Communication– and Control–aware Optimal Quantizer Selection for Multi-agent Control

Mohammad Afshari, Dipankar Maity, and Panagiotis Tsiotras

Abstract— **We consider a multi-agent linear quadratic optimal control problem. Due to communication constraints, the agents are required to quantize their local state measurements before communicating them to the rest of the team, thus resulting in a decentralized information structure. The optimal controllers are to be synthesized under this decentralized and quantized information structure. The agents are given a set of quantizers with varying quantization resolutions—higher resolution incurs higher communication cost and vice versa. The team must optimally select the quantizer to prioritize agents with 'high-quality' information for optimizing the control performance under communication constraints. We show that there exist a sepatation between the optimal solution to the control problem and the choice of the optimal quantizer. We show that the optimal controllers are linear and the optimal selection of the quantizers can be determined by solving a linear program.**

I. Introduction

Networked control systems are widely used in various applications, such as sensor networks, intelligent transportation systems, self-deriving vehicles, and robotics [1]. These systems often employ quantization to reduce the communication bandwidth required to close the feedback loop from the sensor to the controller [2]–[5]. For multi-agent systems with multiple controllers and sensors, the need for quantization is even more pronounced to judiciously utilize communication resources. The quantization process aims to strike a balance between control performance and communication constraints. Higher resolution quantizers incur less quantization error, leading to better control performance but at the expense of larger communication bandwidth required to transmit their output. Conversely, coarser quantizers require fewer bits to be transmitted but result in degraded control performance.

While the trade-off between quantization bit-rate and optimal control performance for single-agent systems has been investigated [6]–[8], this trade-off for multi-agent systems is not equally well understood. This knowledge gap primarily stems from the fact that determining the optimal design for the quantizer and the controller, even for a linear-quadratic single agent, is a computationally intractable problem [5], [8]. For multi-agent systems, the problem becomes significantly more challenging due to the decentralized information structure [9]–[11].

The work was supported by the ARL grant DCIST CRA W911NF-17-2-0181.

M. Afshari and P. Tsiotras are with the D. Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0150, USA. Email: {mafshari, tsiotras}@gatech.edu

Dipankar Maity is with Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, Charlotte, NC, 28223-0001, USA. Email: dmaity@charlotte.edu

For a single agent, the primary challenge lies in the design of the quantizers. While several works have been able to characterize the optimal controller, the optimal quantizer design problem remains to be unsolved. Recently, the works in [12], [13] considered a different formulation where quantizers are not designed but rather chosen from a given set. These works, primarily focusing on the single agent case, aimed to design the optimal controller and then select the optimal quantizer to minimize a weighted cost function combining control and communication costs.

In this paper, we adopt the framework of [12] and extend it to the multi-agent case. Here, each agent must select the optimal quantizer at each time instance to maintain a balance between control performance and communication constraints/costs. While the agents share quantized states with the team, they retain the true state values to themselves, thus resulting in a decentralized information structure.

The contribution and significance of this work lie in deriving the optimal controller and the optimal selection of the quantizers in decentralized settings. We show that the optimal controller for each agent has two components: one that depends on the common information communicated by each agent to others, and another that solely depends on the local information each agent processes. We show that the optimal selection of the quantizers is time-varying for finite-horizon problems, and it can be determined by solving a linear program.

The rest of the paper is organized as follows: We formulate the problem in Section II. We discuss the decentralized information structure and the quantization scheme in Section III. The optimal controller is derived in Section IV and the optimal selection of the quantizers are obtained in Section V. Finally, we conclude the work in Section VI.

A. Notations

Given a matrix A, $A \succeq 0$ and $A \succ 0$ denote that A is positive semi-definite and positive definite, respectively. $\text{vec}(v_1, \ldots, v_n)$ denotes the column vector formed by vertically stacking the vectors v_i's. Given any vector valued process $\{y_t\}_{t\geq 0}$ and any time instances t_1, t_2 such that $t_1 \leq t_2$, $y_{t_1:t_2}$ is a shorthand notation for $\text{vec}(y_{t_1}, y_{t_1+1}, \ldots, y_{t_2})$.

II. Problem formulation

Consider a system of n agents (see Fig. 1) evolving in discrete time with linear dynamics. Let $x_t^i \in \mathbb{R}^{d_x}$ denote the state and $u_t^i \in \mathbb{R}^{d_u}$ denote the control action of agent i, $i \in N \coloneqq \{1, 2, \ldots, n\}$ at time t. The dynamics of each

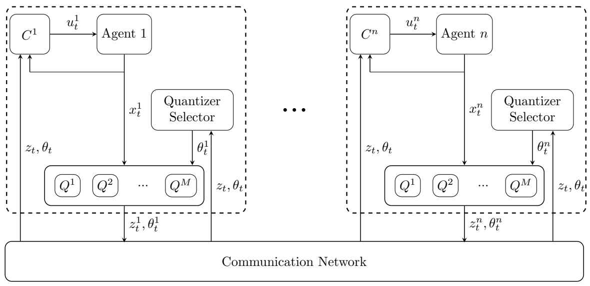

Fig. 1. The n agent model with all-to-all communication framework.

agent are given by

$$x_{t+1}^i = A^i x_t^i + B^i u_t^i + w_t^i, \quad (1)$$

where A^i and B^i are matrices of compatible dimensions and $\{w_t^i\}_{t=0}^{T-1}$ is a zero-mean i.i.d noise process with finite second moment Σ_w^i. We do not assume that w_t^i is Gaussian. The initial state x_0^i is a random vector with zero mean and finite second moment Σ_x^i. For convenience of notation, we often use w_{-1}^i to denote x_0^i.

Assumption 1: For all $i, j \in \{1, 2, \ldots, n\}$ and $t, s \in \{-1, 0, \ldots, T-1\}$, we assume that w_t^i and w_s^j are independent for $i \neq j$ or $t \neq s$.

By concatenating the linear dynamics for all of the agents we may write

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad (2)$$

where $A = \text{diag}(A^1, \ldots, A^n)$, $B = \text{diag}(B^1, \ldots, B^n)$, and $w_t = \text{vec}(w_t^1, \ldots, w_t^n)$. In (2), $x_t = \text{vec}(x_t^1, \ldots, x_t^n)$ and $u_t = \text{vec}(u_t^1, \ldots, u_t^n)$ are the vectors representing the states and controls of all the agents.

