International Journal of Control

ISSN: 0020-7179 (Print) 1366-5820 (Online) Journal homepage: http://www.tandfonline.com/loi/tcon20

Decentralised output-feedback LQG control with one-step communication delay

Yan Wang, Junlin Xiong & Wei Ren

To cite this article: Yan Wang, Junlin Xiong & Wei Ren (2018) Decentralised output-feedback LQG control with one-step communication delay, International Journal of Control, 91:8, 1920-1930, DOI: 10.1080/00207179.2017.1334965

To link to this article: https://doi.org/10.1080/00207179.2017.1334965

Accepted author version posted online: 24 May 2017.
Published online: 08 Jun 2017.

INTERNATIONAL JOURNAL OF CONTROL
2018, VOL. 91, NO. 8, 1920–1930
https://doi.org/10.1080/00207179.2017.1334965

Taylor & Francis
Taylor & Francis Group

Decentralised output-feedback LQG control with one-step communication delay

Yan Wang, Junlin Xiong and Wei Ren

Department of Automation, University of Science and Technology of China, Hefei, China

ABSTRACT
This paper studies the output-feedback LQG control of a two-player system with one-step communication delay. A novel information pattern is considered and a new controller structure is introduced. The proposed controller consists of two parts: the first part is based on estimated system state; the second part is based on current local measurement output. The form of the optimal controller is established using the method of independence decomposition. A necessary condition is established to construct the optimal controller gains. Two iterative algorithms are used to find the optimal gains numerically. Finally, the effectiveness of the theoretical results is illustrated through a numerical example.

ARTICLE HISTORY
Received 12 September 2016
Accepted 22 May 2017

KEYWORDS
Communication delay; decentralised control; local estimator; output-feedback

1. Introduction

Generally speaking, large-scale systems are typically composed of subsystems that are coupled via their system dynamics or common performance criterion. Numerous physical examples of such systems are found in engineering fields, including power grid (Aldeen, Saha, Alpcan, & Evans, 2015), teams of autonomous vehicles (Fax & Murray, 2004; Jin & Ray, 2014), satellite formation and aircraft (Esfahani & Khorasani, 2016). For large-scale systems where subsystems cannot obtain full system information, decentralised control is an effective control technique (Sandell, Varaiya, Athans, & Safonov, 1978). To achieve the best system performance, optimal decentralised control problems are usually modelled as optimisation problems with constraints. Unfortunately, those optimisation problems have been shown to be intractable in many cases (see Blondel & Tsitsiklis, 2000; Papadimitriou & Tsitsiklis, 1986). As a result, research attention has been paid to classes of the problems to which the corresponding optimisation problems are tractable. Several classes of these problems have been found in previous works such as Ho and Chu (1972), Qi, Salapaka, Voulgaris, and Khammash (2004), Rotkowitz and Lall (2006) and Lessard and Lall (2014). In Ho and Chu (1972), the partially nested information pattern was defined. Under this information pattern, the optimal decentralised control policies were proved to be linear. Linear control policies are easy to be found. In Qi et al. (2004), Rotkowitz and Lall (2006), and Lessard and Lall (2014), the definition of quadratic invariance was introduced. The optimal control problems with quadratic invariance constraint can be recast as convex optimisation problems. Convex optimisation problems can be solved by existing algorithms.

Decentralised linear quadratic Gaussian (LQG) control is an important and hot research topic. The output-feedback LQG control with one-step delay-sharing pattern has been studied by matrix minimum principle in Kurtaran and Sivan (1974), and by the second-guessing technique in Toda and Aoki (1975). For multiple steps delay-sharing pattern, the structural results to the optimal control have been established in Nayyar, Mahajan, and Teneketzis (2011). A common feature of the results in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011) is that the estimator of each subsystem needs to estimate the global system state. Hence, the system state is estimated repeatedly by the subsystems in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011). For a nested information pattern, the optimal decentralised output-feedback LQG controllers for large-scale systems with two interconnected subsystems have been designed in Lessard and Lall (2015). The result in Lessard and Lall (2015) is suitable for the two-player systems in which the system matrices are lower triangular, and is unlikely to be extended to the general two-player systems. For the delay model defined over a communication graph, the explicit optimal solution to decentralised state-feedback LQG control has been provided by information hierarchy graph and dynamic programming in Lamperski and Doyle (2012) and Lamperski and Lessard (2015). Furthermore, the explicit optimal solution to the varying delay case has been derived in Matni

CONTACT Junlin Xiong xiong77@ustc.edu.cn

and Doyle (2013). However, the results in Lamperski and Doyle (2012), Lamperski and Lessard (2015) and Matni and Doyle (2013) are not suitable for output feedback case, because the information independence decomposition is not valid for the output feedback case.

