}_n = \hat{\mathbf{x}}_{n|n-1} + \mathbf{K}_n \mathbf{g}_n
$$

Subsequently, we will derive $\mathbf{P}_n \triangleq \mathbb{E}\{(\mathbf{x}_n - \hat{\mathbf{x}}_n)(\mathbf{x}_n - \hat{\mathbf{x}}_n)^{\mathrm{T}}|\mathcal{M}_n\}$. Using (29) and (30), the dynamics of the estimation error $\mathbf{e}_n = \mathbf{x}_n - \hat{\mathbf{x}}_n$ can be written as follows:

$$
\begin{aligned}
\mathbf{e}_n &= \mathbf{x}_n - \hat{\mathbf{x}}_n^o + \hat{\mathbf{x}}_n^o - \hat{\mathbf{x}}_n \\
&= \mathbf{x}_n - \hat{\mathbf{x}}_n^o + \mathbf{K}_n(\tilde{\mathbf{y}}_n - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\})
\end{aligned} \tag{40}
$$

In light of Lemma 2, we have

$$
\begin{aligned}
\mathbf{P}_n &= \mathbb{E}\left\{\mathbf{e}_n\mathbf{e}_n^{\mathrm{T}} \mid \mathcal{M}_n\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\mathbf{e}_n\mathbf{e}_n^{\mathrm{T}} \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\right\} \mid \mathcal{M}_{n-1}, \mathbf{m}_n\right\}
\end{aligned} \tag{41}
$$

Substituting (40) into the above equation yields

$$
\begin{aligned}
&\mathbb{E}\left\{\mathbf{e}_n\mathbf{e}_n^{\mathrm{T}} \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\right\} \\
&= \mathbb{E}\left\{(\mathbf{x}_n - \hat{\mathbf{x}}_n^o)(\mathbf{x}_n - \hat{\mathbf{x}}_n^o)^{\mathrm{T}} \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\right\} \\
&+ \mathbf{K}_n(\tilde{\mathbf{y}}_n - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\})(\tilde{\mathbf{y}}_n - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\})^{\mathrm{T}}\mathbf{K}_n^{\mathrm{T}},
\end{aligned} \tag{42}
$$

which is due to the fact that both $\tilde{\mathbf{y}}_n$ and $\mathbb{E}\{\tilde{\mathbf{y}}_n|\mathcal{M}_n\}$ are deterministic functions under the condition $\{\mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\}$, and $\mathbb{E}\{\mathbf{x}_n - \hat{\mathbf{x}}_n^o|\mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\} = \mathbb{E}\{\mathbf{x}_n|\mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\} - \hat{\mathbf{x}}_n^o = \hat{\mathbf{x}}_n^o - \hat{\mathbf{x}}_n^o = \mathbf{0}$. We observe that the first term of (42) is the covariance of $p(\mathbf{x}_n|\mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n)$, which can be calculated based on $p(\mathbf{x}_n, \tilde{\mathbf{y}}_n|\mathcal{M}_{n-1})$ in (28) as follows:

$$
\begin{aligned}
\mathbf{P}_n^o &\triangleq \mathrm{Cov}\{\mathbf{x}_n \mid \mathcal{M}_{n-1}, \tilde{\mathbf{y}}_n\} = \mathrm{Cov}\{\mathbf{x}_n \mid \mathcal{M}_{n-1}\} \\
&\quad - \mathrm{Cov}\{\mathbf{x}_n, \tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\} \mathrm{Cov}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_{n-1}\}^{-1} \\
&\quad \times \mathrm{Cov}\{\tilde{\mathbf{y}}_n, \mathbf{x}_n \mid \mathcal{M}_{n-1}\} = \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{K}_n^{\mathrm{T}}
\end{aligned} \tag{43}
$$

Combining (41) and (43), we obtain

$$
\begin{aligned}
\mathbf{P}_n &= \mathbf{P}_n^o + \mathbf{K}_n\mathbb{E}\left\{(\tilde{\mathbf{y}}_n - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\}) \right. \\
&\quad \left. \times (\tilde{\mathbf{y}}_n - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\})^{\mathrm{T}} \mid \mathcal{M}_n\right\} \mathbf{K}_n^{\mathrm{T}} \\
&= \mathbf{P}_{n|n-1} - \mathbf{K}_n\mathbf{K}_n^{\mathrm{T}} + \mathbf{K}_n \mathrm{Cov}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\} \mathbf{K}_n^{\mathrm{T}} \\
&= \mathbf{P}_{n|n-1} - \mathbf{K}_n(\mathbf{I} - \mathrm{Cov}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\})\mathbf{K}_n^{\mathrm{T}}
\end{aligned}
$$

Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

where

$$\operatorname{Cov}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\} = \mathbb{E}\{\tilde{\mathbf{y}}_n\tilde{\mathbf{y}}_n^{\mathrm{T}} \mid \mathcal{M}_n\} - \mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\}\mathbb{E}\{\tilde{\mathbf{y}}_n \mid \mathcal{M}_n\}^{\mathrm{T}}$$