Each agent perfectly observes its own state. However, due to communication constraints and limited power of agents in transmitting information (as we will discuss in detail in Section II-A), the agents must use quantizers when transmitting information to the other agents to reduce the communication bandwidth. We assume that a set of M quantizers are provided to quantize the state value for each agent.[1] The symbols of the m-th quantizer are denoted by $\mathcal{Q}^m = \{q_1^m, \ldots, q_{\ell^m}^m\}$. Associated with the m-th quantizer, let $\mathcal{P}^m = \{p_1^m, \ldots, p_{\ell^m}^m\}$ denote a partition in $\mathbb{R}^{d_x}$ such that p_j^m gets mapped to the symbol q_j^m for each $m \in \{1, \ldots, \ell^m\}$. Thus, the m-th quantizer provides a mapping/encoding $\delta^m : \mathbb{R}^{d_x} \to \mathcal{Q}^m$ such that $\delta^m(x) = q_j^m$ if and only if $x \in p_j^m$.

A. System Performance

The control objective for these agents is to jointly minimize the finite-horizon quadratic cost function

$$J_{\text{Control}} = \mathbb{E}\Big[x_T^\mathsf{T} Q x_T + \sum_{t=0}^{T-1} \big(x_t^\mathsf{T} Q x_t + u_t^\mathsf{T} R u_t\big)\Big], \quad (3)$$

where $Q \succeq 0$ and $R \succ 0$. Let $Q^{ij} \in \mathbb{R}^{d_x \times d_x}$ denote the ij-th block element of Q that couples the states of agents i and j via the term $(x_t^i)^\mathsf{T} Q^{ij} x_t^j$. Similarly, we define R^{ij} to be the ij-th block element of R.

Although the agents' dynamics are decoupled (see (1)), the objective function (3) couples the states and control actions. Hence, the optimal control for each agent depends on the global state x_t, which necessitates each agent to share their local state information x_t^i with the other agents to increase performance. In this work, we assume that agents can broadcast their messages to the entire team (i.e., an all-to-all communication architecture), to better coordinate and lower the control cost of the whole team. However, the agents must use quantizers to judiciously use the communication resources (e.g., bandwidth). In other words, the agents must be prioritized to use the communication resources based on how their state information helps in reducing the global objective function (3).

The (communication) cost of using the m-th quantizer is $\lambda^m > 0$. For instance, $\lambda^m = \log_2(\ell^m)$ denotes the number of bits required to transmit the quantized message from the m-th quantizer. Let us define the new decision variable

$$\theta_t^{im} = \begin{cases} 1, & \text{agent } i \text{ selects quantizer } m \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

In general, the quantizer selection policy could be a randomized policy and, therefore, the outcomes θ_t^{im} are random variables. Hence, the expected communication cost for the entire team at time t can be expressed as

$$J_{\text{Comm},t} = \sum_{i=1}^{n} \sum_{m=1}^{M} \mathbb{E}[\theta_t^{im}]\lambda^m = \sum_{i=1}^{n} (\mathbb{E}[\theta_t^i])^\mathsf{T} \lambda, \quad (5)$$

where $\theta_t^i = \text{vec}(\theta_t^{i1}, \ldots, \theta_t^{iM})$ and $\lambda = \text{vec}(\lambda^1, \ldots, \lambda^M)$. The total communication cost for the entire horizon therefore becomes

$$J_{\text{Comm}} = \sum_{t=0}^{T-1} J_{\text{Comm},t}. \quad (6)$$

For the communication constrained control problem, we will consider the following three variations.

Per-time communication constraint:

$$\begin{aligned} &\min \quad J_{\text{Control}} \\ &\text{subject to } J_{\text{Comm},t} \leq c_t, \qquad t = 0, \ldots, T-1, \end{aligned} \quad \text{(P1)}$$

for given time-varying communication budgets c_t's.

Cumulative communication constraint:

$$\begin{aligned} &\min \quad J_{\text{Control}} \\ &\text{subject to } J_{\text{Comm}} \leq c, \end{aligned} \quad \text{(P2)}$$

for a given cumulative budget $c > 0$.

Weighted cost formulation:

$$\min \quad J_{\text{Control}} + \alpha J_{\text{Comm}}, \quad \text{(P3)}$$

where $\alpha \geq 0$ is a trade-off parameter between the communication and the control costs. The optimization problems in (P1)-(P3) are carried out w.r.t the control variables (i.e., $u_{0:T-1}^{1:n}$) and the quantizer selection variables (i.e., $\theta_{0:T-1}^{1:n}$).

[1]The analysis remains the same when different agents have different sets of quantizers.

In this work, we assume that each agent may select only one quantizer at any given time, which imposes the constraint

$$\sum_{m=1}^{M} \theta_t^{im} = 1, \tag{7}$$

for all $i = 1, \ldots, n$ and $t = 0, \ldots, T-1$.

III. Information Structure and Quantization Scheme

We denote the quantized measurement of agent i at time t as z_t^i. If agent i uses the m-th quantizer to quantize x_t^i, then $z_t^i = \delta^m(x_t^i)$. Using the quantizer selection variables θ_t^{im} defined in (4), we may also express z_t^i as

$$z_t^i = \sum_{m=1}^{M} \theta_t^{im} \delta^m(x_t^i),$$

which explicitly shows how the choice of the quantizer (i.e., θ_t^{im}) affects z_t^i.

Remark 1: Instead of quantizing x_t^i, one may quantize any other signal, say ξ_t^i, that depends on the entire history $x_{0:t}^i, u_{0,t-1}^i$ at time t. In fact, often one quantizes $\xi_t^i = x_t^i - \sum_{j=0}^{t-1}(A^i)^{t-1-j}B^i u_j^i$, which is known as *predictive coding* [14, Definition 3.1]. In this work, we also do not quantize x_t^i, as will be explained later.

While each agent shares its quantized state with others, it retains the true state locally and may use it for synthesizing its control inputs. Therefore, our problem formulation has a decentralized information structure. At time t, agent i observes its own state and selects the quantizer θ_t^i to broadcast z_t^i to all agents. Next, the agents use the broadcast information to take optimal actions to solve (P1)-(P3). The information available to agent i prior to quantization and communication at time t is

$$I_{t^-}^i = \{x_{0:t}^i, u_{0:t-1}^i, z_{0:t-1}, \theta_{0:t-1}\}; \quad I_{0^-}^i = \{x_0^i\}, \tag{8}$$

where t^- indicates that $I_{t^-}^i$ is the available information prior to any decision taken (on control or quantizer selection) at time t, and $z_t \triangleq \text{vec}(z_t^1, \ldots, z_t^n)$ is the vector created by concatenating all communicated signals and $\theta_t \triangleq \text{vec}(\theta_t^1, \ldots, \theta_t^n)$ is the concatenation of all quantizer choices. After the quantized measurements are received by the agents, the available information to agent i is

$$I_t^i = \{x_{0:t}^i, u_{0:t-1}^i, z_{0:t}, \theta_{0:t}\} = I_{t^-}^i \cup \{z_t, \theta_t\}. \tag{9}$$

We may split the information I_t^i into two parts: The information available to all agents, i.e., the *common* information, and the information available to each individual agent, i.e., the *local* information. We denote the *common* and *local* information as I_t^c and $I_t^{i,l}$, respectively:

$$I_t^c = \{\theta_{0:t}, z_{0:t}\}, \tag{10}$$

$$I_t^{i,\ell} = \{x_{0:t}^i, u_{0:t-1}^i\}. \tag{11}$$

Agent i's controller is a measurable function of the local information I_t^i, whereas the quantizer selector is I_{t-1}^c measurable. One may notice that the information set I_t^i is equivalent to the information set $\{w_{-1:t-1}^i, u_{0:t-1}^i, z_{0:t}, \theta_{0:t}\}$, which is expressed in terms of the primitive variables w_t^i's.