Motivated by the above discussion, we want to incorporate and expend the optimal controller design methods in Kurtaran and Sivan (1974), Toda and Aoki (1975), Nayyar et al. (2011), Lamperski and Doyle (2012), Lamperski and Lessard (2015) and Matni and Doyle (2013) to a novel information pattern case (that is, output feedback, local state estimation). In particular, we are interested in the output feedback LQG control with communication delay in which the subsystem only estimates its own subsystem state. Compared to the previous work in Kurtaran and Sivan (1974), Toda and Aoki (1975), Nayyar et al. (2011), Lessard and Lall (2015), Lamperski and Doyle (2012), Lamperski and Lessard (2015) and Matni and Doyle (2013), the advantages of our work are summarised as follows. First, the information pattern in this paper is novel. Compared to the information pattern considered in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011), where global state is estimated, our information pattern is suitable for local state estimation. Hence, the repeated calculation for global estimation in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011) is avoided in this paper. Second, the two-player system model in this paper is more general than the one in Lessard and Lall (2015), and the requirement on system matrices are removed. Third, the output feedback controller design method is studied in this paper, whereas Lamperski and Doyle (2012), Lamperski and Lessard (2015) and Matni and Doyle (2013) investigate the state feedback controller design. Because the measurement output instead of the system state is usually available in engineering practice, the study of output feedback controller design is more desirable than the study of the state feedback case.

This paper focuses on the decentralised output-feedback LQG control problem. A two-player system with one-step communication delay is considered. In the two-player system, the measurement output of each subsystem is transmitted to the other via the network with one-step delay. Each subsystem has a plant, a local estimator and a local controller. The local estimator estimates only the local system state and transmits the estimated local state to the other through the network. Based on the available information, a novel controller structure is proposed. The proposed controller is composed of two parts. By means of information independence decomposition and dynamic programming, the form of the optimal controller is established. A necessary condition for the design of the gains of the optimal controller is developed based on the discrete-time matrix minimum principle. The condition is given in terms of a set of matrix equations. The numerical solution to the gains of the optimal controller is found by two iterative algorithms. The first algorithm is developed based on the penalty function method. The second algorithm is exploited according to the gradient descent method. The value of the corresponding performance criterion achieved by the designed controller is derived. Finally, a numerical example is given to illustrate the effectiveness of the proposed theory.

Notation

$\mathbb{E}[x]$ denotes the expectation of a random variable x. A^{T} is the transpose of the matrix A. Let $\{x_{0:t}\}$ denote the sequence $\{x_0, x_1, \ldots, x_t\}$. $\mathrm{tr}[X]$ is the trace of the square matrix X. The notation $X \succ 0$ (respectively, $X \succeq 0$) means that X is a real symmetric positive definite matrix (respectively, real symmetric positive semi-definite matrix). For a matrix M that is partitioned into $n \times m$ blocks $[M_{ij}]_{1 \le i \le n, 1 \le j \le m}$, M^{rs} denotes the sub-matrix of M, given by $[M_{ij}]_{i \in r, j \in s}$, where $r \subseteq \{1, \ldots, n\}$, $s \subseteq \{1, \ldots, m\}$. For instance, if $r = \{1\}$, $s = \{1, 2\}$, then $M^{\{1\}\{1,2\}} = \begin{bmatrix} M_{11} & M_{12} \end{bmatrix}$.

2. Problem statement

Consider a linear time-varying stochastic system with two interconnected subsystems of the form

$$x^1_{t+1} = A^{11}_t x^1_t + A^{12}_t x^2_t + B^1_t u^1_t + \omega^1_t, \tag{1}$$

$$x^2_{t+1} = A^{21}_t x^1_t + A^{22}_t x^2_t + B^2_t u^2_t + \omega^2_t, \tag{2}$$

$$y^i_t = C^i_t x^i_t + \upsilon^i_t, \quad i \in \{1, 2\}, \tag{3}$$

where $x^i_t \in \mathbb{R}^{n_i}$ is the state of subsystem i at time t; $u^i_t \in \mathbb{R}^{l_i}$ is the control input; $y^i_t \in \mathbb{R}^{m_i}$ is the measurement output; $\omega^i_t \in \mathbb{R}^{n_i}$ and $\upsilon^i_t \in \mathbb{R}^{m_i}$ are the system noise and measurement noise, respectively. All the matrices in (1)–(3) have proper dimensions.