Since $\tilde{y}_{n,k}$ are independent and identically distributed variables obeying $p(\tilde{y}_{n,k}) = \mathcal{N}(\tilde{y}_{n,k};0,1)$, we know that the off-diagonal component of $\operatorname{Cov}\{\tilde{\mathbf{y}}_n|\mathcal{M}_n\}$ is equal to 0. As for the diagonal component, we have $\operatorname{Var}\{\tilde{y}_{n,k}|\mathcal{M}_n\} = \mathbb{E}\{\tilde{y}_{n,k}^2|\mathcal{M}_n\} - (\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_n\})^2$, where $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_n\}$ is given in (34), where $\mathbb{E}\{\tilde{y}_{n,k}^2|\mathcal{M}_n\}$ can be calculated in a similar way like $\mathbb{E}\{\tilde{y}_{n,k}|\mathcal{M}_n\}$. To be specific, $\mathbb{E}\{\tilde{y}_{n,k}^2|\mathcal{M}_n\}$ can be rearranged as

$$\begin{aligned}&\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, m_{n,k}=\tau_i\}\\ &= \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k}=\tau_j\}\end{aligned} \tag{44}$$

and $\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k}=\tau_i\}$ for $i = 2,\ldots,2^L-1$ can be written as

$$\begin{aligned}&\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, \mathcal{Q}_{n,k}=i\}\\ &= \frac{p_{\mathcal{R}_n^i}}{p_{\mathcal{L}_n^i}+p_{\mathcal{R}_n^i}}\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, \tilde{y}_{n,k}\in\mathcal{U}_{i-1}, \mathcal{R}_n^i\}\\ &\quad+\frac{p_{\mathcal{L}_n^i}}{p_{\mathcal{L}_n^i}+p_{\mathcal{R}_n^i}}\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, \tilde{y}_{n,k}\in\mathcal{U}_{i}, \mathcal{L}_n^i\}\end{aligned}$$

Finally, it can be obtained that

$$\mathbb{E}\{\tilde{y}_{n,k}^2 \mid \mathcal{M}_{n-1}, m_n=\tau_i\} = \frac{\Lambda(\tau_i)}{H(\tau_i)}$$

where $\Lambda(\tau)$ is $\lambda(\tau-\Delta)+\lambda(\tau+\Delta)-2\lambda(\tau)$ and

$$\lambda(\tau) = \frac{2}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2}) + \tau\varphi(\tau)$$

Letting $\beta_i \triangleq 1-\operatorname{Var}\{\tilde{y}_{n,k} \mid \mathcal{M}_n\}$, we thus have

$$\begin{aligned}\beta_i &= 1-\left\{\sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\frac{\Lambda(\tau)}{H(\tau)}-\alpha_i^2\right\}\\ &= \alpha_i^2-\sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\frac{\Lambda(\tau)-H(\tau)}{H(\tau)}\\ &= \alpha_i^2-\sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\frac{G(\tau)}{H(\tau)}\end{aligned}$$

where $G(\tau) = g(\tau-\Delta)+g(\tau+\Delta)-2g(\tau)$ and

$$g(\tau) = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{\tau^2}{2})$$

Moreover, for $i=1$ and $i=2^L$, we have

$$G(\tau) = \begin{cases} g(\tau+\Delta)-(1-\Delta\tau)g(\tau), & \tau=-W\\ g(\tau-\Delta)-(1+\Delta\tau)g(\tau), & \tau=W\end{cases}$$

Let $\mathbf{G}_n \triangleq \operatorname{diag}\{\beta_{n,1},\beta_{n,2},\ldots,\beta_{n,M}\}$. Therefore, it is straightforward to see that

$$\mathbf{P}_n = \mathbf{P}_{n|n-1}-\mathbf{K}_n\mathbf{G}_n\mathbf{K}_n^{\mathrm{T}} \tag{45}$$

which is the ECM in (13). The proof is now complete.

Appendix B
Proof of Proposition 2

According to [42], the recursion of the FIM $\mathbf{J}_n$ can be obtained as follows:

$$\mathbf{J}_n = \mathbf{D}_{n-1}^{22}-\mathbf{D}_{n-1}^{21}(\mathbf{J}_{n-1}+\mathbf{D}_{n-1}^{11})^{-1}\mathbf{D}_{n-1}^{12} \tag{46}$$

where

$$\begin{aligned}\mathbf{D}_{n-1}^{11} &= \mathbb{E}\left\{-\Delta_{\mathbf{x}_{n-1}}^{\mathbf{x}_{n-1}}\log p(\mathbf{x}_n \mid \mathbf{x}_{n-1})\right\}\\ \mathbf{D}_{n-1}^{12} &= \mathbb{E}\left\{-\Delta_{\mathbf{x}_{n-1}}^{\mathbf{x}_{n}}\log p(\mathbf{x}_n \mid \mathbf{x}_{n-1})\right\}\\ \mathbf{D}_{n-1}^{21} &= \mathbb{E}\left\{-\Delta_{\mathbf{x}_{n}}^{\mathbf{x}_{n-1}}\log p(\mathbf{x}_n \mid \mathbf{x}_{n-1})\right\} = \left(\mathbf{D}_{n-1}^{12}\right)^{\mathrm{T}}\\ \mathbf{D}_{n-1}^{22} &= \mathbb{E}\left\{-\Delta_{\mathbf{x}_{n}}^{\mathbf{x}_{n}}\log p(\mathbf{x}_n \mid \mathbf{x}_{n-1})\right\}\\ &\quad+\mathbb{E}\left\{-\Delta_{\mathbf{x}_{n}}^{\mathbf{x}_{n}}\log \mathbb{P}\{\mathbf{m}_n \mid \mathbf{x}_n\}\right\} = \mathbf{D}_{n-1}^{22,a}+\mathbf{D}_{n-1}^{22,b}\end{aligned}$$