In this work, we restrict ourselves to the *innovation quantization* framework where each agent shares a quantized version of w_{t-1}^i instead of x_t^i at time t. When x_t^i is quantized and shared, the optimal controller synthesis becomes an intractable problem even for a single agent case. This issue becomes significantly more complicated for the decentralized multi-agent case considered in this work. A detailed discussion on quantization of w_t^i instead x_t^i can be found in earlier literature [2] and in our recent works [12], [13]. Therefore, from this point onward, for all $t = 0, \ldots, T-1$, we will consider

$$z_t^i = \sum_{m=1}^{M} \theta_t^{im} \delta^m(w_{t-1}^i). \tag{12}$$

Remark 2: Due to the restriction imposed by (12) (i.e., quantizing w_{t-1}^i instead of x_t^i) we may lose optimality. It is noteworthy that there is no such loss of optimality if we quantize $\sum_{s=0}^{t}(A^i)^{t-s}w_{s-1}^i$ instead of w_{t-1}^i at time t; see for instance [14, Lemma 3.1]. Quantizing/encoding $\sum_{s=0}^{t}(A^i)^{t-s}w_{s-1}^i$ is known as *predictive coding*, where the quantizer removes the contribution of the control before quantization. A brief discussion on the trade-off between computational tractability and optimality for considering (12) instead of predictive coding can be found in [13]. Studying our proposed multi-agent problem in the predictive coding setup is a promising and challenging future direction.

Remark 3: If each agent has a partial and noisy state measurement $y_t^i = C^i x_t^i + v_t^i$, one would need to quantize the *innovation signal* $\xi_t^i = x_t^i - \mathbb{E}[x_t^i \mid y_{0:t}^i, u_{0:t-1}^i]$, as mentioned in [13]. One may verify that $\xi_t^i = w_{t-1}^i$ when $C_t^i = I$ and $v_t^i = 0$, which is the case considered in this work. We refer to our quantization scheme as *innovation quantization* since we quantize the innovation signal.

Under the *innovation quantization* scheme (12), our objective is to find the optimal controller and quantizer selector strategies for each agent to solve the optimization problems in (P1)–(P3).

IV. Optimal Controller

In standard linear-quadratic optimal control problems, the solution typically consists of two components: a state estimator and a feedback gain. The state estimator depends on available information, while the feedback gain depends on system matrices through Riccati equations. However, for multi-agent systems, this is not always the case. [9], [15] showed that the solution is not necessarily linear, and even with linear strategies, certainty equivalent controllers might not be optimal. In our paper, we show that the optimal controller is indeed a certainty equivalent controller.

Given that we have both local and common information, we define the estimators and the Riccati equations upfront for subsequent uses. To that end, following [16], [17] we define the following estimates based on the common information

$$\hat{u}_t = \mathbb{E}[u_t | I_t^c], \quad \hat{x}_t = \mathbb{E}[x_t | I_t^c]. \tag{13}$$

Additionally, we also define following error variables with respect to the conditional expectations defined above

$$\tilde{x}_t = x_t - \hat{x}_t, \tag{14a}$$
$$\tilde{u}_t = u_t - \hat{u}_t. \tag{14b}$$

Lemma 1: The state estimates and estimation errors evolve as follows:

$$\hat{x}_{t+1} = A\hat{x}_t + B\hat{u}_t + \hat{w}_t, \tag{15}$$
$$\tilde{x}_{t+1} = A\tilde{x}_t + B\tilde{u}_t + \tilde{w}_t, \tag{16}$$

where $\hat{w}_t = \mathbb{E}[w_t \mid z_{t+1}, \theta_{t+1}]$ and

$$\tilde{w}_t = \begin{bmatrix} w_t^1 - \mathbb{E}[w_t^1 \mid z_t^1, \theta_{t+1}^1] \\ \vdots \\ w_t^n - \mathbb{E}[w_t^n \mid z_t^n, \theta_{t+1}^n] \end{bmatrix}. \tag{17}$$

Proof: The proof is presented in Appendix A. ■

We define a global Riccati equation whose solution (P_t) is used by all the agents in their controllers, and we also define local Riccati equations ($\tilde{P}_t^i$) for each agent, as follows.

$$\begin{aligned} &P_t = Q + A^\top P_{t+1} A - L_t^\top (R + B^\top P_{t+1} B) L_t, \\ &P_T = Q, \\ &L_t = (R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A, \end{aligned} \tag{18}$$

and for each individual agent i, we define

$$\begin{aligned} &\tilde{P}_t^i = Q^{ii} + (A^i)^\top \tilde{P}_{t+1}^i A^i - (\tilde{L}_t^i)^\top (R^{ii} + (B^i)^\top \tilde{P}_{t+1}^i B^i) \tilde{L}_t^i, \\ &\tilde{P}_T^i = Q^{ii}, \\ &\tilde{L}_t^i = (R^{ii} + (B^i)^\top \tilde{P}_{t+1}^i B^i)^{-1} (B^i)^\top \tilde{P}_{t+1}^i A^i. \end{aligned} \tag{19}$$

The main result of this section is summarized in the following theorem.

Theorem 1: The optimal controller for the i-th agent is

$$u_t^i = -L_t^i \hat{x}_t - \tilde{L}_t^i \tilde{x}_t^i, \tag{20}$$

where L_t^i is the i-th block-row of the matrix L_t defined in (18) and $\tilde{L}_t^i$ is defined in (19).

Furthermore, the optimal control cost under (20) is

$$\begin{aligned} J_{\text{Control}} = \operatorname{tr}(P_0 \Sigma_x) &+ \sum_{t=0}^{T-1} \operatorname{tr}(P_{t+1} \Sigma_w) \\ &+ \sum_{i=1}^{n} \sum_{t=0}^{T-1} \mathbb{E}[(\beta_t^i)^\top \theta_t^i], \end{aligned} \tag{21}$$

where β_t^i is a constant given in (29).