Define the following vectors

$$x \triangleq \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}, \quad u \triangleq \begin{bmatrix} u^1 \\ u^2 \end{bmatrix}, \quad \omega \triangleq \begin{bmatrix} \omega^1 \\ \omega^2 \end{bmatrix},$$
$$y \triangleq \begin{bmatrix} y^1 \\ y^2 \end{bmatrix}, \quad \upsilon \triangleq \begin{bmatrix} \upsilon^1 \\ \upsilon^2 \end{bmatrix},$$

then, the global system dynamics can be written as

$$x_{t+1} = A_t x_t + B_t u_t + \omega_t, \tag{4}$$

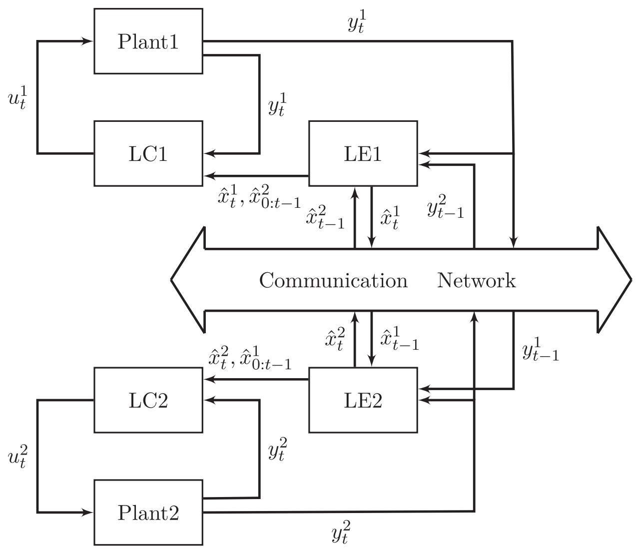

Figure 1. The large-scale system is composed of two subsystems. Each subsystem has a plant, a local estimator (LE) and a local controller (LC). The two subsystems exchange information via a one-step delay communication network.

$$y_t = C_t x_t + \upsilon_t, \tag{5}$$

where $A_t = \begin{bmatrix} A_t^{11} & A_t^{12} \\ A_t^{21} & A_t^{22} \end{bmatrix}$, $B_t = \begin{bmatrix} B_t^1 & 0 \\ 0 & B_t^2 \end{bmatrix}$, $C_t = \begin{bmatrix} C_t^1 & 0 \\ 0 & C_t^2 \end{bmatrix}$. The following assumptions are made about the initial state and noises. The initial state x_0 is a Gaussian random vector with $x_0 \sim \mathcal{N}(0, \Sigma_0)$, where $\Sigma_0 \succ 0$. The disturbances ω_t and υ_t are independent Gaussian noises with $\omega_t \sim \mathcal{N}(0, W_t)$, $\upsilon_t \sim \mathcal{N}(0, V_t)$, where $W_t \succ 0$, $V_t \succ 0$. Furthermore, x_0, ω_{t_1} and υ_{t_2} are pairwise uncorrelated for all t_1 and t_2.

Remark 2.1: The control problems for two-player systems are basic and important research problems. Some special decentralised control problems in discrete-time interconnected systems composed of two subsystems were addressed in Duan, Wang, and Huang (2007). The stabilisation problem of second-order switched positive systems consisting of two unstable subsystems has been studied in Zheng and Feng (2011). A class of two-player output-feedback problems with a nested information pattern has been studied in Lessard and Lall (2015).

In the large-scale system (4)–(5), the subsystems exchange information through a network with one-step delay. The system configuration is depicted in Figure 1, where each subsystem has a plant, a local estimator (LE) and a local controller (LC). Because y_t^j is transmitted to LEi via network with one-step delay, the available measurement output to LEi is $\{y_t^i, y_{0:t-1}\}$. In order to use the standard Kalman filter in LEi, $\{y_{0:\,t-1}\}$ is used to estimate the local state x_t^i. The estimated value of x_t^i is denoted by $\hat{x}_t^i$. The estimated local state $\hat{x}_t^i$ computed in LEi is transmitted to LEj via the same communication network. As a result, the estimated state available to LEi is given by $\mathcal{E}_t^i \triangleq \{\hat{x}_t^i, \hat{x}_{0:t-1}\}$, where $\hat{x}_t$ is obtained by the following standard Kalman filter (Åström, 2012):

$$\hat{x}_{t+1} = A_t \hat{x}_t + B_t u_t + K_t (y_t - C_t \hat{x}_t), \quad \hat{x}_0 = 0, \tag{6}$$

$$P_{t+1} = A_t P_t A_t^{\mathrm{T}} + W_t - A_t P_t C_t^{\mathrm{T}} (C_t P_t C_t^{\mathrm{T}} + V_t)^{-1} C_t P_t A_t^{\mathrm{T}}, \quad P_0 = \Sigma_0, \tag{7}$$

$$K_t = A_t P_t C_t^{\mathrm{T}} (C_t P_t C_t^{\mathrm{T}} + V_t)^{-1}, \tag{8}$$

where $P_t = \mathbb{E}[e_t e_t^{\mathrm{T}}]$ and $e_t \triangleq x_t - \hat{x}_t$ is the estimation error.