and Δ_y^x denotes the operator of the second derivative, namely $\Delta_y^x = \nabla_x\nabla_y^{\mathrm{T}}$ with ∇ standing for the gradient operator.

Given the state-space model (1), it is not difficult to obtain that $\mathbf{D}_{n-1}^{11} = \mathbf{A}_{n-1}^{\mathrm{T}}\Sigma_{\mathbf{w}}^{-1}\mathbf{A}_{n-1}$, $\mathbf{D}_{n-1}^{12} = -\mathbf{A}_{n-1}^{\mathrm{T}}\Sigma_{\mathbf{w}}^{-1}$, $\mathbf{D}_{n-1}^{21} = -\Sigma_{\mathbf{w}}^{-1}\mathbf{A}_{n-1}$ and $\mathbf{D}_{n-1}^{22,a} = \Sigma_{\mathbf{w}}^{-1}$. Applying the Matrix Inversion Lemma[1] to (46), we have

$$\begin{aligned}\mathbf{J}_n &= \mathbf{D}_{n-1}^{22,b}+\Sigma_{\mathbf{w}}^{-1}-\Sigma_{\mathbf{w}}^{-1}\mathbf{A}_{n-1}\\ &\quad\times(\mathbf{J}_{n-1}+\mathbf{A}_{n-1}^{\mathrm{T}}\Sigma_{\mathbf{w}}^{-1}\mathbf{A}_{n-1})^{-1}\mathbf{A}_{n-1}^{\mathrm{T}}\Sigma_{\mathbf{w}}^{-1}\\ &= \mathbf{D}_{n-1}^{22,b}+(\Sigma_{\mathbf{w}}+\mathbf{A}_{n-1}\mathbf{J}_{n-1}^{-1}\mathbf{A}_{n-1}^{\mathrm{T}})^{-1}\end{aligned} \tag{47}$$

To evaluate $\mathbf{D}_{n-1}^{22,b}$, we need to first compute $\mathbb{P}\{\mathbf{m}_n|\mathbf{x}_n\}$ as follows:

$$\begin{aligned}\mathbb{P}\{\mathbf{m}_n \mid \mathbf{x}_n\} &= \mathbb{P}\{\cup_{k=1}^{M}\{m_{n,k}=\tau_i\} \mid \mathbf{x}_n\}\\ &= \prod_{k=1}^{M}\mathbb{P}\{m_{n,k}=\tau_i \mid \mathbf{x}_n\}\end{aligned}$$

According to the relationship between $\{m_{n,k}=\tau_i\}$ and $\cup_{j=1}^{2^L}\{\mathcal{Q}(\tilde{y}_{n,k})=\tau_j,\mathcal{F}_j^i\}$ mentioned in (31), we have

$$\begin{aligned}\mathbb{P}\{m_{n,k}=\tau_i \mid \mathbf{x}_n\} &= \left\{\cup_{j=1}^{2^L}\{\mathcal{Q}(\tilde{y}_{n,k})=\tau_j,\mathcal{F}_j^i\} \mid \mathbf{x}_n\right\}\\ &= \sum_{j=1}^{2^L}\mathbb{P}\{\mathcal{F}_j^i\}\,\mathbb{P}\{\mathcal{Q}(\tilde{y}_{n,k})=\tau_j \mid \mathbf{x}_n\}\end{aligned}$$

where $\mathbb{P}\{\mathcal{F}_j^i\} = p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}$. Based on the Bayes' theorem, the probability $\mathbb{P}\{\mathcal{Q}(\tilde{y}_n) \mid \mathbf{x}_n\}$ can be written as

$$\begin{aligned}\mathbb{P}\{\mathcal{Q}(\tilde{y}_{n,k}) \mid \mathbf{x}_n\} &= \int_{\mathbb{R}} p(\tilde{y}_{n,k},\mathcal{Q}(\tilde{y}_{n,k}) \mid \mathbf{x}_n)\,\mathrm{d}\tilde{y}_{n,k}\\ &= \int_{\mathbb{R}}\mathbb{P}\{\mathcal{Q}(\tilde{y}_{n,k}) \mid \tilde{y}_{n,k}\}\,p(\tilde{y}_{n,k} \mid \mathbf{x}_n)\,\mathrm{d}\tilde{y}_{n,k}\end{aligned} \tag{48}$$

where

$$p(\tilde{y}_{n,k}|\mathbf{x}_n) = \mathcal{N}(\tilde{y}_{n,k};\mu_{n,k},\sigma_{n,k}^2).$$