Proof: A proof sketch is presented in Appendix C. ■

Using (14), one may express (20) in the form of $u_t^i = -\tilde{L}_t^i x_t^i - G_t^i \hat{x}_t$, where an expression for G_t^i can be obtained from L_t^i and $\tilde{L}_t^i$. This demonstrates that the choice of quantizers (i.e., $\theta_{0:t}^{1:n}$) affects u_t^i only through the term $\hat{x}_t$. Furthermore, it can be verified that when the cost is decoupled (i.e., $Q^{ij} = 0$ and $R^{ij} = 0$), we have $G_t^i = 0$, as expected.

Theorem 1 not only reveals how the optimal controller is affected by the quantization process, but also demonstrates how the control performance (i.e., J_{Control}) is influenced by the choice of quantizers. This enables us to optimize the quantizer selection policy further to minimize J_{Control}, a discussion of which will be provided in Section V. We conclude this section with the following remarks.

Remark 4: The optimal controller for the i-th agent consists of two parts: the $-L_t^i \hat{x}_t$ part that depends on the common information and the part $-\tilde{L}_t^i \tilde{x}_t^i$, that depends on the local information.

Remark 5: In the case of non-quantized communication, the optimal control cost is $\operatorname{tr}(P_0 \Sigma_x) + \sum_{t=0}^{T-1} \operatorname{tr}(P_{t+1} \Sigma_w)$, and therefore, the adverse effects of the quantization on the control performance is quantified by the term $\sum_{i=1}^{n} \sum_{t=0}^{T-1} \mathbb{E}[(\beta_t^i)^\top \theta_t^i]$. A similar observation is also made in [18], where the communication suffered from packet dropouts and delays instead of quantization.

V. Optimal Quantizer Selection

In this section, we derive the optimal quantizer selection strategies for the agents. Let $\mu_t^i(\cdot \mid I_{t-}^i)$ denote the quantizer selection policy, which is assumed to be a randomized policy without loss of any generality. In other words, we have $\mathbb{P}(\theta_t^{im} = 1 \mid I_{t-}^i) = \mu_t^i(m \mid I_{t-}^i)$, for all $m = 1, \ldots, M$. For notational convenience, we define μ_t^{im} to denote $\mu_t^i(m \mid I_{t-}^i)$. For $\mu_t^i(\cdot \mid I_{t-}^i)$ to be a valid randomized strategy, we impose $\sum_{m=1}^{M} \mu_t^{im} = 1$ for all t. Finally, we define $\mu_t = (\mu_t^1, \ldots, \mu_t^n)$ and $\mu_t^i = (\mu_t^{i1}, \ldots, \mu_t^{iM})$.

Optimizing J_{Control} in (21) is equivalent to optimizing the last term only since the first two terms are constants. At this point we consider each of the optimization problems (P1)–(P3) separately and discuss their corresponding optimal quantizer selections.

A. Per-time and Cumulative Communication Constraints

In this section we consider (P1) and (P2) and derive the optimal quantizer selection strategies for these two cases. Using (21) and the definition of μ_t^i, we may rewrite (P1) as

$$\begin{aligned} \min \quad & \sum_{t=0}^{T-1} \sum_{i=1}^{n} (\mu_t^i)^\top \beta_t^i \\ \text{subject to} \quad & \left. \begin{aligned} &\sum_{i=1}^{n} (\mu_t^i)^\top \lambda \le c_t, \\ &\mathbb{1}^\top \mu_t^i = 1, \quad \mu_t^i \ge 0, \end{aligned} \right\} \begin{aligned} &t = 0, \ldots, T-1, \\ &i = 1, \ldots, n, \end{aligned} \end{aligned} \tag{22}$$

which is a linear programming (LP) problem in μ. The constraints $\mathbb{1}^\top \mu_t^i = 1$ and $\mu_t^i \ge 0$ are to ensure that μ_t^i is a valid probability distribution. Since the cost function can be decoupled in t and the constraints are already decoupled, the optimal selection strategy at any given time t can be found by solving the following optimization

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} (\mu_t^i)^\top \beta_t^i \\ \text{subject to} \quad & \left. \begin{aligned} &\sum_{i=1}^{n} (\mu_t^i)^\top \lambda \le c_t, \\ &\mathbb{1}^\top \mu_t^i = 1, \quad \mu_t^i \ge 0, \end{aligned} \right\} \begin{aligned} &t = 0, \ldots, T-1, \\ &i = 1, \ldots, n. \end{aligned} \end{aligned} \tag{23}$$

This results in a linear program and can be solved efficiently.

Remark 6: Although it may appear that the optimal selection of the quantizers at time t is not concerned with the system's future performance, this is not the case. The β_t^i variable encapsulates the effects of the selected quantizer at time t on the future performance.

In a similar fashion, the cumulative communication constrained problem (P2) is expressed as

$$\begin{aligned} &\min \quad \sum_{t=0}^{T-1}\sum_{i=1}^{n}(\mu_t^i)^\mathsf{T}\beta_t^i && (24)\\ &\text{subject to} \quad \left.\begin{aligned}&\sum_{t=0}^{T-1}\sum_{i=1}^{n}(\mu_t^i)^\mathsf{T}\lambda \le c,\\ &\mathbb{1}^\mathsf{T}\mu_t^i = 1, \quad \mu_t^i \ge 0,\end{aligned}\right\} \begin{aligned}&t=0,\dots,T-1,\\ &i=1,\dots,n.\end{aligned} \end{aligned}$$

which is also a linear program. However, unlike the previous case, the optimal choice at time t cannot be decoupled.

B. Weighted Cost Formulation

Following the same steps as in the previous section, the weighted cost formulation yields

$$\min \quad \sum_{t=0}^{T-1}\sum_{i=1}^{n}(\mu_t^i)^\mathsf{T}(\beta_t^i + \alpha\lambda), \tag{25}$$

$$\text{subject to } \mathbb{1}^\mathsf{T}\mu_t^i = 1, \quad \mu_t^i \ge 0, \qquad \begin{aligned}&t=0,\dots,T-1,\\ &i=0,\dots,n.\end{aligned}$$

This problem is particularly interesting as the class of deterministic policies (i.e., $\mu_t^{im} \in \{0,1\}$) always contains the optimal policy. In particular, agent i's optimal quantizer at time t is

$$m^* = \text{argmin}_m\{\beta_t^{im} + \alpha\lambda^m\}. \tag{26}$$

It is noteworthy that (P3) can be though of a Lagrangian relaxation of (P2). Therefore, one might be tempted to solve (P2) via (P3). However, (P3) will always return a deterministic selection policy (for every value of α) which is not necessarily an optimal solution to (P2). In other words, one may not be able to recover the solution of (P2) from (P3) by simply varying α. A detailed discussion on this is beyond the scope of this letter and will be addressed elsewhere.