A lemma associated with the Kalman filter is stated in the following.

Lemma 2.1: (*Åström, 2012*) *The sequence*

$$\phi_t = y_t - C_t \hat{x}_t, \quad \textit{for } t = 0, \ldots, N-1,$$

is a zero-mean uncorrelated Gaussian process with covariance $Y_t = C_t P_t C_t^{\mathrm{T}} + V_t$. *Moreover,* ϕ_t *is uncorrelated with the past measurements, that is* $\mathbb{E}[\phi_{t_1} y_{t_2}^{\mathrm{T}}] = 0$, *for* $t_2 < t_1$.

Remark 2.2: If the global state x_t is estimated in each LE such as in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011), then the estimated local state $\hat{x}_t^i$ is not needed to be transmitted to LEj via the network. This means that the estimated state available to LEi is $\{\hat{x}_{0:t}\}$. However, in this case, the repeated calculation for estimation is conducted across over the two LEs. In other words, the amount of calculation for estimation is larger than that of the case considered in this paper.

The estimated state available to LEi and $\{y_t^i\}$ are delivered to LCi through the physical connection without delay as illustrated in Figure 1. Thus, the information for LCi has two parts:

$$\mathcal{D}_t^i = \{y_t^i\} \cup \mathcal{E}_t^i, \quad i = 1, 2. \tag{9}$$

Based on (9), the controller to be designed in this paper is of the form:

$$u_t^i = \gamma_t^i(\mathcal{E}_t^i) + F_t^i y_t^i, \quad i = 1, 2, \tag{10}$$

where F_t^i is the gain matrix and γ_t^i is a Borel-measurable function.

Remark 2.3: Although y_t^i is not used to estimate the system state x_t^i, it is used to generate the control input u_t^i directly. Thus, the second term in (10) plays the role of the correction.

The objective of this paper is to find the control sequences $\{u^i_{0:N-1}\}$, $i = 1, 2$, satisfying (10) such that the quadratic performance criterion

$$J \triangleq \mathbb{E}\left[\sum_{t=0}^{N-1}(x_t^{\mathsf{T}}Q_tx_t + u_t^{\mathsf{T}}R_tu_t) + x_N^{\mathsf{T}}Q_Nx_N\right] \tag{11}$$

is minimised, where $Q_t \succeq 0$ and $R_t \succ 0$ are given matrices. In other words, our control problem can be formulated as the following optimisation problem:

$$\begin{aligned} &\min \quad J \\ &\text{subject to} \quad (1), (2), (3), (10). \end{aligned} \tag{12}$$

3. Controller design

In this section, the main results of this paper are presented. In Section 3.1, the form of the optimal policies $\gamma_t^1(\mathcal{E}_t^1)$ and $\gamma_t^2(\mathcal{E}_t^2)$ are found. In Section 3.2, a necessary condition to (12) with respect to the gains of the controller is established. In Section 3.3, two algorithms are given to compute the optimal controller gains. In Section 3.4, the value of the optimal performance criterion is derived.

3.1 The form of optimal policies γ_t^1 and γ_t^2

The optimal policies γ_t^1 and γ_t^2 are shown to be linear in this subsection. Moreover, the form of optimal γ_t^1 and γ_t^2 are presented.

Lemma 3.1: *Consider the optimisation problem (12), the optimal policies γ_t^1 and γ_t^2 are of the form*

$$\begin{bmatrix}\gamma_t^1(\mathcal{E}_t^1)\\ \gamma_t^2(\mathcal{E}_t^2)\end{bmatrix} = \begin{bmatrix}(T_t^{\{1\}} - F_t^1C_t^1)\zeta_t^{\{1\}}\\ (T_t^{\{2\}} - F_t^2C_t^2)\zeta_t^{\{2\}}\end{bmatrix} + (T_t^{\{1,2\}} - F_tC_t)\zeta_t^{\{1,2\}}, \tag{13}$$

where

$$\begin{aligned} \mathcal{F} &= \{\{1\}, \{2\}, \{1, 2\}\},\\ \zeta_0^r &= 0, \quad \textit{for} \quad r \in \mathcal{F},\\ \zeta_{t+1}^{\{1\}} &= \hat{\omega}_t^1, \quad \zeta_{t+1}^{\{2\}} = \hat{\omega}_t^2,\\ \zeta_{t+1}^{\{1,2\}} &= \sum_{r\in\mathcal{F}}(A_t^{\{1,2\}r} + B_t^{\{1,2\}r}T_t^r)\zeta_t^r, \end{aligned} \tag{14}$$

the gain matrices $T_t^{\{1\}}$, $T_t^{\{2\}}$ and $F_t = diag\{F_t^1, F_t^2\}$ are to be chosen in the optimisation procedure, and $T_t^{\{1,2\}}$ are computed, recursively, as follows:

$$X_N = Q_N, \tag{15}$$

$$\Omega_t = R_t + B_t^{\mathsf{T}}X_{t+1}B_t, \tag{16}$$

$$T_t^{\{1,2\}} = -\Omega_t^{-1}B_t^{\mathsf{T}}X_{t+1}A_t, \tag{17}$$

$$X_t = Q_t + A_t^{\mathsf{T}}X_{t+1}A_t - T_t^{\{1,2\}^{\mathsf{T}}}\Omega_tT_t^{\{1,2\}}. \tag{18}$$

