Citation information: DOI: 10.1109/TAC.2026.3673082, IEEE Transactions on Automatic Control

Moving-Horizon Estimation for Multi-Sensor Systems Under Probabilistic Caching Mechanism: A Co-Design Scheme

Weihao Song, Zidong Wang, Zhongkui Li, and Hongli Dong

***Abstract*—In practice, the cache, capable of storing frequently accessed data, is widely deployed in edge servers to guarantee quick retrieval and improve overall system performance. In this paper, the moving-horizon state estimation problem is investigated for a class of multi-sensor systems under the effects of limited caching capacity and sensor resolution. The measurement information collected by multiple sensors is first transmitted to an edge server for state estimation purposes and then stored in the cache for future use. To accommodate the limited caching capacity, the probabilistic caching mechanism (PCM) is harnessed to manage the cached content, under which only a portion of the measurement information is probabilistically selected and retained in the cache. By solving the least-squares optimization problem, a novel moving-horizon state estimator is proposed under the PCM. Sufficient conditions are derived to guarantee that the estimation error is exponentially ultimately bounded in the mean-square sense. To improve the estimation accuracy, the parameters of both the estimator and the PCM are jointly designed by addressing a constrained optimization problem with the assistance of the particle swarm optimization method. Finally, two examples are given to showcase the effectiveness of the proposed algorithm.**

***Index Terms*—Moving-horizon estimation, multi-sensor systems, probabilistic caching mechanism, sensor resolution, exponential ultimate boundedness.**

I. Introduction

With the rapid development of the fifth generation mobile networks and the Internet of Things, the limitations of centralized cloud computing paradigm, such as high latency and low efficiency, have become increasingly conspicuous in recent years. To tackle these challenges, edge servers have emerged as a powerful alternative by bringing computational resources to the network edge, which is in close proximity to end users or data sources [12]. Owing primarily to the distinctive advantages in latency reduction, real-time data processing, and efficiency enhancement, edge servers have demonstrated substantial application potential across a wide range of realms such as environmental monitoring, autonomous vehicles, intelligent transportation, and industrial automation [17], [19]. In these representative applications, timely and accurate decisions are crucial for guaranteeing the overall system performance, particularly in terms of efficiency and safety.

This work was supported in part by the National Natural Science Foundation of China under Grants 62203016, 62425301, U2241214, T2121002, and 62373008, in part by the China Postdoctoral Science Foundation under Grant 2021TQ0009, in part by the Royal Society of the UK, and in part by the Alexander von Humboldt Foundation of Germany. *(Corresponding author: Zhongkui Li.)*

Weihao Song is with the School of Advanced Manufacturing and Robotics, Peking University, Beijing 100871, China, and also with the Department of Computer Science, Brunel University of London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (Email: weihao.song@pku.edu.cn)

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

Zhongkui Li is with the School of Advanced Manufacturing and Robotics, Peking University, Beijing 100871, China. (Email: zhongkli@pku.edu.cn)

Hongli 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)

It should be noted that the accurate knowledge of the system state is essential for operators or automated systems to make informed decisions. Nevertheless, in real-world applications, it is often difficult (if not impossible) to directly obtain such state information due mainly to physical constraints or budget limitations [3], [14], [25]. A technically viable and economically sustainable approach is to infer the states of interest by leveraging the available, though possibly noisy, measurement signals. As such, the state estimation problem has garnered substantial research attention from both the signal processing and control communities, and a rich body of effective state estimation methods, tailored to different types of systems and noise models, has been reported in the literature, see e.g., [4], [5], [8], [16], [27] and the references therein.

Recently, the moving-horizon estimation, also known as receding-horizon estimation, has stirred a great deal of research interest due to its versatility in dealing with a wide range of system complexities, including model uncertainties and disturbance/state constraints [1], [11], [21], [26]. The fundamental idea behind the moving-horizon estimation is to recursively reconstruct the system states by solving an optimization problem based on a series of past measurements within a moving window/horizon of constant size. Compared with the computation-intensive full information estimation where all the past measurements are utilized, the moving-horizon estimation is able to achieve a balance between estimation accuracy and computational efficiency [18]. Consequently, such an estimation framework has found successful applications in many practical fields including, but not limited to, robotics and industrial process control [15], [23].

So far, a wealth of literature has been reported on the design of moving-horizon estimation schemes under various system settings and network-induced phenomena. For example, the moving-horizon estimation problem has been investigated in [31] for the networked time-delay systems, where the measurement transmissions are scheduled by the Round-Robin protocol. In [29], the event-triggered distributed moving-horizon estimation algorithm has been designed for linear discrete-time systems over wireless sensor networks. It is worth mentioning that, in most existing literature, there is an implicit assumption that the actual measurement output is identical to the true value. Nevertheless, this assumption is often unrealistic due to the inherent limitations of sensor resolution, which determines the smallest detectable variation. The sensor-resolution-induced measurement deviations, if not properly handled, would inevitably lead to performance degradation [20], [30]. Consequently, it makes practical sense to investigate the moving-horizon estimation problem under the sensor resolution effects.

On another research frontier, caches are usually deployed in edge servers to reduce the time of retrieving frequently accessed content, thereby enhancing both system performance and user experience [10]. The cache-enabled real-time property is particularly desired by latency-sensitive applications (e.g., mobile games and video streaming) since the delays in the service delivery would significantly affect the user retention and satisfaction in a negative way. Notably, in practical engineering, it is virtually impossible to cache all the potentially required data due to the inherent physical size limitation that restricts the storage capacity available for caching [7], [24]. As

such, it is of great necessity to develop efficient caching strategies that determine which data should be cached. In this regard, a promising approach known as the probabilistic caching mechanism (PCM) has been considered in [9], where each content potentially requested by the user is cached based on a certain probability. Due to its outstanding adaptability and flexibility, the PCM has been widely employed in the computer network and wireless communication fields [6], [22]. Nevertheless, in the specific context of state estimation where the limited caching capacity becomes a major concern, the corresponding issues remain underexplored and deserve further investigation.

Summarizing the above discussions, in this paper, we endeavor to deal with the moving-horizon estimation problem for linear discrete-time systems under the PCM and sensor resolution effects. This is far from a trivial task due to the following three fundamental difficulties: 1) how to construct a unified and practical state estimation framework under the consideration of PCM and sensor resolution effects? 2) how to derive sufficient conditions that ensure the boundedness of the estimation error under the PCM? and 3) how to co-design the parameters involved in both the PCM and the moving-horizon estimator to optimize the estimation performance? The main purpose of this paper is to offer concrete solutions to these difficulties by conducting a systematic investigation.

In light of the identified challenges, the primary contributions of this paper can be emphasized as follows:

1) the moving-horizon state estimation problem is, for the first time, explored for multi-sensor systems under the sensor resolution effects and the PCM, which is able to accommodate the engineering practice with efficient management of limited caching capacity;
2) some sufficient conditions are established to ensure the exponential ultimate boundedness of the estimation error in the mean-square sense, and the asymptotic upper bound is derived to examine the impact of the PCM; and
3) a constrained optimization problem is constructed to jointly design the parameters of both the moving-horizon estimator and the PCM, where the particle swarm optimization (PSO) method combined with a family of matrix inequalities is utilized to optimize the estimation performance while satisfying the parameter constraints.

Notations: $\mathbb{R}^n$ represents the n-dimensional Euclidean space. For a matrix A, A^T and A^{-1} denote, respectively, its transpose and inverse, and $\sigma_{\min}(A)$ signifies its minimal eigenvalue. $\mathbb{E}\{\cdot\}$ denotes the mathematical expectation. $\mathrm{col}\{a_i\}_{i=1}^N$ stands for a column vector formed by stacking $a_1, a_2, \ldots, a_N$ on top of one another. For a vector x, $\|x\|$ and $\|x\|_Q$ represent, respectively, its Euclidean norm and weighted norm defined by $\|x\|_Q \triangleq \sqrt{x^TQx}$, where Q a positive definite matrix. $\mathrm{mod}(a, b)$ returns the nonnegative remainder when the integer a is divided by the positive integer b. $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote, respectively, the floor function and the ceil function.