We conclude the discussion on quantization selection by the remark that the optimal selection strategy can be found by solving a *centralized* linear program. This LP can be solved offline, similar to the computation of the Riccati equations that can be carried out offline as well. This significantly aids the practical implementation of the framework, where one does not need to carry out an online optimization at every time instance.

VI. Conclusions

In this letter, we revisited a decentralized linear-quadratic optimal control problem with communication constraints. We derived the optimal controllers as well as the optimal choice of quantizers for the agents. We analytically quantified the degradation in control performance due to the communication constraints. We demonstrated that the optimal controller can be designed based on the solution of the matrix Riccati equations, while the optimal quantizers can be determined by solving a linear program. Furthermore, this linear program can be further simplified depending on the nature of the communication constraints.

References

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007.

[2] V. S. Borkar and S. K. Mitter, "LQG control with communication constraints," in *Communications, computation, control, and signal processing: a tribute to Thomas Kailath*, pp. 365–373, Springer, 1997.

[3] S. Tatikonda, *Control under Communication Constraints*. PhD thesis, Massachusetts Institute of Technology, 2000.

[4] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 304–307, 2003.

[5] G. N. Nair and R. J. Evans, "Stabilizability of stochastic linear systems with finite feedback data rates," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 413–436, 2004.

[6] S. Tatikonda, A. Sahai, and S. Mitter, "Stochastic linear control over a communication channel," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1549–1561, 2004.

[7] T. Tanaka, K. H. Johansson, T. Oechtering, H. Sandberg, and M. Skoglund, "Rate of prefix-free codes in LQG control systems," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 2399–2403, 2016.

[8] V. Kostina and B. Hassibi, "Rate-cost tradeoffs in control," *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4525–4540, 2019.

[9] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[10] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

[11] S. Yüksel and T. Basar, "Stochastic networked control systems," *AMC*, vol. 10, p. 12, 2013.

[12] D. Maity and P. Tsiotras, "Optimal controller synthesis and dynamic quantizer switching for linear-quadratic-Gaussian systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 382–389, 2021.

[13] D. Maity and P. Tsiotras, "Optimal quantizer scheduling and controller synthesis for partially observable linear systems," *SIAM Journal on Control and Optimization*, vol. 61, no. 4, pp. 2682–2707, 2023.

[14] S. Yüksel, "Jointly optimal LQG quantization and control policies for multi-dimensional systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1612–1617, 2013.

[15] H. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, 1971.

[16] M. Afshari and A. Mahajan, "Optimal local and remote controllers with unreliable uplink channels: An elementary proof," *IEEE Transactions on Automatic Control*, vol. 65, no. 8, pp. 3616–3622, 2020.

[17] M. Afshari and A. Mahajan, "Decentralized linear quadratic systems with major and minor agents and non-gaussian noise," *IEEE Transactions on Automatic Control*, vol. 68, no. 8, pp. 4666–4681, 2023.

[18] D. Maity, M. H. Mamduhi, S. Hirche, and K. H. Johansson, "Optimal LQG control of networked systems under traffic-correlated delay and dropout," *IEEE Control Systems Letters*, vol. 6, pp. 1280–1285, 2021.

[19] K. J. Åström, *Introduction to Stochastic Control Theory*. Courier Corporation, 2012.

Appendix

A. Proof of Lemma 1

We start by

$$\hat{x}_0 = \mathbb{E}[x_0 \mid I_0^c] = \mathbb{E}[x_0 \mid z_0, \theta_0].$$

Next, we use (2) to derive the definition of $\hat{x}_{t+1}$. We have,

$$\begin{aligned}\hat{x}_{t+1} &= \mathbb{E}[x_{t+1} \mid I_{t+1}^c] = \mathbb{E}[Ax_t + Bu_t + w_t \mid I_t^c, z_{t+1}, \theta_{t+1}]\\ &\overset{(a)}{=} \mathbb{E}[Ax_t + Bu_t \mid I_t^c] + \mathbb{E}[w_t \mid z_{t+1}, \theta_{t+1}],\end{aligned}$$

where (a) follows the fact that w_t is independent of I_t^c and (x_t, u_t) are independent of z_{t+1}, θ_{t+1} which only depend on w_t. Finally,

$$\hat{x}_{t+1} = A\hat{x}_t + B\hat{u}_t + \hat{w}_t,$$

where $\hat{w}_t = \mathbb{E}[w_t \mid z_{t+1}, \theta_{t+1}]$. Then, the error process evolves as

$$\begin{aligned}\tilde{x}_{t+1} &= Ax_t + Bu_t + w_t - A\hat{x}_t + B\hat{u}_t + \hat{w}_t \\ &= A\tilde{x}_t + B\tilde{u}_t + \tilde{w}_t,\end{aligned}$$

where $\tilde{w}_t = w_t - \mathbb{E}[w_t \mid z_{t+1}, \theta_{t+1}]$, which by Assumption 1 can be further simplified to

$$\tilde{w}_t = \begin{bmatrix} w_t^1 - \mathbb{E}[w_t^1 \mid z_t^1, \theta_{t+1}^1] \\ \vdots \\ w_t^n - \mathbb{E}[w_t^n \mid z_t^n, \theta_{t+1}^n] \end{bmatrix}. \tag{27}$$

B. Some Useful Lemmas

Lemma 2 (Completion of Square): Given a linear dynamics $x_{t+1} = Ax_t + Bu_t + w_t$, with w_t being independent from (x_t, u_t), we may write

$$\begin{aligned}&\mathbb{E}\Big[x_T^\intercal Q x_T + \sum_{t=0}^{T-1} x_t^\intercal Q x_t + u_t^\intercal R u_t\Big] = \\ &\quad \mathbb{E}\left[\sum_{t=0}^{T-1} (u_t + L_t x_t)^\intercal P_{t+1} (u_t + L_t x_t)\right] + \operatorname{tr}\left(P_0 \mathbb{E}[x_0 x_0^\intercal]\right) \\ &+ \sum_{t=0}^{T-1} \operatorname{tr}\left(P_{t+1} \mathbb{E}[w_t w_t^\intercal]\right)\end{aligned}$$

where P_t follows the Riccati equation (18).