Proof: The control input u_t^i in (10) can be decomposed into

$$u_t^i = \tilde{u}_t^i + F_t^i\phi_t^i, \quad i = 1, 2. \tag{19}$$

As a result,

$$\begin{aligned} \tilde{u}_t^i &= F_t^iy_t^i + \gamma_t^i(\mathcal{E}_t^i) - F_t^i(y_t^i - C_t^i\hat{x}_t^i)\\ &= \gamma_t^i(\mathcal{E}_t^i) + F_t^iC_t^i\hat{x}_t^i, \quad i = 1, 2. \end{aligned} \tag{20}$$

Denote $\tilde{u}_t = \begin{bmatrix}\tilde{u}_t^1\\ \tilde{u}_t^2\end{bmatrix}$. Using $e_t = x_t - \hat{x}_t$ and substituting (19) into (11), we have

$$\begin{aligned} J =& \mathbb{E}\left\{\hat{x}_N^{\mathsf{T}}Q_N\hat{x}_N + \sum_{t=0}^{N-1}(\hat{x}_t^{\mathsf{T}}Q_t\hat{x}_t + \tilde{u}_t^{\mathsf{T}}R_t\tilde{u}_t)\right.\\ &\left. + \sum_{t=0}^{N}e_t^{\mathsf{T}}Q_te_t + \sum_{t=0}^{N-1}\phi_t^{\mathsf{T}}F_t^{*^{\mathsf{T}}}R_tF_t^*\phi_t + 2\sum_{t=0}^{N-1}\phi_t^{\mathsf{T}}F_t^{*^{\mathsf{T}}}R_t\tilde{u}_t\right\}. \end{aligned} \tag{21}$$

A routine computation gives that

$$\mathbb{E}[e_t^{\mathsf{T}}Q_te_t] = \mathrm{tr}\left(Q\mathbb{E}[e_te_t^{\mathsf{T}}]\right) = \mathrm{tr}(Q_tP_t), \tag{22}$$

$$\begin{aligned} \mathbb{E}[\phi_t^{\mathsf{T}}F_t^{*^{\mathsf{T}}}R_tF_t^*\phi_t] &= \mathrm{tr}\left(F_t^{*^{\mathsf{T}}}R_tF_t^*\mathbb{E}[\phi_t\phi_t^{\mathsf{T}}]\right)\\ &= \mathrm{tr}[F_t^{*^{\mathsf{T}}}R_tF_t^*(C_tP_tC_t^{\mathsf{T}} + V_t)]. \end{aligned} \tag{23}$$

Using the fact that $\mathcal{E}_t^i$ is a linear combination of $\{y_{0:t-1}\}$ and according to Lemma 2.1, it holds that

$$\mathbb{E}[\tilde{u}_t\phi_t^{\mathsf{T}}] = \mathbb{E}\left[\begin{bmatrix}\gamma_t^1(\mathcal{E}_t^1) + F_t^{1*}C_t^1\hat{x}_t^1\\ \gamma_t^2(\mathcal{E}_t^2) + F_t^{2*}C_t^2\hat{x}_t^2\end{bmatrix}\phi_t^{\mathsf{T}}\right] = 0,$$

which implies that

$$\mathbb{E}\left[\phi_t'F_t^{*^{\mathsf{T}}}R_t\tilde{u}_t\right] = \mathrm{tr}\{F_t^{*^{\mathsf{T}}}R_t\mathbb{E}[\tilde{u}_t\phi_t^{\mathsf{T}}]\} = 0. \tag{24}$$

Because (22), (23) and (24) do not depend on $\tilde{u}_t$, it turns out that the optimal input $\tilde{u}_t$ minimising the performance criterion (11) is identical to the one minimising

$$\tilde{J} = \mathbb{E}\left\{\hat{x}_N^{\mathsf{T}}Q_N\hat{x}_N + \sum_{t=0}^{N-1}(\hat{x}_t^{\mathsf{T}}Q_t\hat{x}_t + \tilde{u}_t^{\mathsf{T}}R_t\tilde{u}_t)\right\}. \tag{25}$$