II. Problem Formulation

A. System description

Consider a class of discrete-time systems with the following dynamics:

$$\begin{aligned} x_{k+1} &= Gx_k + \mu_k \\ \bar{y}_{i,k} &= H_i x_k + \nu_{i,k},\ i = 1, 2, \ldots, N \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ signifies the system state vector and $\bar{y}_{i,k} \in \mathbb{R}^{n_y}$ denotes the ideal measurement output for the i-th sensor. $\mu_k \in \mathcal{U} \triangleq \{\mu : \|\mu\| \le \mu_{\max}; \mu \in \mathbb{R}^{n_x}\}$ and $\nu_{i,k} \in \mathcal{V}^i \triangleq \{\nu : \|\nu\| \le \nu_{i,\max}; \nu \in \mathbb{R}^{n_y}\}$ represent, respectively, the process noise and the measurement noise for the i-th sensor. G and H_i are constant matrices with compatible dimensions.

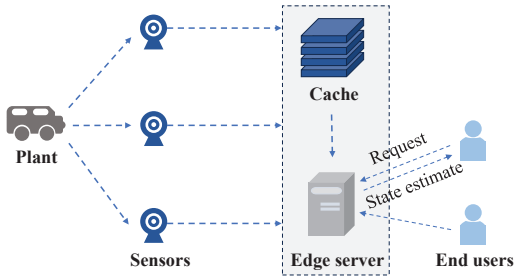

Fig. 1: An illustrative diagram for the considered scenario, where the vehicle icon represents a generic dynamic system.

In practical applications, the ability of sensors to capture arbitrarily small variations in the quantity being measured is constrained, and the smallest detectable variation is typically referred to as the sensor resolution.

Definition 1 ([30]): If the l-th component ($l = 1, 2, \ldots, n_y$) of the measurement output of a sensor belongs to the set $\{nm^l | n = 0, \pm1, \pm2, \ldots, \pm s\}$ where s is a given positive integer, then $m \triangleq [m^1, m^2, \ldots, m^{n_y}]^T$ is named as the sensor resolution.

Taking the sensor resolution into account and following the line of [20], [30], the actual measurement output for the l-th component of the i-th sensor, denoted by $y_{i,k}^l$, can be described as follows:

$$y_{i,k}^l = \begin{cases} \left\lfloor \frac{\bar{y}_{i,k}^l}{m_i^l} \right\rfloor m_i^l, & \bar{y}_{i,k}^l \ge m_i^l \\ 0, & \bar{y}_{i,k}^l \in (-m_i^l, m_i^l) \\ \left\lceil \frac{\bar{y}_{i,k}^l}{m_i^l} \right\rceil m_i^l, & \bar{y}_{i,k}^l \le -m_i^l \end{cases} \tag{2}$$

where $\bar{y}_{i,k}^l$ represents the ideal measurement output for the l-th component of the i-th sensor, and m_i^l stands for the corresponding sensor resolution. Denote by $d_{i,k}^l$ the difference between the actual measurement and the ideal measurement, namely, $d_{i,k}^l \triangleq y_{i,k}^l - \bar{y}_{i,k}^l$. Clearly, one has

$$|d_{i,k}^l| < m_i^l. \tag{3}$$

For notational convenience, let us denote

$$\begin{aligned} \bar{y}_k &= \mathrm{col}\{\bar{y}_{i,k}\}_{i=1}^N,\ y_k = \mathrm{col}\{y_{i,k}\}_{i=1}^N \\ d_k &= \mathrm{col}\{d_{i,k}\}_{i=1}^N,\ \nu_k = \mathrm{col}\{\nu_{i,k}\}_{i=1}^N. \end{aligned}$$

B. Characterization of the PCM

In this paper, as depicted in Fig. 1, the state estimation process at the edge server is carried out in response to the requests initiated by the end users. Moreover, the edge server and the cache are combined to enable the fast and efficient delivery of the queried/retrieved information to the end users. For this purpose, the frequently required data are stored in the cache, where the information replacement follows the well-known first-in-first-out rule (i.e., the oldest information is discarded and the newest information is cached at each update).

In the context of state estimation, the measurements collected by the multi-sensor systems will be transmitted to the edge server (fusion center) for further processing and to the cache for future use. Nevertheless, in practical engineering, due to the limited storage capacity of the cache, only a portion of newly collected measurements can be stored in the cache. In this paper, it is supposed that only the measurements of S ($1 \le S < N$) sensors are permitted to be cached to cope with the storage constraints. Clearly, there are R possible

cases for the cached measurements, and each case corresponds to one caching mode, where

$$R = \binom{N}{S} = \frac{N!}{(N-S)!S!} \tag{4}$$

and $N!$ denotes the factorial of N.

More specifically, under the PCM, each caching mode can be characterized by a matrix selected from the following finite set:

$$\mathcal{Y} = \{\bar{\Xi}_1, \bar{\Xi}_2, \cdots, \bar{\Xi}_R\} \tag{5}$$

where the matrix $\bar{\Xi}_r$ ($r = 1, 2, \ldots, R$), corresponding to the r-th caching mode, is a diagonal matrix composed of S diagonal elements "1" and $N - S$ diagonal elements "0". For example, if $N = 3$ and $S = 2$, the matrix $\bar{\Xi}_1$ is defined as $\mathrm{diag}\{1, 1, 0\}$, which indicates that only the measurements from the first and second sensors are stored in the cache. Similarly, $\bar{\Xi}_2$ and $\bar{\Xi}_3$ are respectively defined as $\mathrm{diag}\{1, 0, 1\}$ and $\mathrm{diag}\{0, 1, 1\}$.

To describe the caching mode at each time instant, a random variable β_k, taking values in $\{1, 2, \ldots, R\}$, is defined with the following probability distribution:

$$\Pr\{\beta_k = r\} = \xi_r, \ r = 1, 2, \ldots, R \tag{6}$$

$$\xi_r \in [0, 1], \ \sum_{r=1}^{R} \xi_r = 1 \tag{7}$$

where ξ_r denotes the caching probability allocated to the r-th caching mode. It is clear that based on the specific caching mode at time instant k, the available measurement information at the edge server can be denoted by

$$y_k^a = \Xi_{\beta_k} y_k \tag{8}$$

where $\Xi_{\beta_k} = \bar{\Xi}_{\beta_k} \otimes I_{n_y}$ and $\otimes$ signifies the Kronecker product.

Remark 1: It is worth mentioning that under the PCM, only a subset of the sensor measurements can be randomly selected and stored in the cache to make efficient use of the limited storage resources. To facilitate the modeling and analysis of such a mechanism, it is essential to determine all the possible combinations of measurements that can be cached (i.e., all the possible caching modes) as well as their corresponding caching probabilities [9]. On the other hand, when retrieving data from the cache, we are able to determine which sensors the cached measurements belong to by resorting to certain unique information such as the packet identifier. Accordingly, at the edge server, the available measurement information obtained from the cache can be uniformly described by (8), where the uncached measurements are represented by zeros.

Remark 2: In the existing literature, the probabilistic caching strategy is usually explored to optimize the cache hit probability, i.e., the probability that a random file request can be successfully satisfied [6]. Nevertheless, in practical engineering, it might be more desirable to investigate the effect of PCM onto the task completion performance in specific scenarios. More specifically, in the context of state estimation, it is warranted to examine the impact of the caching probabilities defined in (6) on the estimation performance and design proper parameters to optimize the designated performance indices.

C. Moving-horizon state estimation problem

In this paper, we are going to consider the moving-horizon state estimation problem for the discrete-time systems (1) under the probabilistic caching mechanism. Assume that the fixed size of the moving window is represented by a positive integer M and is compatible with the limited storage space of the cache, which implies that the required historical measurements can be retrieved from the cache. More specifically, at each time instant $k \geq M$, we aim to obtain the estimates of the states $x_{k-M}, x_{k-M+1}, \ldots, x_k$, denoted respectively by $\hat{x}_{k-M|k}, \hat{x}_{k-M+1|k}, \ldots, \hat{x}_{k|k}$, based on the prior prediction $\bar{x}_{k-M}$, the newly received measurements y_k as well as the previously cached measurements y_s^a ($s = k-M, k-M+1, \ldots, k-1$).

For convenience, let us formulate the moving-horizon state estimation problem under the PCM as follows.