Proof: The proof follows from [19]. ■

Lemma 3 (Conditional Independence): For any fixed feasible control strategy g and quantization strategy δ, the random vectors $x_t^{1:n}, u_t^{1:n}$ are conditionally independent given the common information I_t^c. That is, for the collection of events $E_t^{1:n}$, where E_t^i is any measurable subset of $\mathbb{R}^{d_x+d_u}$,

$$\mathbb{P}\big(\{x_{0:t}^i, u_{0:t}^i \in E_{0:t}^i\}_{i\in N} \mid I_t^c\big) = \prod_{i=1}^n \mathbb{P}\big((x_{0:t}^i, u_{0:t}^i) \in E_{0:t}^i | I_t^c\big)$$

Proof: Arbitrarily fix a control and quantization strategies for all of the agents and define the following sigma algebra:

$$\begin{aligned}\mathcal{F}_t^c &= \sigma(z_{0:t}, \theta_{0:t}), \\ \mathcal{F}_t^i &= \sigma(x_0^i, w_{0:t}^i, z_{0:t}, \theta_{0:t}).\end{aligned}$$

It follows from the independence of $\{x_0^i, w_{0:n}^i\}_{i\in N}$ that $\{\mathcal{F}_t^i\}_{i\in N}$ are conditionally independent given $\mathcal{F}_t^c$. Next, we show the result by induction. At time $t = 0$, the above statement is true because the initial conditions are independent. Suppose this statement is true at time t. At time $t+1$ we have

$$\begin{aligned}&\mathbb{P}\big(\{x_{0:t+1}^i, u_{0:t+1}^i \in E_{0:t+1}^i\}_{i\in N} \mid I_{t+1}^c\big) \\ &\quad = \mathbb{P}\big(\{x_{0:t+1}^i, u_{0:t+1}^i \in E_{0:t+1}^i\}_{i\in N} \mid I_t^c, \{z_{t+1}^i, \theta_{t+1}^i\}_{i\in N}\big) \\ &\quad = \mathbb{P}\big(\{x_{0:t+1}^i, u_{0:t+1}^i \in E_{0:t+1}^i\}_{i\in N} \mid I_t^c, \\ &\qquad\qquad \{\sum_{m=1}^M \theta_{t+1}^{im} \delta^m(w_t^i), \theta_{t+1}^i\}_{i\in N}\big) \\ &\quad = \mathbb{P}\big(\{(\mathcal{G}^x(x_{0:t}^i, w_t^i, I_t^c), \mathcal{G}^u(x_{0:t}^i, I_t^c)) \in E_{t+1}^i, \\ &\qquad x_{0:t}^i, u_{0:t}^i \in E_{0:t}^i\}_{i\in N} \mid I_t^c, \{\sum_{m=1}^M \theta_{t+1}^{im} \delta^m(w_t^i), \theta_{t+1}^i\}_{i\in N}\big),\end{aligned}$$

where $\mathcal{G}^x$ and $\mathcal{G}^u$ are functions given by the choice of the control and quantization strategies. In the last equality, $\{x_{0:t}^i$ and $u_{0:t}^i$ are independent of w_t^i and conditionally independent of I_t^c for $i \neq j$ by the induction assumption. Furthermore, $(x_{0:t}^i, w_t^i, I_t^c) \in \mathcal{F}_t^i$ and hence is conditionally independent of $\mathcal{F}_t^c$. Hence, we can write

$$\begin{aligned}&\mathbb{P}\big(\{(\mathcal{L}^x(x_{0:t}^i, w_t^i, I_t^c), \mathcal{L}^u(x_{0:t}^i, I_t^c)) \in E_{t+1}^i \\ &\qquad , x_{0:t}^i, u_{0:t}^i \in E_{0:t}^i\}_{i\in N} \mid I_t^c, \{\sum_{m=1}^M \theta_{t+1}^{im} \delta^m(w_t^i), \theta_{t+1}^i\}_{i\in N}\big) \\ &\quad = \prod_{i=1}^n \mathbb{P}\big((\mathcal{L}^x(x_{0:t}^i, w_t^i, I_t^c), \mathcal{L}^u(x_{0:t}^i, I_t^c)) \in E_{t+1}^i \\ &\qquad , x_{0:t}^i, u_{0:t}^i \in E_{0:t}^i\}_{i\in N} \mid I_t^c, \{\sum_{m=1}^M \theta_{t+1}^{im} \delta^m(w_t^i), \theta_{t+1}^i\big) \\ &\quad = \prod_{i=1}^n \mathbb{P}\big(x_{0:t+1}^i, u_{0:t+1}^i \in E_{0:t+1}^i\}_{i\in N} \mid I_{t+1}^c\big)\end{aligned}$$

■

Corollary 1: At any given t and for given matrices M_x and M_u,

$$\mathbb{E}[\tilde{x}_t^i M_x \tilde{x}_t^j] = 0, \quad \mathbb{E}[\tilde{u}_t^i M_u \tilde{u}_t^j] = 0. \quad \text{for} \quad i \neq j.$$

for any matrices M_x and M_u with compatible dimensions.

Proof: The proof follows directly from Lemma (3). ■

C. Proof of Theorem 1

We write J_{Control} using the estimates in (13) and (14) as follows:

$$\begin{aligned}J_{\text{Control}} &= \mathbb{E}\Big[\hat{x}_T^\intercal Q \hat{x}_T + \tilde{x}_T^\intercal Q \tilde{x}_T \\ &\quad + \sum_{t=0}^{T-1} \left(\hat{x}_t^\intercal Q \hat{x}_t + \tilde{x}_t^\intercal Q \tilde{x}_t + \hat{u}_t^\intercal R \hat{u}_t + \tilde{u}_t^\intercal R \tilde{u}_t\right)\Big] \\ &= \hat{J}_{\text{Control}} + \tilde{J}_{\text{Control}},\end{aligned}$$

where

$$\begin{aligned}\hat{J}_{\text{Control}} &= \mathbb{E}\Big[\hat{x}_T^\intercal Q \hat{x}_T + \sum_{t=0}^{T-1} \left(\hat{x}_t^\intercal Q \hat{x}_t + \hat{u}_t^\intercal R \hat{u}_t\right)\Big] \\ \tilde{J}_{\text{Control}} &= \mathbb{E}\Big[\tilde{x}_T^\intercal Q \tilde{x}_T + \sum_{t=0}^{T-1} \left(\tilde{x}_t^\intercal Q \tilde{x}_t + \tilde{u}_t^\intercal R \tilde{u}_t\right)\Big],\end{aligned}$$

and one can use the tower rule to show that $\mathbb{E}[\hat{x}_t^\mathsf{T} Q\tilde{x}_t] = \mathbb{E}[\hat{u}_t^\mathsf{T} Q\tilde{u}_t] = \mathbb{E}[\hat{x}_T^\mathsf{T} Q\tilde{x}_T] = 0$. For instance, we have