Plugging (19) back into (6) and using $\phi_t = y_t - C_t\hat{x}_t$, we obtain that

$$\hat{x}_{t+1} = A_t\hat{x}_t + B_t\tilde{u}_t + \hat{\omega}_t, \quad \hat{x}_0 = 0, \tag{26}$$

where $\hat{\omega}_t = (B_tF_t + K_t)\phi_t$. Consider the system dynamics (26), the performance criterion (25) and the control input (20). The optimal $\tilde{u}_t$ is the solution to the following optimisation problem:

$$\begin{aligned} \min \quad & \tilde{J} \\ \text{subject to} \quad & \hat{x}_{t+1} = A_t\hat{x}_t + B_t\tilde{u}_t + \hat{\omega}_t, \\ & \tilde{u}_t^i = \gamma_t^i(\mathcal{E}_t^i) + F_t^{i*}C_t^i\hat{x}_t^i. \end{aligned} \tag{27}$$

The information structure of problem (27) is partially nested in terms of Definition 3 in Ho and Chu (1972). Using Theorem 2 in Ho and Chu (1972), the optimal $\tilde{u}_t^i$ to the problem (27) is linear. According to (20), the optimal $\gamma_t^i(\mathcal{E}_t^i)$ to the problem (12) is linear.

Considering problem (27) and using the state and input decomposition results in Lamperski and Lessard (2015), we have

$$\hat{x}_t = \begin{bmatrix} \zeta_t^{\{1\}} \\ \zeta_t^{\{2\}} \end{bmatrix} + \zeta_t^{\{1,2\}}, \quad \tilde{u}_t = \begin{bmatrix} \tilde{\varphi}_t^{\{1\}} \\ \tilde{\varphi}_t^{\{2\}} \end{bmatrix} + \tilde{\varphi}_t^{\{1,2\}}, \tag{28}$$

where $\zeta_t^{\{1\}}$ and $\tilde{\varphi}_t^{\{1\}}$ are the functions of $\hat{\omega}_{t-1}^1$; $\zeta_t^{\{2\}}$ and $\tilde{\varphi}_t^{\{2\}}$ are the functions of $\hat{\omega}_{t-1}^2$; $\zeta_t^{\{\{1,2\}\}}$ and $\tilde{\varphi}_t^{\{1,2\}}$ are the functions of $\{\hat{\omega}_{0:t-2}\}$.

It follows from Lemma 2.1 that $\{\hat{\omega}_{0:t-1}\}$ is independent of $\hat{\omega}_t^{\{1\}}$ and $\hat{\omega}_t^{\{2\}}$. By the technology of dynamic programming in Lamperski and Doyle (2012), the optimal solution $\tilde{\varphi}_t^{\{1,2\}}$ to problem (27) is given by

$$\tilde{\varphi}_t^{\{1,2\}} = T_t^{\{1,2\}}\zeta_t^{\{1,2\}},$$

where $T_t^{\{1,2\}}$ is given by (15)–(18). As a result, the optimal $\tilde{u}_t$ has the following form:

$$\tilde{u}_t = \begin{bmatrix} T_t^{\{1\}}\zeta_t^{\{1\}} \\ T_t^{\{2\}}\zeta_t^{\{2\}} \end{bmatrix} + T_t^{\{1,2\}}\zeta_t^{\{1,2\}}.$$

Using (20), it follows that

$$\begin{aligned} \begin{bmatrix} \gamma_t^1(\mathcal{E}_t^1) \\ \gamma_t^2(\mathcal{E}_t^2) \end{bmatrix} &= \tilde{u}_t - F_tC_t\hat{x}_t \\ &= \begin{bmatrix} T_t^{\{1\}}\zeta_t^{\{1\}} \\ T_t^{\{2\}}\zeta_t^{\{2\}} \end{bmatrix} + T_t^r\zeta_t^{\{1,2\}} - F_tC_t\left(\begin{bmatrix} \zeta_t^{\{1\}} \\ \zeta_t^{\{2\}} \end{bmatrix} + \zeta_t^{\{1,2\}}\right) \\ &= \begin{bmatrix} (T_t^{\{1\}} - F_t^1C_t^1)\zeta_t^{\{1\}} \\ (T_t^{\{2\}} - F_t^2C_t^2)\zeta_t^{\{2\}} \end{bmatrix} + (T_t^{\{1,2\}} - F_tC_t)\zeta_t^{\{1,2\}}. \end{aligned}$$

The proof is completed. ∎

In the following, for easy of notation, we denote $T_t^{\{1,2\}}$ and $\zeta_t^{\{1,2\}}$ by T_t and ζ_t, respectively. In addition, define

$$H_t = \text{diag}\{T_t^{\{1\}}, T_t^{\{2\}}\}.$$

3.2 Optimisation condition of gain matrices

According to Lemma 3.1, a necessary optimisation condition to problem (12) with respect to matrix variables F_t, and H_t is presented in this subsection by means of matrix minimum principle.