Problem 1: Suppose that the prior prediction $\bar{x}_{k-M}$, the current measurements y_k, and the cached measurements y_s^a ($s = k-M, k-M+1, \ldots, k-1$) are available at a time instant $k \geq M$. At each time instant k, determine the optimal state estimates $\hat{x}_{k-M|k}, \hat{x}_{k-M+1|k}, \ldots, \hat{x}_{k|k}$ by minimizing the following least-squares cost function

$$\begin{aligned} J_k = &\|\hat{x}_{k-M|k} - \bar{x}_{k-M}\|_{\Omega}^2 + \|y_k - \hat{y}_{k|k}\|^2 \\ &+ \sum_{s=k-M}^{k-1} \|y_s^a - \hat{y}_{s|k}^a\|^2 \end{aligned} \tag{9}$$

under the constraints defined below

$$\begin{cases} \hat{y}_{k|k} = H\hat{x}_{k|k} \\ \hat{y}_{s|k}^a = \Xi_{\beta_s} H\hat{x}_{s|k}, \ s \in [k-M, k-1] \\ \hat{x}_{s+1|k} = G\hat{x}_{s|k}, \ s \in [k-M, k-1] \\ \bar{x}_{k-M} = G\hat{x}_{k-M-1|k-1} \end{cases} \tag{10}$$

where $H = \begin{bmatrix} H_1^T & H_2^T & \cdots & H_N^T \end{bmatrix}^T$ and Ω denotes a positive definite matrix to be determined later.

Remark 3: It should be emphasized that the cost function defined in (9) is composed of three terms. Specifically, the first term $\|\hat{x}_{k-M|k} - \bar{x}_{k-M}\|_{\Omega}^2$ penalizes the difference between the state estimate and the prior prediction, where the parameter matrix Ω is utilized to adjust the relative emphasis on the model-based prediction and the measurement-based correction. The second term $\|y_k - \hat{y}_{k|k}\|^2$ imposes a penalty for the distance between the newly received measurements y_k and the estimated measurements. The third term $\|y_s^a - \hat{y}_{s|k}^a\|^2$ is introduced to penalize the discrepancy between the previously cached measurements y_s^a and the corresponding estimates. Note that at each time instant, the estimator can receive the latest measurements from all sensors, whereas at later time instants, only the cached subset of these measurements remains available. On the other hand, the conventional moving-horizon estimation methods are no longer suitable for the scenario considered in this paper due mainly to the measurement deviations incurred by the sensor resolution and the partially available measurements caused by the PCM. In this sense, it is of critical necessity to design a novel moving-horizon estimator that can handle these identified complexities and better accommodate the needs of practical engineering.

Before proceeding further, let us define the estimation error as $\tilde{x}_{k-M} = x_{k-M} - \hat{x}_{k-M|k}$ and present the following definition with respect to the exponential ultimate boundedness in the mean-square sense.

Definition 2 ([28]): The dynamics of the estimation error $\tilde{x}_k$ is said to be exponentially ultimately bounded in the mean-square sense if there exist three positive constants α, γ, and ρ such that

$$\mathbb{E}\{\|\tilde{x}_k\|^2\} \leq \alpha^k \gamma + \rho$$

where $\alpha \in [0, 1)$. α and ρ denote, respectively, the decay rate and the asymptotic upper bound of $\mathbb{E}\{\|\tilde{x}_k\|^2\}$.

In this paper, we are focused on investigating the moving-horizon state estimation problem for multi-sensor systems under the PCM and sensor resolution effects. The primary purpose of this paper is to co-design the parameters of the estimator and the PCM such that: 1) the exponential ultimate boundedness is guaranteed for the estimation error dynamics in the mean-square sense; and 2) the estimation performance is optimized by minimizing the asymptotic upper bound.

III. Main Results

In this section, we first establish an explicit expression for the estimation error dynamics by solving Problem 1. Then, with the aid of stochastic analysis techniques, some sufficient conditions are found to ensure the mean-square exponential ultimate boundedness of estimation error. Subsequently, we concentrate on the co-design problem of moving-horizon estimator and PCM to optimize the estimation performance.

For ease of notation, we denote

$$\begin{aligned}
\bar{Y}_{k-1} &= \mathrm{col}\{\bar{y}_s\}_{s=k-M}^{k-1},\ D_{k-1} = \mathrm{col}\{d_s\}_{s=k-M}^{k-1}\\
U_{k-1} &= \mathrm{col}\{\mu_s\}_{s=k-M}^{k-1},\ V_{k-1} = \mathrm{col}\{\nu_s\}_{s=k-M}^{k-1}\\
Y_{k-1}^a &= \mathrm{col}\{y_s^a\}_{s=k-M}^{k-1},\ \mathcal{H}_M = HG^M\\
\mathcal{H}_{M-1} &= \begin{bmatrix} H^T & (HG)^T & \cdots & (HG^{M-1})^T \end{bmatrix}^T\\
\Phi_{k-1} &= \mathrm{diag}\{\Xi_{\beta_{k-M}}, \Xi_{\beta_{k-M+1}}, \cdots, \Xi_{\beta_{k-1}}\}\\
\mathcal{G}_{M-2} &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 0\\ H & 0 & \cdots & 0 & 0\\ HG & H & \cdots & 0 & 0\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ HG^{M-2} & HG^{M-3} & \cdots & H & 0 \end{bmatrix}\\
\mathcal{G}_{M-1} &= \begin{bmatrix} HG^{M-1} & HG^{M-2} & \cdots & HG & H \end{bmatrix}.
\end{aligned}$$

Based on the above notations, the system dynamics (1), and the PCM (6), one has

$$\begin{aligned}
Y_{k-1}^a &= \Phi_{k-1}\bar{Y}_{k-1} + \Phi_{k-1}D_{k-1}\\
\bar{Y}_{k-1} &= \mathcal{H}_{M-1}x_{k-M} + \mathcal{G}_{M-2}U_{k-1} + V_{k-1}\\
\bar{y}_k &= \mathcal{H}_M x_{k-M} + \mathcal{G}_{M-1}U_{k-1} + \nu_k.
\end{aligned} \tag{11}$$

In the following theorem, we will determine the optimal solution to the moving-horizon state estimation problem defined in Problem 1.

Theorem 1: Suppose that the parameter matrix Ω, the prior prediction $\bar{x}_{k-M}$, the current measurements y_k, and the cached measurements Y_{k-1}^a are available at a time instant $k \geq M$. At each time instant k, the unique solution to Problem 1 can be described by

$$\begin{aligned}
\hat{x}_{k-M|k} = &\left(\Omega + \mathcal{H}_M^T\mathcal{H}_M + \mathcal{H}_{M-1}^T\Phi_{k-1}\mathcal{H}_{M-1}\right)^{-1}\\
&\times\left(\Omega\bar{x}_{k-M} + \mathcal{H}_M^T y_k + \mathcal{H}_{M-1}^T\Phi_{k-1}Y_{k-1}^a\right).
\end{aligned} \tag{12}$$

Proof: It is clear that the least-squares cost function (9) can be rewritten as

$$\begin{aligned}
J_k = &\|\hat{x}_{k-M|k} - \bar{x}_{k-M}\|_\Omega^2 + \|y_k - \mathcal{H}_M\hat{x}_{k-M|k}\|^2\\
&+ \|Y_{k-1}^a - \Phi_{k-1}\mathcal{H}_{M-1}\hat{x}_{k-M|k}\|^2.
\end{aligned} \tag{13}$$

The necessary condition for the minimum of (13) is given by

$$\begin{aligned}
\frac{\partial J_k}{\partial \hat{x}_{k-M|k}} = &\ 2\Omega(\hat{x}_{k-M|k} - \bar{x}_{k-M}) - 2\mathcal{H}_M^T(y_k - \mathcal{H}_M\hat{x}_{k-M|k})\\
&- 2\mathcal{H}_{M-1}^T\Phi_{k-1}^T(Y_{k-1}^a - \Phi_{k-1}\mathcal{H}_{M-1}\hat{x}_{k-M|k}) = 0.
\end{aligned}$$

On the other hand, it can be verified that the Hessian matrix with respect to (13) is a positive definite matrix. Consequently, the optimal estimate that minimizes the least-squares cost function (9) can be designed as in (12), which completes the proof. ∎