[1] $(A+UCV)^{-1} = A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

Moreover, since $\{\mathcal{Q}(\tilde{y}_{n,k}) = \tau_j\}$ is equivalent to $\{\tilde{y}_{n,k} \in \mathcal{U}_{j-1}, \mathcal{R}^j_{n,k}\} \cup \{\tilde{y}_{n,k} \in \mathcal{U}_j, \mathcal{L}^j_{n,k}\}$, we can rearrange (48) as follows:

$$\begin{aligned}
&\mathbb{P}\left\{\mathcal{Q}(\tilde{y}_{n,k}) = \tau_j \mid \mathbf{x}_n\right\} \\
&= \int_{\mathcal{U}_{j-1}} \frac{\tilde{y}_{n,k} - \tau_{j-1}}{\Delta} \mathcal{N}(\tilde{y}_{n,k}; \mu_{n,k}, \sigma^2_{n,k})\, \mathrm{d}\tilde{y}_{n,k} \\
&+ \int_{\mathcal{U}_j} \frac{\tau_{j+1} - \tilde{y}_{n,k}}{\Delta} \mathcal{N}(\tilde{y}_{n,k}; \mu_{n,k}, \sigma^2_{n,k})\, \mathrm{d}\tilde{y}_{n,k}
\end{aligned} \tag{49}$$

for $j = 2, \ldots, 2^L - 1$. Let $\psi_{n,k}(\tau; \mathbf{x}_n) \triangleq \psi(\frac{\tau-\mu_{n,k}}{\sigma_{n,k}})$ where $\psi(\tau) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{\tau^2}{2}) + \tau\varphi(\tau)$ and $\Psi_{n,k}(\tau; \mathbf{x}_n) \triangleq \psi_{n,k}(\tau - \Delta; \mathbf{x}_n) + \psi_{n,k}(\tau + \Delta; \mathbf{x}_n) - 2\psi_{n,k}(\tau; \mathbf{x}_n)$. Then, by computing (49), we obtain

$$\mathbb{P}\left\{\mathcal{Q}(\tilde{y}_{n,k}) = \tau_j \mid \mathbf{x}_n\right\} = \frac{\sigma^2_{n,k}}{\Delta}\Psi_{n,k}(\tau_j; \mathbf{x}_n)$$

In particular, $\mathbb{P}\{\mathcal{Q}(\tilde{y}_{n,k}) = \tau_j | \mathbf{x}_n\}$ is reduced to the first integral term for $j = 1$ and the second term for $j = 2^L$. Consequently, one has

$$\Psi_{n,k}(\tau; \mathbf{x}_n) = \begin{cases}
\psi_{n,k}(\tau + \Delta; \mathbf{x}_n) - \psi_{n,k}(\tau; \mathbf{x}_n) \\
-\frac{\Delta}{\sigma_{n,k}}\varphi(\frac{\tau-\mu_{n,k}}{\sigma_{n,k}}), & \tau = -W \\
\psi_{n,k}(\tau - \Delta; \mathbf{x}_n) - \psi_{n,k}(\tau; \mathbf{x}_n) \\
+\frac{\Delta}{\sigma_{n,k}}\varphi(\frac{\tau-\mu_{n,k}}{\sigma_{n,k}}), & \tau = W \\
\psi_{n,k}(\tau - \Delta; \mathbf{x}_n) + \psi_{n,k}(\tau + \Delta; \mathbf{x}_n) \\
-2\psi_{n,k}(\tau; \mathbf{x}_n), & \text{otherwise}
\end{cases}$$

Let $\Upsilon_{n,k} \triangleq \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\Psi_{n,k}(\tau_j; \mathbf{x}_n)$. With these notations, we have

$$\mathbb{P}\{m_{n,k}|\mathbf{x}_n\} = \frac{\sigma^2_{n,k}}{\Delta}\Upsilon_{n,k}.$$

According to the chain rule, we know that

$$\begin{aligned}
&-\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\log\mathbb{P}\left\{\mathbf{m}_n \mid \mathbf{x}_n\right\} \\
&= \sum_{k=1}^{M} -\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\log\mathbb{P}\left\{m_{n,k} \mid \mathbf{x}_n\right\} \\
&= \sum_{k=1}^{M}\left(\frac{\nabla_{\mathbf{x}_n}\mathbb{P}\{m_{n,k}|\mathbf{x}_n\}}{\mathbb{P}\{m_{n,k}|\mathbf{x}_n\}}\right)^2_{\mathbf{I}} - \frac{\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\mathbb{P}\{m_{n,k}|\mathbf{x}_n\}}{\mathbb{P}\{m_{n,k}|\mathbf{x}_n\}} \\
&= \sum_{k=1}^{M}\left(\frac{\nabla_{\mathbf{x}_n}\Upsilon_{n,k}}{\Upsilon_{n,k}}\right)^2_{\mathbf{I}} - \frac{\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\Upsilon_{n,k}}{\Upsilon_{n,k}}
\end{aligned} \tag{50}$$