Theorem 3.1: *Consider the optimisation problem (12), suppose $F^*_{0:N-1}$ and $H^*_{0:N-1}$ are the optimal solutions. Then, $F^*_{0:N-1}$ and $H^*_{0:N-1}$ satisfy the following conditions:*

- *H_0^* is any block diagonal matrix with proper dimension, and satisfies $\|H_0^*\| < +\infty$.*
- *For $t = 0, \ldots, N-2$,*

$$(B_t{}^{\mathsf{T}}\Phi_{t+1}B_t + R_t)F_tY_t + B_t{}^{\mathsf{T}}\Phi_{t+1}K_tY_t + S_t = 0, \tag{29}$$

$$\begin{aligned} &(B_{t+1}^{\mathsf{T}}M_{t+2}B_{t+1} + R_{t+1})H_{t+1}\hat{W}_t \\ &\quad + B_{t+1}^{\mathsf{T}}M_{t+2}A_{t+1}\hat{W}_t + \tilde{S}_t = 0. \end{aligned} \tag{30}$$

- *For $t = N-1$,*

$$\begin{aligned} &(B_{N-1}{}^{\mathsf{T}}Q_NB_{N-1} + R_{N-1})F_{N-1}Y_{N-1} \\ &\quad + B_{N-1}{}^{\mathsf{T}}Q_NK_{N-1}Y_{N-1} + S_{N-1} = 0. \end{aligned} \tag{31}$$

In (29)–(31),

$$\begin{aligned} \Phi_t &= Q_t + H_t^{\mathsf{T}}R_tH_t + (A_t + B_tH_t)^{\mathsf{T}}M_{t+1}(A_t + B_tH_t), \\ \hat{W}_t &= (B_tF_t + K_t)Y_t(B_tF_t + K_t)^{\mathsf{T}}, \\ S_t &\in \begin{bmatrix} 0 & S_t^{12} \\ S_t^{21} & 0 \end{bmatrix}, \; S_t^{12} \in \mathbb{R}^{l_1\times m_2}, \; S_t^{21} \in \mathbb{R}^{l_2\times m_1}, \\ \tilde{S}_t &\in \begin{bmatrix} 0 & \tilde{S}_t^{12} \\ \tilde{S}_t^{21} & 0 \end{bmatrix}, \; \tilde{S}_t^{12} \in \mathbb{R}^{l_1\times n_2}, \; \tilde{S}_t^{21} \in \mathbb{R}^{l_2\times n_1}, \end{aligned}$$

and $M_t^* \in \mathbb{R}^{n\times n}$ is obtained by the following recursion equations:

$$M_t^* = Q_t^{\mathsf{T}} + T_t{}^{\mathsf{T}}R_tT_t + (A_t + B_tT_t)^{\mathsf{T}}M_{t+1}^*(A_t + B_tT_t), \quad t = N-1, \ldots, 1, \tag{32}$$

$$M_N^* = Q_N^{\mathsf{T}}. \tag{33}$$

Proof: Equations (28) can be rewritten in the following form:

$$\hat{x}_t = \hat{\omega}_{t-1} + \zeta_t, \tag{34}$$

$$\tilde{u}_t = H_t\hat{\omega}_{t-1} + T_t\zeta_t, \tag{35}$$

where $H_t = \text{diag}\{T_t^{\{1\}}, T_t^{\{2\}}\}$. Recall that, $\hat{\omega}_{-1} = \begin{bmatrix} \zeta_0^{\{1\}} \\ \zeta_0^{\{2\}} \end{bmatrix} = 0$ (see Lemma 3.1). It follows from (35) and $\hat{\omega}_{-1} = 0$ that the optimal H_0^* can be any block diagonal matrix with proper dimension satisfying $\|H_0^*\| < +\infty$.