Recalling the definition of the estimation error $\tilde{x}_{k-M}$ and (12), it is not difficult to obtain that

$$\begin{aligned}
\tilde{x}_{k-M} &= x_{k-M} - \hat{x}_{k-M|k}\\
&= \left(\Omega + \mathcal{H}_M^T\mathcal{H}_M + \mathcal{H}_{M-1}^T\Phi_{k-1}\mathcal{H}_{M-1}\right)^{-1}\\
&\quad\times\Big(\Omega x_{k-M} + \mathcal{H}_M^T\mathcal{H}_M x_{k-M} + \mathcal{H}_{M-1}^T\Phi_{k-1}\\
&\quad\times\mathcal{H}_{M-1}x_{k-M} - \Omega\bar{x}_{k-M} - \mathcal{H}_M^T\bar{y}_k - \mathcal{H}_M^T d_k\\
&\quad - \mathcal{H}_{M-1}^T\Phi_{k-1}\bar{Y}_{k-1} - \mathcal{H}_{M-1}^T\Phi_{k-1}D_{k-1}\Big).
\end{aligned} \tag{14}$$

Based on (1), (10) and (11), the estimation error dynamics (14) can be further expressed as

$$\begin{aligned}
&\tilde{x}_{k-M}\\
&= \left(\Omega + \mathcal{H}_M^T\mathcal{H}_M + \mathcal{H}_{M-1}^T\Phi_{k-1}\mathcal{H}_{M-1}\right)^{-1}\big(\Omega G\tilde{x}_{k-M-1}\\
&\quad + \Omega\mu_{k-M-1} - \mathcal{H}_M^T\mathcal{G}_{M-1}U_{k-1} - \mathcal{H}_M^T\nu_k - \mathcal{H}_M^T d_k\\
&\quad - \mathcal{H}_{M-1}^T\Phi_{k-1}\mathcal{G}_{M-2}U_{k-1} - \mathcal{H}_{M-1}^T\Phi_{k-1}V_{k-1}\\
&\quad - \mathcal{H}_{M-1}^T\Phi_{k-1}D_{k-1}\big).
\end{aligned} \tag{15}$$

Remark 4: It is clear from (15) that due to the presence of PCM, a stochastic diagonal matrix Φ_{k-1} is involved in the estimation error dynamics. In other words, the estimation error $\tilde{x}_{k-M}$ is essentially a stochastic sequence, which makes it necessary to analyze the behaviors of the estimation error from a statistical perspective. On the other hand, the measurement deviations induced by the sensor resolution effects are also reflected in the estimation error dynamics. In this sense, the estimator parameter and the PCM should be meticulously designed to secure the expected estimation performance.

In what follows, based on the established estimation error dynamics (15), our attention will be centered on the analysis and design problems of the moving-horizon state estimator (12) under the PCM characterized by (6) and (7).

A. Analysis of the moving-horizon estimator under the PCM

Before proceeding further, the following lemma, which can be easily derived from Propositions 1 and 2 in [32], is given to facilitate the reformulation of estimation error dynamics (15).

Lemma 1: The PCM characterized by $\{\beta_{k-s}\}_{1\leq s\leq M}$ can be mapped to the variable $\phi_{k-1} \in \mathcal{S} \triangleq \{1, 2, \cdots, R^M\}$ by resorting to the following relationship:

$$\phi_{k-1} = \sum_{s=1}^{M} R^{s-1}(\beta_{k-s} - 1) + 1$$

where R denotes the number of all possible caching modes defined in (4). In this context, given the value of ϕ_{k-1}, the caching modes can be revealed by the functions $\mathcal{T}_s(\phi_{k-1})$ defined as follows:

$$\beta_{k-s} = \mathcal{T}_s(\phi_{k-1}) \triangleq \mathrm{mod}\left(\left\lfloor\frac{\phi_{k-1} - 1}{R^{s-1}}\right\rfloor, R\right) + 1.$$

In addition, the stochastic process ϕ_{k-1} is a Markov chain with the following transition probability:

$$p_{ab} \triangleq \Pr\{\phi_k = b|\phi_{k-1} = a\} = \begin{cases} \xi_{\mathcal{T}_1(b)} & \text{if } 1 \leq \omega(a,b) \leq R\\ 0 & \text{otherwise} \end{cases}$$

where $a, b \in \mathcal{S}$, $\omega(a,b) = b - R(a - 1 - R^{M-1}(\mathcal{T}_M(a) - 1))$ and $\xi_{\mathcal{T}_1(b)}$ denotes the caching probability defined in (6). Specifically, $\omega(a,b)$ serves as an indicator for whether the transition is possible.

It follows from Lemma 1 that the stochastic matrix Φ_{k-1} can be determined by the newly introduced variable ϕ_{k-1}. Therefore, we rewrite Φ_{k-1} as $\Phi_{\phi_{k-1}}$ and obtain the following dynamics:

$$\begin{aligned}
\tilde{x}_{k-M} = &\left(\Omega + \mathcal{H}_M^T\mathcal{H}_M + \mathcal{H}_{M-1}^T\Phi_{\phi_{k-1}}\mathcal{H}_{M-1}\right)^{-1}\big(\Omega G\tilde{x}_{k-M-1}\\
&+ \Omega\mu_{k-M-1} - \mathcal{H}_M^T\mathcal{G}_{M-1}U_{k-1} - \mathcal{H}_M^T\nu_k - \mathcal{H}_M^T d_k\\
&- \mathcal{H}_{M-1}^T\Phi_{\phi_{k-1}}\mathcal{G}_{M-2}U_{k-1} - \mathcal{H}_{M-1}^T\Phi_{\phi_{k-1}}V_{k-1}\\
&- \mathcal{H}_{M-1}^T\Phi_{\phi_{k-1}}D_{k-1}\big).
\end{aligned} \tag{16}$$

To facilitate the subsequent analysis, let us define

$$\bar{V}_{k+1} = \begin{bmatrix} V_k^T & \nu_{k+1}^T \end{bmatrix}^T,\ \bar{D}_{k+1} = \begin{bmatrix} D_k^T & d_{k+1}^T \end{bmatrix}^T$$

$$\zeta_{k+1} = \begin{bmatrix} \tilde{x}_{k-M}^T & \mu_{k-M}^T & U_k^T & \bar{V}_{k+1}^T & \bar{D}_{k+1}^T \end{bmatrix}^T$$
$$K_{\phi_k} = \left(\Omega + \mathcal{H}_M^T \mathcal{H}_M + \mathcal{H}_{M-1}^T \Phi_{\phi_k} \mathcal{H}_{M-1}\right)^{-1}$$
$$\bar{\mathcal{H}}_{\phi_k,M} = \mathcal{H}_M^T \mathcal{G}_{M-1} + \mathcal{H}_{M-1}^T \Phi_{\phi_k} \mathcal{G}_{M-2}$$
$$\tilde{\mathcal{H}}_{\phi_k,M} = \begin{bmatrix} \mathcal{H}_{M-1}^T \Phi_{\phi_k} & \mathcal{H}_M^T \end{bmatrix}.$$

Now, we are poised to provide a sufficient condition that assures the exponential ultimate boundedness of the estimation error in the mean-square sense.

Theorem 2: Consider the estimation error dynamics (16) and the PCM characterized by (6) and (7). Suppose that the estimator parameter matrix Ω and the caching probabilities ξ_r $(r = 1, 2, \ldots, R)$ are given. If there exist positive definite matrices P_a $(a \in \mathcal{S})$ as well as positive scalars $0 < \theta < 1$ and ϵ_i $(i = 1, 2, 3, 4)$ satisfying the following matrix inequalities:

$$\Lambda_a = \begin{bmatrix} \Lambda_a^{11} & \Lambda_a^{12} & \Lambda_a^{13} & \Lambda_a^{14} & \Lambda_a^{15} \\ * & \Lambda_a^{22} & \Lambda_a^{23} & \Lambda_a^{24} & \Lambda_a^{25} \\ * & * & \Lambda_a^{33} & \Lambda_a^{34} & \Lambda_a^{35} \\ * & * & * & \Lambda_a^{44} & \Lambda_a^{45} \\ * & * & * & * & \Lambda_a^{55} \end{bmatrix} < 0 \tag{17}$$