where $(\star)^2_{\mathbf{I}} \triangleq (\star)(\star)^{\mathrm{T}}$ for any column vector $\star$. Let $\dot{\psi}_{n,k}(\tau; \mathbf{x}_n) \triangleq \dot{\psi}(\frac{\tau-\mu_{n,k}}{\sigma_{n,k}})$ where $\dot{\psi}(\tau) = \varphi(\tau)$ and

$$\dot{\Psi}_{n,k}(\tau; \mathbf{x}_n) \triangleq \begin{cases}
\dot{\psi}_{n,k}(\tau + \Delta; \mathbf{x}_n) - \dot{\psi}_{n,k}(\tau; \mathbf{x}_n) \\
\quad -\frac{\Delta}{\sigma_{n,k}}\frac{1}{\sqrt{2\pi}}\exp(-\frac{(\tau-\mu_{n,k})^2}{2\sigma^2_{n,k}}), & \tau = -W \\
\dot{\psi}_{n,k}(\tau - \Delta; \mathbf{x}_n) - \dot{\psi}_{n,k}(\tau; \mathbf{x}_n) \\
\quad +\frac{\Delta}{\sigma_{n,k}}\frac{1}{\sqrt{2\pi}}\exp(-\frac{(\tau-\mu_{n,k})^2}{2\sigma^2_{n,k}}), & \tau = W \\
\dot{\psi}_{n,k}(\tau - \Delta; \mathbf{x}_n) + \dot{\psi}_{n,k}(\tau + \Delta; \mathbf{x}_n) \\
\quad -2\dot{\psi}_{n,k}(\tau; \mathbf{x}_n), & \text{otherwise}
\end{cases}$$

Denoting

$$\dot{\Upsilon}_{n,k} \triangleq \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\dot{\Psi}_{n,k}(\tau_j; \mathbf{x}_n),$$

the first-order derivative of $\Upsilon_{n,k}$ can be written as

$$\begin{aligned}
\nabla_{\mathbf{x}_n}\Upsilon_{n,k} &= \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\nabla_{\mathbf{x}_n}\Psi_{n,k}(\tau_j; \mathbf{x}_n) \\
&= -\sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}(\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\frac{\dot{\Psi}_{n,k}(\tau_j; \mathbf{x}_n)}{\sigma_v} \\
&= -(\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\dot{\Upsilon}_{n,k}\sigma_v^{-1}
\end{aligned}$$

and

$$\left(\frac{\nabla_{\mathbf{x}_n}\Upsilon_{n,k}}{\Upsilon_{n,k}}\right)^2_{\mathbf{I}} = (\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\left(\frac{\dot{\Upsilon}_{n,k}}{\Upsilon_{n,k}}\right)^2\sigma_v^{-2}\mathbf{C}^{(k,:)}_n \tag{51}$$

Let $\ddot{\psi}_{n,k}(\tau; \mathbf{x}_n) \triangleq \ddot{\psi}(\frac{\tau-\mu_{n,k}}{\sigma_{n,k}})$ where $\ddot{\psi}(\tau) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{\tau^2}{2})$ and

$$\ddot{\Psi}_{n,k}(\tau; \mathbf{x}_n) \triangleq \begin{cases}
\ddot{\psi}_{n,k}(\tau + \Delta; \mathbf{x}_n) - \ddot{\psi}_{n,k}(\tau; \mathbf{x}_n) \\
\quad +\frac{\Delta}{\sigma_{n,k}}\frac{\tau-\mu_{n,k}}{\sigma_{n,k}}\ddot{\psi}_{n,k}(\tau; \mathbf{x}), & \tau = -W \\
\ddot{\psi}_{n,k}(\tau - \Delta; \mathbf{x}_n) - \ddot{\psi}_{n,k}(\tau; \mathbf{x}_n) \\
\quad -\frac{\Delta}{\sigma_{n,k}}\frac{\tau-\mu_{n,k}}{\sigma_{n,k}}\ddot{\psi}_{n,k}(\tau; \mathbf{x}), & \tau = W \\
\ddot{\psi}_{n,k}(\tau - \Delta; \mathbf{x}_n) + \ddot{\psi}_{n,k}(\tau + \Delta; \mathbf{x}_n) \\
\quad -2\ddot{\psi}_{n,k}(\tau; \mathbf{x}_n), & \text{otherwise}
\end{cases}$$

Denoting

$$\ddot{\Upsilon}_{n,k} \triangleq \sum_{j=1}^{2^L} p^{\theta_{ji}}(1-p)^{L-\theta_{ji}}\ddot{\Psi}_{n,k}(\tau_j; \mathbf{x}_n),$$

the second-order derivative of $\Upsilon_{n,k}$ can be written as

$$\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\Upsilon_{n,k} = (\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\ddot{\Upsilon}_{n,k}\sigma_v^{-2}\mathbf{C}^{(k,:)}_n$$

and

$$\frac{\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\Upsilon_{n,k}}{\Upsilon_{n,k}} = (\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\frac{\ddot{\Upsilon}_{n,k}}{\Upsilon_{n,k}}\sigma_v^{-2}\mathbf{C}^{(k,:)}_n \tag{52}$$