$$\begin{aligned}\mathbb{E}[\hat{u}_t^\mathsf{T} R\tilde{u}_t] &= \mathbb{E}\left[\mathbb{E}[\hat{u}_t^\mathsf{T} R\tilde{u}_t \mid I_t^c]\right]\\ &\overset{(a)}{=} \mathbb{E}\left[\hat{u}_t^\mathsf{T} R\mathbb{E}[\tilde{u}_t \mid I_t^c]\right]\\ &\overset{(b)}{=} 0\end{aligned}$$

where (a) follows from the fact that $\hat{u}_t$ is I_t^c measurable and (b) follows from (13) and (14). We further use Corollary 1 to show

$$\tilde{J}_{\text{Control}} = \sum_{i=1}^n \mathbb{E}\Big[(\tilde{x}_T^i)^\mathsf{T} Q^{ii}\tilde{x}_T^i + \sum_{t=0}^{T-1}\big((\tilde{x}_t^i)^\mathsf{T} Q^{ii}\tilde{x}_t^i + (\tilde{u}_t^i)^\mathsf{T} R^{ii}\tilde{u}_t^i\big)\Big].$$

Next, we use the following lemma.

Lemma 4: For any given matrix $\hat{M}$ and $\tilde{M}$ with proper dimensions, we can show that

$$\begin{aligned}&\mathbb{E}[\begin{bmatrix}\hat{x}_t^\mathsf{T} & \hat{u}_t^\mathsf{T}\end{bmatrix}\hat{M}\hat{w}_t] = 0\\ &\mathbb{E}[\begin{bmatrix}(\tilde{x}^i)_t^\mathsf{T} & (\tilde{u}_t^i)^\mathsf{T}\end{bmatrix}\tilde{M}\tilde{w}_t^i] = 0\end{aligned}$$

Proof: The proof stems from the fact that, for any fixed quantization and control strategies, $\hat{x}_t$, $\hat{u}_t$, $\tilde{x}_t^i$, and $\tilde{u}_t^i$ are functions of $w_{0:t-1}$, while $\hat{w}_t$ and $\tilde{w}_t$ are functions of w_t. Therefore, by Assumption 1, we can show that

$$\mathbb{E}[\begin{bmatrix}\hat{x}_t^\mathsf{T} & \hat{u}_t^\mathsf{T}\end{bmatrix}\hat{M}\hat{w}_t] = \mathbb{E}[\begin{bmatrix}\hat{x}_t^\mathsf{T} & \hat{u}_t^\mathsf{T}\end{bmatrix}]\hat{M}\mathbb{E}[\hat{w}_t] = 0,$$

where the last equality holds because

$$\mathbb{E}[\hat{w}_t] = \mathbb{E}[\mathbb{E}[w_t|z_{t+1},\theta_{t+1}]] = \mathbb{E}[w_t] = 0.$$

Similarly one can show that

$$\mathbb{E}[\begin{bmatrix}(\tilde{x}_t^i)^\mathsf{T} & (\tilde{u}_t^i)^\mathsf{T}\end{bmatrix}\tilde{M}\tilde{w}_t^i] = \mathbb{E}[\begin{bmatrix}(\tilde{x}_t^i)^\mathsf{T} & (\tilde{u}_t^i)^\mathsf{T}\end{bmatrix}]\tilde{M}\mathbb{E}[\tilde{w}_t^i] = 0,$$

where the last equality holds because $\mathbb{E}[\tilde{w}_t^i] = \mathbb{E}[w_t^i - \hat{w}_t^i] = 0$. ■

Now, we can perform the completion of squares of Lemma 2 to both $\hat{J}_{\text{Control}}$ and $\tilde{J}_{\text{Control}}$. We get

$$\begin{aligned}J_{\text{Control}} &= \operatorname{tr}(P_0\Sigma_{\hat{x}}) + \sum_{t=0}^{T-1}\operatorname{tr}(P_{t+1}\Sigma_{\hat{w}})\\ &+ \sum_{i=1}^n\left[\operatorname{tr}\left(\tilde{P}_0^i\Sigma_{\tilde{x}}^i\right) + \sum_{t=0}^{T-1}\operatorname{tr}\left(\tilde{P}_{t+1}^i\Sigma_{\tilde{w}}^i\right)\right]\\ &+ \mathbb{E}\left[\sum_{t=0}^{T-1}(\hat{u}_t + L_t\hat{x}_t)^\mathsf{T} P_{t+1}(\hat{u}_t + L_t\hat{x}_t)\right]\\ &+ \sum_{i=1}^n\mathbb{E}\left[\sum_{t=0}^{T-1}(\tilde{u}_t^i + \tilde{L}_t^i\tilde{x}_t^i)^\mathsf{T}\tilde{P}_{t+1}^i(\tilde{u}_t^i + \tilde{L}_t^i\tilde{x}_t^i)\right],\end{aligned}$$

where

$$\begin{aligned}&\Sigma_{\hat{x}} \triangleq \mathbb{E}[\hat{x}_0(\hat{x}_0)^\mathsf{T}] \quad \Sigma_{\hat{w}} \triangleq \mathbb{E}[\hat{w}_t(\hat{w}_t)^\mathsf{T}]\\ &\Sigma_{\tilde{x}} \triangleq \mathbb{E}[\tilde{x}_t(\tilde{x}_t)^\mathsf{T}] = \operatorname{diag}(\Sigma_{\tilde{x}^1},\ldots,\Sigma_{\tilde{x}^n}), \quad \Sigma_{\tilde{x}}^i \triangleq \mathbb{E}[\tilde{x}_0^i(\tilde{x}_0^i)^\mathsf{T}]\\ &\Sigma_{\tilde{w}_t} \triangleq \mathbb{E}[\tilde{w}_t(\tilde{w}_t)^\mathsf{T}] = \operatorname{diag}(\Sigma_{\tilde{w}_t}^1,\ldots,\Sigma_{\tilde{w}_t}^n), \quad \Sigma_{\tilde{w}_t}^i \triangleq \mathbb{E}[\tilde{w}_t^i(\tilde{w}_t^i)^\mathsf{T}],\end{aligned}$$

and the matrices P_t, L_t, $\tilde{P}_t^i$ and $\tilde{L}_t^i$ are defined in (18) and (19). It is worth noting that P_t is the common estimate error covariance and $\tilde{P}_t^i$ is the quantization error covariance at each agent.