where

$$\Lambda_a^{11} = \bar{\Lambda}_a^{11} + \theta P_a,\ \bar{\Lambda}_a^{11} = G^T \Omega \tilde{P}_a \Omega G - P_a,\ \Lambda_a^{12} = G^T \Omega \tilde{P}_a \Omega$$
$$\Lambda_a^{13} = -G^T \Omega \tilde{P}_a \bar{\mathcal{H}}_{a,M},\ \Lambda_a^{14} = -G^T \Omega \tilde{P}_a \tilde{\mathcal{H}}_{a,M}$$
$$\Lambda_a^{15} = \Lambda_a^{14},\ \Lambda_a^{22} = \bar{\Lambda}_a^{22} - \epsilon_1 I,\ \bar{\Lambda}_a^{22} = \Omega \tilde{P}_a \Omega$$
$$\Lambda_a^{23} = -\Omega \tilde{P}_a \bar{\mathcal{H}}_{a,M}, \Lambda_a^{24} = -\Omega \tilde{P}_a \tilde{\mathcal{H}}_{a,M}, \Lambda_a^{25} = \Lambda_a^{24}$$
$$\Lambda_a^{33} = \bar{\Lambda}_a^{33} - \epsilon_2 I,\ \bar{\Lambda}_a^{33} = \bar{\mathcal{H}}_{a,M}^T \tilde{P}_a \bar{\mathcal{H}}_{a,M}$$
$$\Lambda_a^{34} = \bar{\mathcal{H}}_{a,M}^T \tilde{P}_a \tilde{\mathcal{H}}_{a,M},\ \Lambda_a^{35} = \Lambda_a^{34},\ \Lambda_a^{44} = \bar{\Lambda}_a^{44} - \epsilon_3 I$$
$$\bar{\Lambda}_a^{44} = \tilde{\mathcal{H}}_{a,M}^T \tilde{P}_a \tilde{\mathcal{H}}_{a,M},\ \Lambda_a^{45} = \bar{\Lambda}_a^{44},\ \Lambda_a^{55} = \bar{\Lambda}_a^{55} - \epsilon_4 I$$
$$\bar{\Lambda}_a^{55} = \bar{\Lambda}_a^{44},\ \tilde{P}_a = K_a \bar{P}_a K_a,\ \bar{P}_a = \sum_{b \in \mathcal{S}} p_{ab} P_b,$$

then the estimation error is exponentially ultimately bounded in the mean-square sense.

Proof: To begin with, let us choose the following Lyapunov-like function:

$$W_k = \tilde{x}_{k-M}^T P_{\phi_k} \tilde{x}_{k-M}. \tag{18}$$

Defining $\Delta W_k \triangleq W_{k+1} - W_k$, it follows from (16) and (18) that

$$\begin{aligned} &\mathbb{E}\{\Delta W_k | \phi_k = a\} \\ &= \mathbb{E}\left\{\tilde{x}_{k-M+1}^T P_{\phi_{k+1}} \tilde{x}_{k-M+1} - \tilde{x}_{k-M}^T P_{\phi_k} \tilde{x}_{k-M} | \phi_k = a\right\} \\ &= \mathbb{E}\Big\{\Big(\Omega G \tilde{x}_{k-M} + \Omega \mu_{k-M} - \tilde{\mathcal{H}}_{a,M} \bar{V}_{k+1} - \bar{\mathcal{H}}_{a,M} U_k \\ &\quad - \tilde{\mathcal{H}}_{a,M} \bar{D}_{k+1}\Big)^T \tilde{P}_a \Big(\Omega G \tilde{x}_{k-M} + \Omega \mu_{k-M} - \tilde{\mathcal{H}}_{a,M} \bar{V}_{k+1} \\ &\quad - \bar{\mathcal{H}}_{a,M} U_k - \tilde{\mathcal{H}}_{a,M} \bar{D}_{k+1}\Big) - \tilde{x}_{k-M}^T P_a \tilde{x}_{k-M} | \phi_k = a\Big\} \\ &= \mathbb{E}\{\zeta_{k+1}^T \bar{\Lambda}_a \zeta_{k+1} | \phi_k = a\} \end{aligned} \tag{19}$$

where

$$\bar{\Lambda}_a = \begin{bmatrix} \bar{\Lambda}_a^{11} & \Lambda_a^{12} & \Lambda_a^{13} & \Lambda_a^{14} & \Lambda_a^{15} \\ * & \bar{\Lambda}_a^{22} & \Lambda_a^{23} & \Lambda_a^{24} & \Lambda_a^{25} \\ * & * & \bar{\Lambda}_a^{33} & \Lambda_a^{34} & \Lambda_a^{35} \\ * & * & * & \bar{\Lambda}_a^{44} & \Lambda_a^{45} \\ * & * & * & * & \bar{\Lambda}_a^{55} \end{bmatrix}.$$

Based on (17) and (19), it is not difficult to verify that

$$\begin{aligned} &\mathbb{E}\{\Delta W_k | \phi_k = a\} \\ &= \mathbb{E}\Big\{\zeta_{k+1}^T \bar{\Lambda}_a \zeta_{k+1} + \theta W_k - \epsilon_1 \mu_{k-M}^T \mu_{k-M} - \epsilon_2 U_k^T U_k \\ &\quad - \epsilon_3 \bar{V}_{k+1}^T \bar{V}_{k+1} - \epsilon_4 \bar{D}_{k+1}^T \bar{D}_{k+1} - \theta W_k + \epsilon_1 \mu_{k-M}^T \mu_{k-M} \\ &\quad + \epsilon_2 U_k^T U_k + \epsilon_3 \bar{V}_{k+1}^T \bar{V}_{k+1} + \epsilon_4 \bar{D}_{k+1}^T \bar{D}_{k+1} | \phi_k = a\Big\} \\ &= \mathbb{E}\Big\{\zeta_{k+1}^T \Lambda_a \zeta_{k+1} - \theta W_k + \epsilon_1 \mu_{k-M}^T \mu_{k-M} + \epsilon_2 U_k^T U_k \\ &\quad + \epsilon_3 \bar{V}_{k+1}^T \bar{V}_{k+1} + \epsilon_4 \bar{D}_{k+1}^T \bar{D}_{k+1} | \phi_k = a\Big\} \\ &\leq -\theta \mathbb{E}\{W_k | \phi_k = a\} + \bar{\rho} \end{aligned} \tag{20}$$

where

$$\bar{\rho} \triangleq (\epsilon_1 + \epsilon_2 M) \mu_{\max}^2 + \epsilon_3 (M+1) \bar{\nu}_{\max}^2 + \epsilon_4 (M+1) \bar{m}^2 \tag{21}$$

with $\bar{\nu}_{\max}^2 = \sum_{i=1}^N \nu_{i,\max}^2$ and $\bar{m}^2 = \sum_{i=1}^N \sum_{l=1}^{n_y} (m_i^l)^2$.

Taking the mathematical expectation on both sides of (20), it is clear that

$$\mathbb{E}\{W_{k+1}\} \leq (1-\theta)\mathbb{E}\{W_k\} + \bar{\rho}, \tag{22}$$

which further indicates that

$$\begin{aligned} &\mathbb{E}\{W_{k+1}\} \\ &\leq (1-\theta)^{k-M+1} \mathbb{E}\{W_M\} + \sum_{\lambda=M}^{k} (1-\theta)^{k-\lambda} \bar{\rho} \\ &= (1-\theta)^{k-M+1} \mathbb{E}\{W_M\} + \frac{1-(1-\theta)^{k-M+1}}{\theta} \bar{\rho}. \end{aligned} \tag{23}$$

Recalling the definition of W_k in (18), it follows from (23) that

$$\mathbb{E}\left\{\|\tilde{x}_{k-M+1}\|^2\right\} \leq \frac{(1-\theta)^{k-M+1}\left(\mathbb{E}\{W_M\} - \frac{\bar{\rho}}{\theta}\right) + \frac{\bar{\rho}}{\theta}}{\min_{a \in \mathcal{S}}\{\sigma_{\min}(P_a)\}}, \tag{24}$$

which implies that the estimation error dynamics is exponentially ultimately bounded in the mean-square sense and an asymptotic upper bound is

$$\rho = \frac{\bar{\rho}}{\theta \min_{a \in \mathcal{S}}\{\sigma_{\min}(P_a)\}}. \tag{25}$$

Now, the proof is complete. ■

B. Design of the moving-horizon estimator under the PCM

Building on the analysis results established in Theorem 2, we now present a theorem that designs the estimator parameter matrix Ω and minimizes the asymptotic upper bound of the estimation error under the given caching probabilities.