Substituting (51) and (52) into (50) yields

$$\begin{aligned}
&-\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\log\mathbb{P}\left\{\mathbf{m}_n \mid \mathbf{x}_n\right\} \\
&= (\sigma_v^2)^{-1}\sum_{i=1}^{M}(\mathbf{C}^{(k,:)}_n)^{\mathrm{T}}\left(\frac{\dot{\Upsilon}^2_{n,k}}{\Upsilon^2_{n,k}} - \frac{\ddot{\Upsilon}_{n,k}}{\Upsilon_{n,k}}\right)\mathbf{C}^{(k,:)}_n \\
&= \mathbf{C}^{\mathrm{T}}_n\Xi_n^{-1}\mathbf{C}_n
\end{aligned} \tag{53}$$

where

$$\Xi_n = \begin{bmatrix}
\frac{\sigma_v^2}{(\frac{\dot{\Upsilon}_{n,1}}{\Upsilon_{n,1}})^2 - \frac{\ddot{\Upsilon}_{n,1}}{\Upsilon_{n,1}}} & & \\
& \ddots & \\
& & \frac{\sigma_v^2}{(\frac{\dot{\Upsilon}_{n,M}}{\Upsilon_{n,M}})^2 - \frac{\ddot{\Upsilon}_{n,M}}{\Upsilon_{n,M}}}
\end{bmatrix}_{M\times M}$$

Therefore, we can see that

$$\mathbf{D}^{22,b}_{n-1} = \mathbb{E}\{-\Delta^{\mathbf{x}_n}_{\mathbf{x}_n}\log\mathbb{P}\{\mathbf{m}_n|\mathbf{x}_n\}\} = \mathbf{C}^{\mathrm{T}}_n\bar{\Xi}_n^{-1}\mathbf{C}_n$$

where $\bar{\Xi}_n \triangleq \mathbb{E}\{\Xi_n\}$ can be calculated by the Monte Carlo method as in [46], [50]. $\{\hat{\mathbf{x}}^{(t)}_{n|n-1}, \mathbf{P}^{(t)}_{n|n-1}, \mathbf{x}^{(t)}_n, \mathbf{m}^{(t)}_n\}_{t=1}^{T_{\mathrm{MC}}}$ represents the realization of T_{MC} trials of Monte Carlo. Then $\bar{\Xi}_n$ can be approximated by replacing $\{\hat{\mathbf{x}}_{n|n-1}, \mathbf{P}_{n|n-1}, \mathbf{x}_n, \mathbf{m}_n\}$ with these samplings and averaging. In addition, the initial value can be obtained from $\mathbf{J}_0 = \mathbb{E}\{-\Delta^{\mathbf{x}_0}_{\mathbf{x}_0}\log p(\mathbf{x}_0)\} = \mathbb{E}\{-\Delta^{\mathbf{x}_0}_{\mathbf{x}_0}\log\mathcal{N}(\mathbf{x}_0; \mu_0, \Sigma_0)\} = \mathbb{E}\{\Delta^{\mathbf{x}_0}_{\mathbf{x}_0}\frac{1}{2}(\mathbf{x}_0 - \mu_0)^{\mathrm{T}}\Sigma_0^{-1}(\mathbf{x}_0 - \mu_0)\} = \Sigma_0^{-1}$. Now, the proof is complete.

Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

References

[1] G. T. Amariucai, and S. Wei, Feedback-based collaborative secrecy encoding over binary symmetric channels, *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5248-5266, 2012.

[2] T. Aysal, M. Coates, and M. Rabbat, Distributed average consensus using probabilistic quantization, *In 2007 IEEE/SP 14th Workshop on Statistical Signal Processing*, IEEE, pp. 640–644, 2007.

[3] A. Bernstein, K. Steiglitz, and J. E. Hopcroft, Encoding of analog signals for binary symmetric channels, *IEEE Transactions on Information Theory*, vol. 12, no. 4, pp. 425-430, 1966.

[4] A. Bretas, B. Newton, J. London Jr., and B. Carvalho, Cyber-physical power systems state estimation, *Elsevier*, 2021.

[5] B. Chen, D. W. C. Ho, G. Hu, and L. Yu, Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks, *IEEE Transactions on Cybernetics*, vol. 48, no. 6, pp. 1862–1876, 2018.

[6] B. Chen, G. Hu, D. W. C. Ho, and L. Yu, Distributed Kalman filtering for time-varying discrete sequential systems, *Automatica*, vol. 99, pp. 228–236, 2019.

[7] W. Chen, Z. Wang, Q. Liu, D. Yue and G.-P. Liu, A new privacy-preserving average consensus algorithm with two-phase structure: applications to load sharing of microgrids, *Automatica*, vol. 167, art. no. 111715, Sept. 2024.

[8] Y. Chow, H. Teicher, Probability Theory: Independence, Interchangeability, Martingales (3rd ed), New York: Springer-Verlag, 1997.