From the decomposition of J_{Control} in above equation and the fact that $\hat{w}_t$ and $\tilde{w}_t$ are control free, one may conclude that $\hat{u}_t = -L_t\hat{x}_t$ and $\tilde{u}_t^i = -\tilde{L}_t^i\tilde{x}_t^i$ are the optimal choices. Thus, combining the optimal choices for $\hat{u}_t$ and $\tilde{u}_t$, we obtain

$$u_t^* = -L_t\hat{x}_t - \operatorname{diag}(\tilde{L}_t^1,\ldots,\tilde{L}_t^n)\tilde{x}_t,$$

and therefore the optimal input of agent i is

$$u_t^i = -L_t^i\hat{x}_t - \tilde{L}_t^i\tilde{x}_t^i,$$

where L_t^i is the i-th block-row of the matrix L_t. This completes the derivation of the optimal controller. Therefore we have,

$$\begin{aligned}J_{\text{Control}} &= \operatorname{tr}(P_0\Sigma_{\hat{x}}) + \sum_{t=0}^{T-1}\operatorname{tr}(P_{t+1}\Sigma_{\hat{w}})\\ &+ \sum_{i=1}^n\left[\operatorname{tr}\left(\tilde{P}_0^i\Sigma_{\tilde{x}}^i\right) + \sum_{t=0}^{T-1}\operatorname{tr}\left(\tilde{P}_{t+1}^i\Sigma_{\tilde{w}}^i\right)\right].\end{aligned}$$

Let us define $\tilde{P}_0 = \operatorname{diag}(\tilde{P}_0^1,\ldots,\tilde{P}_0^n)$, $\tilde{P}_t = \operatorname{diag}(\tilde{P}_t^1,\ldots,\tilde{P}_t^n)$ in (a), and $\bar{P}_t \triangleq \tilde{P}_t - P_0$ and denote $\bar{P}_t^i$ to be the i-th diagonal block of $\bar{P}_t$. Consequently, we may write

$$\begin{aligned}J_{\text{Control}} &= \operatorname{tr}(P_0\Sigma_x) + \sum_{t=0}^{T-1}\operatorname{tr}(P_{t+1}\Sigma_w)\\ &+ \sum_{i=1}^n\Big(\operatorname{tr}\left(\bar{P}_0^i\Sigma_{\tilde{x}}^i\right) + \sum_{t=0}^{T-1}\operatorname{tr}\left(\bar{P}_{t+1}^i\Sigma_{\tilde{w}_t}^i\right)\Big).\end{aligned}$$

This expression helps to show explicitly how the choice of the quantizers affects J_{Control}. To that end, recall that $\Sigma_{\tilde{w}_t}^i = \mathbb{E}[\tilde{w}_t^i(\tilde{w}_t^i)^\mathsf{T}]$, $\tilde{w}_t^i = w_t^i - \mathbb{E}[w_t^i \mid z_{t+1}^i,\theta_{t+1}^i]$, and $z_{t+1}^i = \sum_{m=1}^M\theta_{t+1}^{ij}\delta^m(w_t^i)$. Therefore,

$$\begin{aligned}\tilde{w}_t^i &= w_t^i - \sum_{m=1}^M\theta_{t+1}^{im}\mathbb{E}[w_t^i \mid \delta^m(w_t^i)]\\ &= \sum_{m=1}^M\theta_{t+1}^{im}(w_t^i - \mathbb{E}[w_t^i \mid \delta^m(w_t^i)]) \triangleq \sum_{m=1}^M\theta_{t+1}^{im}\tilde{w}_t^{im},\end{aligned}$$

where we have used the constraint that $\sum_{m=1}^M\theta_t^{im} = 1$ and $\tilde{w}_t^{im}$ is the quantization error of quantizer m on the signal w_t^i. Consequently,

$$\begin{aligned}\Sigma_{\tilde{w}_t}^i &= \mathbb{E}\left[\sum_{m=1}^M\theta_{t+1}^{im}\tilde{w}_t^{im}(\tilde{w}_t^{im})^\mathsf{T}\right]\\ &= \mathbb{E}\left[\mathbb{E}\Big[\sum_{m=1}^M\theta_{t+1}^{im}\tilde{w}_t^{im}(\tilde{w}_t^{im})^\mathsf{T} \mid \theta_{t+1}^i\Big]\right] \triangleq \mathbb{E}\left[\sum_{m=1}^M\theta_{t+1}^{im}F^{im}\right],\end{aligned}$$

where

$$F^{im} = \sum_{j=1}^{\ell_m}\int_{p_s^m}(w - \mathbb{E}[w \in p_j^m])(w - \mathbb{E}[w \in p_j^m])^\mathsf{T}\mathbb{P}^i(dw) \tag{28}$$

is the quantization error covariance that depends on the partitions of $\mathcal{P}^m$ and the number of the quantization levels ℓ_m of the m-th quantizer and the distribution of the source signal $\mathbb{P}^i$. Notice that F^{im} does not depend on time since the distribution of w_t^i does not change with time due to the i.i.d assumption. Similarly, one may obtain $\Sigma_{\tilde{x}}^i = \mathbb{E}\left[\sum_{m=1}^M \theta_0^{im} F_0^{im}\right]$, where F_0^{im} has the same expression as F^{im}, except $\mathbb{P}^i$ is replaced with the distribution of x_0^i. Finally, we obtain

$$
\begin{aligned}
J_{\text{Control}} &= \operatorname{tr}\left(P_0 \Sigma_x\right) + \sum_{t=0}^{T-1} \operatorname{tr}\left(P_{t+1} \Sigma_w\right) \\
&+ \sum_{i=1}^{n} \mathbb{E}\Big[\sum_{m=1}^{M} \theta_0^{im} \operatorname{tr}\left(\bar{P}_0^i F_0^{im}\right) + \sum_{t=0}^{T-1} \sum_{m=1}^{M} \theta_{t+1}^{im} \operatorname{tr}\left(\bar{P}_{t+1}^i F^{im}\right)\Big].
\end{aligned}
$$

We define the constants

$$
\beta_t^{im} = \begin{cases} \operatorname{tr}\left(\bar{P}_0^i F_0^{im}\right), & t = 0, \\ \operatorname{tr}\left(\bar{P}_t^i F^{im}\right), & \text{otherwise}, \end{cases} \tag{29}
$$

and the vector $\beta_t^i = \operatorname{vec}(\beta_t^{i1}, \ldots, \beta_t^{iM})$, which yields

$$
J_{\text{Control}} = \operatorname{tr}\left(P_0 \Sigma_x\right) + \sum_{t=0}^{T-1} \operatorname{tr}\left(P_{t+1} \Sigma_w\right) + \sum_{i=1}^{n} \sum_{t=0}^{T-1} \mathbb{E}[(\beta_t^i)^\mathsf{T} \theta_t^i].
$$

This completes the proof.