Theorem 3: Consider the estimation error dynamics (16) and the PCM characterized by (6) and (7). Suppose that the caching probabilities ξ_r $(r = 1, 2, \ldots, R)$ and the scalar $0 < \theta < 1$ are given. If there exist positive definite matrices P_a $(a \in \mathcal{S})$, positive scalars ϵ_i $(i = 1, 2, 3, 4)$, and positive definite matrix Ω satisfying the following matrix inequalities

$$\Pi_a = \begin{bmatrix} \Pi_a^{11} & \Pi_a^{12} \\ * & \Pi_a^{22} \end{bmatrix} < 0, \tag{26}$$
$$P_a \geq I \tag{27}$$

where

$$\Pi_a^{11} = \mathrm{diag}\{(\theta-1)P_a, -\epsilon_1 I, -\epsilon_2 I, -\epsilon_3 I, -\epsilon_4 I\}$$
$$\Pi_a^{12} = \begin{bmatrix} \Omega G & \Omega & -\bar{\mathcal{H}}_{a,M} & -\tilde{\mathcal{H}}_{a,M} & -\tilde{\mathcal{H}}_{a,M} \end{bmatrix}^T$$
$$\Pi_a^{22} = \bar{P}_a - 2K_a^{-1},$$

then the estimation error is exponentially ultimately bounded in the mean-square sense. In addition, the asymptotic upper bound can be minimized by resorting to the following optimization problem:

$$\min_{P_a, \epsilon_i, \Omega} \frac{\bar{\rho}}{\theta} \tag{28}$$

under the matrix inequality constraints (26) and (27), where $\bar{\rho}$ is defined in (21).

Proof: Noting that $(\bar{P}_a - K_a^{-1})\bar{P}_a^{-1}(\bar{P}_a - K_a^{-1}) \geq 0$, it is easy to obtain that

$$\bar{\Pi}_a \triangleq \begin{bmatrix} \Pi_a^{11} & \Pi_a^{12} \\ * & -K_a^{-1}\bar{P}_a^{-1}K_a^{-1} \end{bmatrix} \leq \Pi_a.$$

Based on the renowned Schur Complement lemma, the condition (17) is satisfied if and only if the following matrix inequalities

$$\bar{\Pi}_a < 0$$

hold. Therefore, according to Theorem 2 as well as the conditions (26) and (27), the estimation error is exponentially ultimately bounded in the mean-square sense and the asymptotic upper bound can be expressed by $\frac{\bar{\rho}}{\theta}$. The proof is now complete. ■

C. Co-design of the moving-horizon estimator and the PCM

In the previous subsection, the caching probabilities ξ_r ($r = 1, 2, \ldots, R$) and the scalar θ are all given in advance. Obviously, the caching probability, as an important parameter in the PCM, would exert a substantial impact on the estimation performance. Next, our focus is on resolving the co-design problem of the estimator and the PCM to minimize the asymptotic upper bound of the estimation error.

Corollary 1: Consider the estimation error dynamics (16). Based on Theorem 3, if the caching probabilities ξ_r ($r = 1, 2, \ldots, R$) and the scalar θ are all designable parameters, then the asymptotic upper bound can be minimized by resorting to the following constrained optimization problem:

$$\begin{aligned} &\min_{\xi_r, \theta, P_a, \epsilon_i, \Omega} \frac{\bar{\rho}}{\theta} \\ &\text{s.t. } (7), (26), (27) \\ &\qquad 0 < \theta < 1, \epsilon_i > 0, P_a > 0, \Omega > 0 \end{aligned} \tag{29}$$

where $\bar{\rho}$ is defined in (21).

Proof: The proof proceeds in a manner similar to that of Theorem 3 and is therefore omitted for brevity. ■

Note that the objective function in (29) is constructed based on an asymptotic upper bound of the estimation error, which captures the impact of the design variables on the worst-case estimation performance. Although the solution obtained from this upper-bound-based formulation may not exactly coincide with that minimizing the actual estimation error, it provides a tractable and meaningful surrogate for the joint design of the caching strategy and the estimator parameter under unknown-but-bounded noises and sensor resolution effects. It should also be noted that the value of $\bar{\rho}$ is indirectly affected by ξ_r and θ via the variables ϵ_i, which are determined by the matrix inequality constraints (26) and (27) involving ξ_r and θ. Moreover, the objective function and constraints in (29) are both nonlinear since the caching probabilities ξ_r ($r = 1, 2, \ldots, R$) and the scalar θ are treated as parameters to be designed. Consequently, it would be difficult (if not impossible) to directly find a solution to the constrained optimization problem formulated in (29). In this paper, similar to the methodology in [13], we will resort to the well-known PSO algorithm and propose a co-design framework that integrates the PSO algorithm and the linear matrix inequality technique.

In the PSO algorithm, each particle, serving as a candidate solution, adjusts its position within the search space based on its own experiences and those of the swarm to pursue the optimal solution [2]. In this paper, L particles are selected. Recalling the second constraint $\sum_{r=1}^{R} \xi_r = 1$ in (7), it is clear that there are only $R-1$ free variables with respect to the caching probabilities. Therefore, the position of the l-th particle ($l = 1, 2, \ldots, L$) can be detailed by

$$p^{(l)} = \begin{bmatrix} \xi_1^{(l)} & \xi_2^{(l)} & \cdots & \xi_{R-1}^{(l)} & \theta^{(l)} \end{bmatrix}^T \tag{30}$$

where the parameter with superscript (l) denotes the corresponding value in the l-th particle. Notably, one has $\xi_R^{(l)} = 1 - \sum_{r=1}^{R-1} \xi_r^{(l)}$.

Now, we proceed to define the following fitness function:

$$f(p^{(l)}) = \begin{cases} \bar{\rho}_\theta^{(l)}, & \text{if all the constraints in (29) are satisfied} \\ +\infty, & \text{otherwise} \end{cases} \tag{31}$$

where $\bar{\rho}_\theta^{(l)}$ is the obtained minimal value of $\frac{\bar{\rho}}{\theta}$ given the l-th particle.

Following the standard PSO algorithm, the position update of the l-th particle can be governed by [2]:

$$\begin{aligned} \omega^{(l)}(\iota) = \chi\omega^{(l)}(\iota-1) &+ w_1\tau_1\left(p_{ib}^{(l)}(\iota-1) - p^{(l)}(\iota-1)\right) \\ &+ w_2\tau_2\left(p_{gb}(\iota-1) - p^{(l)}(\iota-1)\right) \end{aligned} \tag{32}$$

$$p^{(l)}(\iota) = p^{(l)}(\iota-1) + \omega^{(l)}(\iota) \tag{33}$$

where $\omega^{(l)}(\iota)$ and $p^{(l)}(\iota)$ denote, respectively, the velocity and the position of the l-th particle at the ι-th iteration; χ stands for the inertia weight; w_1 and w_2 are acceleration coefficients; τ_1 and τ_2 are uniformly distributed random variables within the interval $[0, 1]$; $p_{ib}^{(l)}$ signifies the historical best position found by the l-th particle; and p_{gb} indicates the historical best position found by the entire swarm.

Next, the proposed PSO-based co-design scheme for the moving-horizon estimator and the PCM is summarized in Algorithm 1. Note that the PSO algorithm is employed only in the offline design stage, and its computational cost is mainly determined by the number of particles and iterations, as well as the problem dimension. Once the optimization is completed, the designed parameters can be directly applied online in real time. Specifically, the received measurements are cached according to the designed probabilities ξ_r, and the state estimator (12) is executed using the current measurements y_k, the cached measurements Y_{k-1}^a, and the designed parameter matrix Ω. Therefore, the proposed state estimation scheme is suitable for online implementation.

Algorithm 1 PSO-based co-design algorithm for moving-horizon estimator and PCM.

Input: The number of selected particles L, the maximum allowable number of iterations $\iota_{\max}$, and other parameters χ, w_1 and w_2.