[9] M. Colla, L. Tiziano, and M. Semo, Design and implementation of industrial automation control systems: A survey, *In 2009 7th IEEE International Conference on Industrial Informatics*, IEEE, pp. 570–575, 2009.

[10] R. Caballero-Águila, A. Hermoso-Carazo and J. Linares-Pérez, Networked distributed fusion estimation under uncertain outputs with random transmission delays, packet losses and multi-packet processing, *Signal Processing*, vol. 156, pp. 71-83, Mar. 2019.

[11] W. Chen, D. Ding, H. Dong and G. Wei, Distributed resilient filtering for power systems subject to denial-of-service attacks, *IEEE Transactions on Systems, Man, and Cybernetics-Systems*, vol. 49, no. 8, pp. 1688-1697, Aug. 2019.

[12] D. Ciuonzo, A. Aubry, and V. Carotenuto, Rician MIMO channel- and jamming-aware decision fusion, *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866–3880, 2017.

[13] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan and Z. Wang, A review: state estimation based on hybrid models of Kalman filter and neural network, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2173682, 2023.

[14] M. Fu and L. Xie, The sector bound approach to quantized feedback control, *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1698-1711, 2005.

[15] X. Ge, Q. Han and F. Yang, Event-based set-membership leader-following consensus of networked multi-agent systems subject to limited communication resources and unknown-but-bounded noise, *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5045–5054, 2017.

[16] L. Hu, Z. Wang and X. Liu, Dynamic state estimation of power systems with quantization effects: A recursive filter Approach, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1604-1614, 2016.

[17] N. Harel and T. Routtenberg, Non-Bayesian estimation with partially quantized observations." *In 2017 22nd International Conference on Digital Signal Processing (DSP)*, IEEE, pp. 1–5, 2017.

[18] A. Haydari and Y. Yilmaz, Deep reinforcement learning for intelligent transportation systems: A survey, *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[19] B. M. Kurkoski and H. Yagi, Single-bit quantization of binary-input, continuous-output channels, *in Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2088–2092, 2017.

[20] H. Li, and Y. Shi, Robust H_∞ filtering for nonlinear stochastic systems with uncertainties and Markov delays, *Automatica*, vol. 48, no. 1, pp. 159–166, 2012.

[21] W. Li, Y. Jia, and J. Du, Variance-constrained state estimation for non-linearly coupled complex networks, *IEEE Transactions on Cybernetics*, vol. 48, no. 2, pp. 818–824, 2018.

[22] W. Li and F. Yang, Information fusion over network dynamics with unknown correlations: An overview, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, art. no. 100003, Jun. 2023.

[23] A. Liu, L. Yu, W. Zhang and M. Z. Q. Chen, Moving horizon estimation for networked systems with quantized measurements and packet dropouts," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 7, pp. 1823-1834, 2013.

[24] H. Leung, C. Seneviratne and M. Xu, A novel statistical model for distributed estimation in wireless sensor networks, *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3154-3164, 2015.

[25] Q. Liu and Z. Wang, Moving-horizon estimation for linear dynamic networks with binary encoding schemes, *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1763-1770, 2021.

[26] Q. Li, Z. Wang, J. Hu, and W. Sheng, Distributed state and fault estimation over sensor networks with probabilistic quantizations: The dynamic event-triggered case, *Automatica*, vol. 131, p. 109784, 2021.

[27] Q. Liu, Z. Wang, H. Dong, and C. Jiang, Remote estimation for energy harvesting systems under multiplicative noises: a binary encoding scheme with probabilistic bit flips, *IEEE Transactions on Automatic Control*, 2022.

[28] Z. Matousek, M. Babjak, and J. Ochodnicky, Walsh-Hadamard sequences for binary encoding of radar signals, *IEEE International Conference on Military Technologies (ICMT)*, pp. 1-6, 2015.

[29] E. Msechu, S. Roumeliotis, A. Ribeiro, and G. Giannakis, Decentralized quantized Kalman filtering with scalable communication cost", *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3727-3741, 2008.

[30] C. Montella, The kalman filter and related algorithms: A literature review, *Research Gate*, 2011.

[31] R. Niu and P. K. Varshney, Target location estimation in sensor networks with quantized data, *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4519–4528, 2006.

[32] T. Nguyen and T. Nguyen, Optimal quantizer structure for maximizing mutual information under constraints, *IEEE Transactions on Communications*, vol. 69, np. 11, pp. 7406–7413, 2021.

[33] H. C. Papadopoulos, G. W. Wornell and A. V. Oppenheim, Sequential signal encoding from noisy measurements using quantizers with dynamic bias control, *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 978-1002, 2001.

[34] J. Park, J. Kim, H. Yoo, W. Sung, Construction of phase tracking codebooks based on the Lloyd-Max vector quantization, *In2011 IEEE Wireless Communications and Networking Conference*, IEEE, pp. 1983–1987, 2011.

[35] A. Rajagopal and S. Chitraganti, State estimation and control for networked control systems in the presence of correlated packet drops, *International Journal of Systems Science*, vol. 54, no. 11, pp. 2352–2365, 2023.