1: Initialize $p^{(l)}(0)$ and $\omega^{(l)}(0)$ for each particle $l = 1, 2, \ldots, L$.
2: **for** $\iota = 1$ to $\iota_{\max} + 1$ **do**
3: **for** $l = 1$ to L **do**
4: Calculate the fitness function $f\left(p^{(l)}(\iota-1)\right)$ of the l-th particle according to (31).
5: Update the individual best position $p_{ib}^{(l)}(\iota-1)$ based on the values of $f\left(p^{(l)}(s)\right)$ for $0 \leq s \leq \iota - 1$.
6: **end for**
7: Update the global best position $p_{gb}(\iota-1)$ based on the values of $f\left(p^{(l)}(s)\right)$ for $0 \leq s \leq \iota - 1$ and $1 \leq l \leq L$.
8: **if** $\iota \leq \iota_{\max}$ **then**
9: **for** $l = 1$ to L **do**
10: Update the position $p^{(l)}(\iota)$ and the velocity $\omega^{(l)}(\iota)$ for the l-th particle based on (32) and (33).
11: **end for**
12: **end if**
13: **end for**
14: According to the found global best position $p_{gb}(\iota_{\max})$, determine the optimal values for ξ_r and θ.
15: According to the obtained parameters, design the matrix Ω for the moving-horizon estimator by resorting to Theorem 3.

Output: The designed caching probabilities ξ_r ($r = 1, 2, \ldots, R$) for the PCM and the designed parameter matrix Ω for the estimator.

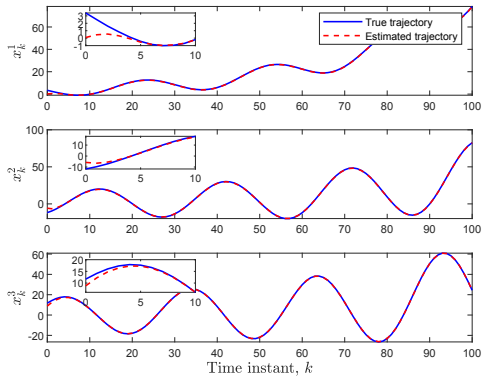

Fig. 2: True and estimated trajectories of x_k.

IV. Numerical Simulations

In this section, two examples are provided to showcase the feasibility and effectiveness of the proposed moving-horizon state estimator under the PCM.

Example 1: The discrete-time linear system (1) under consideration is composed of four sensors (i.e., $N = 4$) and is specified by the following parameters:

$$G = \begin{bmatrix} 1.02 & 0.05 & -0.03 \\ -0.04 & 1.01 & 0.2 \\ 0.1 & -0.2 & 0.98 \end{bmatrix},$$
$$H_1 = \begin{bmatrix} 0.18 & 0.01 & 0 \end{bmatrix}, \ H_2 = \begin{bmatrix} 0 & 0.2 & 0 \end{bmatrix},$$
$$H_3 = \begin{bmatrix} 0.5 & 0 & 0 \end{bmatrix}, \ H_4 = \begin{bmatrix} 0 & 0.04 & 0.8 \end{bmatrix}.$$

The process and measurement noises μ_k and $\nu_{i,k}$ $(i = 1, 2, 3, 4)$ are, respectively, generated from uniform distributions over the intervals $[-0.05,\ 0.05]$ and $[-0.1,\ 0.1]$. Then, the values of the noise-related parameters $\mu^2_{\max}$ and $\nu^2_{i,\max}$ are 7.5×10^{-3} and 0.01, respectively. The size of the window is selected as $M = 2$, and the sensor resolutions are set as $m^1_i = 0.15$ $(i = 1, 2, 3, 4)$.

Considering the limited caching capacity, only the measurements of two sensors are allowed to be stored in the cache, i.e., $S = 2$. According to (4), there are a total of six possible caching modes in this scenario. More specifically, one has

$$\bar{\Xi}_1 = \mathrm{diag}\{1, 1, 0, 0\},\ \bar{\Xi}_2 = \mathrm{diag}\{1, 0, 1, 0\},$$
$$\bar{\Xi}_3 = \mathrm{diag}\{1, 0, 0, 1\},\ \bar{\Xi}_4 = \mathrm{diag}\{0, 1, 1, 0\},$$
$$\bar{\Xi}_5 = \mathrm{diag}\{0, 1, 0, 1\},\ \bar{\Xi}_6 = \mathrm{diag}\{0, 0, 1, 1\}.$$

By implementing Algorithm 1 with parameters $L = 100$, $\iota_{\max} = 80$, $\chi = 0.9$, $w_1 = 0.6$ and $w_2 = 0.4$, the caching probabilities for the PCM and the parameter matrix for the moving-horizon estimator are obtained as follows:

$$\xi_1 = 0.0003,\ \xi_2 = 0.0001,\ \xi_3 = 0.1245,$$
$$\xi_4 = 0.0002,\ \xi_5 = 0.2352,\ \xi_6 = 0.6398,$$
$$\Omega = \begin{bmatrix} 1.1505 & -0.2044 & -0.2266 \\ -0.2044 & 1.0496 & 0.6817 \\ -0.2266 & 0.6817 & 2.6830 \end{bmatrix}.$$

Based on the proposed scheme with these designed parameters, the simulation results are displayed in Fig. 2, which intuitively illustrate the effectiveness of the proposed state estimation strategy.

To further showcase the superiority of the proposed scheme, we compare its performance against the following three schemes: 1) *Scheme I*: each caching mode is randomly activated with equal probability; 2) *Scheme II*: all measurements, including the latest ones, are randomly cached; and 3) *Scheme III*: the measurements of all sensors within a smaller window $M = 1$ are available, meaning that the measurements at time instants k and $k-1$ are all retained. For a fair comparison, the parameters in *Schemes II* and *III* are also optimized using the PSO algorithm.

The simulation results, in terms of estimation errors over 1000 Monte Carlo runs, are provided in Table I, from which we are able to see that the proposed scheme exhibits the best estimation performance among all the compared schemes. This aligns with our expectations, since in the proposed scheme: 1) higher caching probabilities are assigned to the caching modes offering more information; 2) all the latest measurements are guaranteed to be available; and 3) a larger window allows the estimator to more accurately capture the system's dynamic behavior. These results confirm that the proposed state estimation algorithm, together with the joint design of parameters, can deliver good performance.

TABLE I: Comparison under different schemes.

Scheme	Ours	I	II	III
$\mathbb{E}\{\sum_{k=0}^{100} \|\tilde{x}_k\|\}$	44.5513	49.2058	47.6757	50.0976

Example 2: Consider the problem of tracking a moving target using a multi-sensor system with six sensors (i.e., $N = 6$). The target's motion is described by a linear dynamic model with the following state transition matrix [29]:

$$G = \begin{bmatrix} 1 & \tau_s & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tau_s \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\tau_s = 1$ signifies the sampling period, and the state vector, defined by $x_k \triangleq [p_{x,k}, v_{x,k}, p_{y,k}, v_{y,k}]^T$, contains the position and velocity components along the X and Y axes.

The measurement matrices are specified as follows:

$$H_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \ H_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}, \ H_3 = \begin{bmatrix} 0.4 & 0 & 0.8 & 0 \end{bmatrix}$$
$$H_4 = \begin{bmatrix} 0.7 & 0 & 0.5 & 0 \end{bmatrix}, \ H_5 = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \ H_6 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

The process and measurement noises μ_k and $\nu_{i,k}$ are, respectively, generated from uniform distributions over the intervals $[-0.2,\ 0.2]$ and $[-0.05,\ 0.05]$. The size of the window is selected as $M = 3$, and the sensor resolutions are set as $m^1_i = 0.01$.

Given the capacity constraints, only one sensor's measurement is allowed to be stored in the cache (i.e., $S = 1$), yielding six possible caching modes in this setup. Based on Algorithm 1 with parameters $L = 100$, $\iota_{\max} = 60$, $\chi = 0.9$, $w_1 = 0.6$ and $w_2 = 0.4$, the caching probabilities for the PCM are obtained by

$$\xi_1 = 0.1952,\ \xi_2 = 0.1408,\ \xi_3 = 0.2032,$$
$$\xi_4 = 0.2240,\ \xi_5 = 0.0909,\ \xi_6 = 0.1460.$$

The comparative results in terms of root mean square errors (RMSEs) of the position and velocity estimates at time instant k, averaged over 1000 Monte Carlo runs, are displayed in Fig. 3, which further demonstrate the practicality and effectiveness of the proposed scheme.

V. Conclusions

In this paper, the moving-horizon state estimation problem has been handled for linear discrete-time systems subject to PCM and sensor resolution effects. A mathematical model, incorporating the caching modes and the corresponding caching probabilities, has been carefully

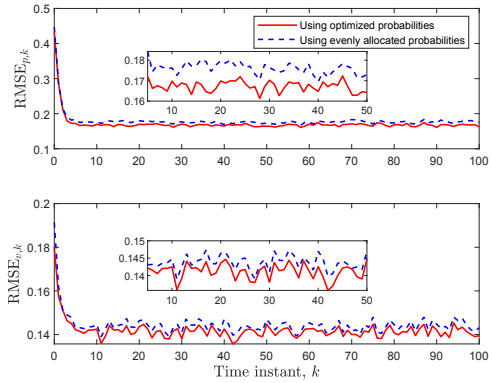

Fig. 3: Behaviors of $\mathrm{RMSE}_{p,k}$ and $\mathrm{RMSE}_{v,k}$.

established to capture the behavior of the PCM in the context of multi-sensor systems. By formulating a least-squares optimization problem, the estimator has been designed to exploit both the newly received measurements and the previously cached measurements. Based on the analysis of the estimation error dynamics, sufficient conditions have been rigorously derived to guarantee that the estimation error is exponentially ultimately bounded in the mean-square sense. Moreover, by resorting to the PSO method with customized particle structure and fitness function, a constrained optimization problem has been solved to co-design the caching probabilities and the estimator parameter matrix, thereby enhancing the estimation performance. Finally, simulation results have been presented to confirm the viability and effectiveness of the proposed scheme. Future studies would focus on extending the current results to time-varying or nonlinear systems, designing adaptive strategies capable of dynamically adjusting caching probabilities, and exploring the interaction between caching mechanisms and communication constraints.

REFERENCES

[1] A. Alessandri and M. Awawdeh, Moving-horizon estimation with guaranteed robustness for discrete-time linear systems and measurements subject to outliers, *Automatica*, vol. 67, pp. 85–93, 2016.

[2] D. Bratton and J. Kennedy, Defining a standard for particle swarm optimization, in *Proceedings of the 2007 IEEE Swarm Intelligence Symposium*, Honolulu, HI, USA, 2007, pp. 120–127.

[3] R. Caballero-Águila, J. Hu and J. Linares-Pérez, Distributed estimation for uncertain systems subject to measurement quantization and adversarial attacks, *Information Fusion*, vol. 120, art. no. 103044, 2025.

[4] R. Caballero-Águila and J. Linares-Pérez, Quadratic estimation for stochastic systems in the presence of random parameter matrices, time-correlated additive noise and deception attacks, *Journal of the Franklin Institute*, vol. 360, no. 15, pp. 11141–11164, 2023.

[5] B. Chen, W.-A. Zhang and L. Yu, Distributed finite-horizon fusion Kalman filtering for bandwidth and energy constrained wireless sensor networks, *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 797–812, 2014.

[6] Z. Chen, N. Pappas and M. Kountouris, Probabilistic caching in wireless D2D networks: Cache hit optimal versus throughput optimal, *IEEE Communications Letters*, vol. 21, no. 3, pp. 584–587, 2017.

[7] P. Chennakesavula, Y.-W. P. Hong and A. Scaglione, Efficient caching by linear compression for parameter estimation in wireless sensor networks, *IEEE Transactions on Signal Processing*, vol. 70, pp. 1155–1169, 2022.

[8] P. Duan, Z. Duan, G. Chen and L. Shi, Distributed state estimation for uncertain linear systems: A regularized least-squares approach, *Automatica*, vol. 117, art. no. 109007, 2020.

[9] J. Gao, S. Zhang, L. Zhao and X. Shen, The design of dynamic probabilistic caching with time-varying content popularity, *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1672–1684, 2021.

[10] B. Jedari, G. Premsankar, G. Illahi, M. D. Francesco, A. Mehrabi and A. Ylä-Jääski, Video caching, analytics, and delivery at the wireless edge: A survey and future directions, *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 431–471, 2021.

[11] L. Ji, J. B. Rawlings, W. Hu, A. Wynn and M. Diehl, Robust stability of moving horizon estimation under bounded disturbances, *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3509–3514, 2016.

[12] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob and A. Ahmed, Edge computing: A survey, *Future Generation Computer Systems*, vol. 97, pp. 219–235, 2019.

[13] J.-Y. Li, Z. Wang, R. Lu and Y. Xu, A component-based coding-decoding approach to set-membership filtering for time-varying systems under constrained bit rate, *Automatica*, vol. 152, art. no. 110874, 2023.

[14] W. Li, Y. Jia and J. Du, Resilient filtering for nonlinear complex networks with multiplicative noise, *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2522–2528, 2019.

[15] A. Liu, W.-A. Zhang, M. Z. Q. Chen and L. Yu, Moving horizon estimation for mobile robots with multirate sampling, *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1457–1467, 2017.

[16] X. Lv, P. Duan, Z. Duan, G. Chen and L. Shi, Stochastic event-triggered variational Bayesian filtering, *IEEE Transactions on Automatic Control*, vol. 68, no. 7, pp. 4321–4328, 2023.

[17] Y. Ma, L. Dai, H. Yang, J. Zhao, R. Gao and Y. Xia, Cloud-edge cooperative MPC for large-scale complex systems with input nonlinearity, *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 3835–3851, 2025.

[18] M. A. Müller, Nonlinear moving horizon estimation in the presence of bounded disturbances, *Automatica*, vol. 79, pp. 306–314, 2017.

[19] G. Premsankar, M. Di Francesco and T. Taleb, Edge computing for the Internet of Things: A case study, *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1275–1284, 2018.

[20] Y. Shen, Z. Wang, H. Dong, H. Liu and X. Liu, Joint state and unknown input estimation for a class of artificial neural networks with sensor resolution: An encoding-decoding mechanism, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 36, no. 2, pp. 3671–3681, 2025.

[21] Z. Sun, C. Han and X. Hu, Receding horizon estimation for multi-rate sampled data systems under component-based event-triggered mechanisms: Handling delayed and degraded measurements, *IET Control Theory & Applications*, vol. 18, no. 13, pp. 1699–1709, 2024.

[22] S. Tarnoi, V. Suppakitpaisarn, W. Kumwilaisak and Y. Ji, Performance analysis of probabilistic caching scheme using Markov chains, in *Proceedings of 2015 IEEE 40th Conference on Local Computer Networks*, Clearwater Beach, FL, USA, 2015, pp. 46–54.

[23] S. Wang, L. Chen, D. Gu and H. Hu, An optimization based moving horizon estimation with application to localization of autonomous underwater vehicles, *Robotics and Autonomous Systems*, vol. 62, no. 10, pp. 1581–1596, 2014.

[24] X. Xia, F. Chen, J. Grundy, M. Abdelrazek, H. Jin and Q. He, Constrained app data caching over edge server graphs in edge computing environment, *IEEE Transactions on Services Computing*, vol. 15, no. 5, pp. 2635–2647, 2022.

[25] Y. Xu, R. Lu, P. Shi, J. Tao and S. Xie, Robust estimation for neural networks with randomly occurring distributed delays and Markovian jump coupling, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 845–855, 2018.

[26] Y. Xu, J. Zhou, H. Rao, R. Lu and L. Xie, Reset moving horizon estimation for quantized discrete time systems, *IEEE Transactions on Automatic Control*, vol. 66, no. 9, pp. 4199–4205, 2021.

[27] Y. Yang, Z. Li, X. Zhao and Y. Li, Worst-case performance of Kalman filtering against cyber-attacks with true and contaminated information, *Automatica*, vol. 176, art. no. 112257, 2025.

[28] E. Yaz and A. Azemi, Observer design for discrete and continuous non-linear stochastic systems, *International Journal of Systems Science*, vol. 24, no. 12, pp. 2289–2302, 1993.

[29] D. Yu, Y. Xia and D.-H. Zhai, Distributed moving-horizon estimation with event-triggered communication over sensor networks, *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 7982–7988, 2023.

[30] J. Zhang, X. He and D. Zhou, Filtering for stochastic uncertain systems with non-logarithmic sensor resolution, *Automatica*, vol. 89, pp. 194–200, 2018.

[31] L. Zou, Z. Wang, Q.-L. Han and D. Zhou, Moving horizon estimation for networked time-delay systems under Round-Robin protocol, *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 5191–5198, 2019.

[32] L. Zou, Z. Wang, Q.-L. Han and D. Zhou, Moving horizon estimation of networked nonlinear systems with random access protocol, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 5, pp. 2937–2948, 2021.