[36] A. Ribeiro, G. Giannakis, and S. Roumeliotis, SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations." *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4782-4795, 2006.

[37] S. M. Ross, Introduction to Probability Models, *New York*, NY, USA: Academic press, 2006.

[38] Rudin, W. (1991), Functional analysis, international series in pure and applied mathematics, New York: McGraw-Hill, Inc., ISBN 0-07-054236-8, MR 1157815C

[39] Y. S. Shmaliy, F. Lehmann, S. Zhao, and C. K. Ahn, Comparing robustness of the Kalman, H_∞, and UFIR filters, *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3447–3458, 2018.

[40] B. Sinopoli, L. Schenato, and M. Franceschetti, Kalman filtering with intermittent observations, *IEEE transactions on Automatic Control*, vol. 49, no. 9, pp. 1453-1464, 2004.

[41] S. Sun, Distributed optimal linear fusion predictors and filters for systems with random parameter matrices and correlated noises, *IEEE Transactions on Signal Processing*, vol. 68, pp. 1064–1074, 2020.

[42] P. Tichavsky, C. H. Muravchik and A. Nehorai, Posterior Cramer-Rao bounds for discrete-time nonlinear filtering, *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1386-1396, 1998.

[43] L. Wang, Z. Wang, D. Zhao, Y. Liu and G. Wei, Recursive filtering for discrete-time stochastic complex networks under bit-rate constraints: a locally minimum variance approach, *IEEE Transactions on Automatic Control*, vol. 69, no. 5, pp. 3441–3448, May. 2024.

[44] Y. Wang, H.-J. Liu and H.-L. Tan, An overview of filtering for sampled-data systems under communication constraints, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100011, Sep. 2023.

[45] Y.-A. Wang, B. Shen, L. Zou and Q.-L. Han, A survey on recent advances in distributed filtering over sensor networks subject to communication constraints, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, art. no. 100007, Jun. 2023.

[46] Y. Weng, Y. Zhao and B. Zhao, Target tracking using quantized innovations in wireless sensor networks, *2010 5th International Conference on Computer Science & Education*, pp. 69-74, 2010.

[47] J. Xiao, S. Cui, Z. Luo and A. J. Goldsmith, Power scheduling of universal decentralized estimation in sensor networks, *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413-422, 2006.

[48] T. Zhang, Q. Liu, J. Liu, Z. Wang and H. Li, Multiple-bipartite consensus for networked Lagrangian systems without using neighbours' velocity information in the directed graph, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2210185, 2023.

Citation information: DOI: 10.1109/TAC.2024.3496573, IEEE Transactions on Automatic Control

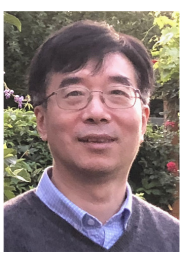
[49] X.-M. Zhang, Q.-L. Han, and X. Yu, Survey on recent advances in networked control systems, *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1740–1752, 2016.

[50] L. Zuo, R. Niu and P. K. Varshney, Posterior crlb based sensor selection for target tracking in sensor networks, *2007 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. II-1041-II-1044, 2007.

Qinyuan Liu received the B.Eng. degree in measurement and control technology and instrumentation from Huazhong University of Science and Technology, Wuhan, China, in 2012, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2017.

He is currently a Professor in the School of Computer Science and Technology, Tongji University, Shanghai, China. From Jul. 2015 to Sep. 2016, he was a Researcher Assistant in the Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. From Jan. 2016 to Jan. 2017, he was an international researcher in the Department of Computer Science, Brunel University London, UK. His research interests include networked control systems, multi-agent systems, and distributed filtering. He is an active reviewer for many international journals.

Yao Nie received the B.S. degree in engineering from the School of Computer Science and Technology, Tongji University, Shanghai, China, in 2012. He is currently working toward the Ph.D. degree in computer science and technology at Tongji University, Shanghai, China. His general research interests are in topics related to filtering theory and stochastic dynamical systems.



Zidong Wang (SM'03-F'14) received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published a number of papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, the Editor-in-Chief for *Systems Science & Control Engineering*, and an Associate Editor for 12 international journals including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics-Part C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Hongli Dong received the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2012.

From 2009 to 2010, she was a Research Assistant with the Department of Applied Mathematics, City University of Hong Kong, Hong Kong. From 2010 to 2011, she was a Research Assistant with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From 2011 to 2012, she was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, London, U.K. From 2012 to 2014, she was an Alexander von Humboldt Research Fellow with the University of Duisburg–Essen, Duisburg, Germany. She is currently a Professor with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing, China. She is also the Director of the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Daqing. Her current research interests include robust control and networked control systems.

Dr. Dong is a very active reviewer for many international journals.

Changjun Jiang received the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1995.

He is currently a Professor with the Department of Computer Science and Technology, Tongji University, Shanghai, China. He has published over 100 publications. His current research interests include concurrence theory, Petri nets, formal verification of software, cluster, grid technology, program testing, intelligent transportation systems, and service-oriented computing.