Plugging (34)–(35) into (21), we have

$$J = \bar{J} + \mathbb{E}\Big[\sum_{t=0}^{N} e_t^{\mathsf{T}} Q_t e_t\Big],$$

where

$$\begin{aligned}
\bar{J} = \mathbb{E}\Bigg\{ & \sum_{t=0}^{N-1}\Big[\Big(\hat{\omega}_{t-1} + \zeta_t\Big)^{\mathsf{T}} Q_t\Big(\hat{\omega}_{t-1} + \zeta_t\Big) \\
& + \Big(H_t\hat{\omega}_{t-1} + T_t\zeta_t\Big)^{\mathsf{T}} R_t\Big(H_t\hat{\omega}_{t-1} + T_t\zeta_t\Big) \\
& + \Big(F_t\phi_t\Big)^{\mathsf{T}} R_t F_t\phi_t\Big] + \Big(\hat{\omega}_{N-1} + \zeta_N\Big)^{\mathsf{T}} Q_N\Big(\hat{\omega}_{N-1} + \zeta_N\Big)\Bigg\} \\
= \mathbb{E}\Bigg\{ & \sum_{t=1}^{N-1}\Big[\zeta_t^{\mathsf{T}}\Big(Q_t + T_t^{\mathsf{T}} R_t T_t\Big)\zeta_t + \hat{\omega}_{t-1}^{\mathsf{T}}\Big(Q_t + H_t^{\mathsf{T}} R_t H_t\Big)\hat{\omega}_{t-1}\Big] \\
& + \zeta_N^{\mathsf{T}} Q_N\zeta_N + \hat{\omega}_{N-1}^{\mathsf{T}} Q_N\hat{\omega}_{N-1} + \sum_{t=0}^{N-1}\Big(F_t\phi_t\Big)^{\mathsf{T}} R_t F_t\phi_t\Bigg\}.
\end{aligned}$$

The first equality follows from direct computation; the second equality holds due to the independence of ζ_t and $\hat{\omega}_{t-1}$.

Define $\Sigma_t \triangleq \mathbb{E}[\zeta_t\zeta_t^{\mathsf{T}}]$ for $t = 0, \ldots, N$. It turns out that $\bar{J} = \sum_{t=0}^{N-1} L_{t+1}$, where L_{t+1} is given by

$$\begin{aligned}
L_{t+1} = {} & \text{tr}\Big[(Q_{t+1} + T_{t+1}^{\mathsf{T}} R_{t+1} T_{t+1})\Sigma_{t+1}\Big] \\
& + \text{tr}\Big[(Q_{t+1} + H_{t+1}^{\mathsf{T}} R_{t+1} H_{t+1})\hat{W}_t\Big] \\
& + \text{tr}(F_t^{\mathsf{T}} R_t F_t Y_t).
\end{aligned} \tag{36}$$

In (36), $T_N = 0$ and $H_N = 0$. In addition, according to (14), one has that Σ_t can be computed by the following recursion equations:

$$\Sigma_0 = \Sigma_1 = 0, \tag{37}$$

$$\begin{aligned}
\Sigma_{t+1} = {} & (A_t + B_t T_t)\Sigma_t(A_t + B_t T_t)^{\mathsf{T}} \\
& + (A_t + B_t H_t)\hat{W}_{t-1}(A_t + B_t H_t)^{\mathsf{T}}, \\
& \text{for } t = 1, \ldots, N-1.
\end{aligned} \tag{38}$$

Thus, the optimal $F_{0:N-1}^*$ and $H_{0:N-1}^*$ of the minimising problem (12) are the optimal solution to the following optimisation problem:

$$\begin{aligned}
\min \quad & \bar{J} = \sum_{t=0}^{N-1} L_{t+1} \\
\text{subject to} \quad & (37),\ (38).
\end{aligned} \tag{39}$$

Now (37) and (38) are viewed as the state dynamic equations, where the state is Σ_t, the gain matrices F_{t-1} and H_t play the role of inputs, and the performance criterion is $\bar{J}$. The aim is to find optimal $F_{0:N-1}^*$ and $H_{0:N-1}^*$ to minimise $\bar{J}$. This optimisation problem can be dealt with by the discrete matrix minimum principle (Athans, 1967). The Hamiltonian function for the optimisation problem is

$$\begin{aligned}
h_t = L_t & + \text{tr}\Big[(A_t + B_t T_t)\Sigma_t(A_t + B_t T_t)^{\mathsf{T}} M_{t+1}^{\mathsf{T}}\Big] \\
& + \text{tr}\Big\{(A_t + B_t H_t)\hat{W}_{t-1}(A_t + B_t H_t)^{\mathsf{T}} M_{t+1}^{\mathsf{T}}\Big\} \\
& + \text{tr}\Big[2F_{t-1}S_{t-1}^{\mathsf{T}}\Big] + \text{tr}\Big[2H_t\tilde{S}_t^{\mathsf{T}}\Big], \\
& \text{for } t = 1, \ldots, N,
\end{aligned}$$

where S_t and $\tilde{S}_t$ are the Lagrange multipliers matrices and $M_t \in R^{n\times n}$ is the costate matrix. A necessary condition to problem (12) is given by

$$\frac{\partial h_t}{\partial F_{t-1}}\Big|_* = 0, \quad \frac{\partial h_t}{\partial H_t}\Big|_* = 0, \tag{40}$$

$$\frac{\partial h_t}{\partial M_{t+1}}\Big|_* = \Sigma_t^*, \quad \Sigma_1 = \Sigma_1^*, \tag{41}$$

$$\frac{\partial h_t}{\partial \Sigma_t}\Big|_* = M_t^*, \quad 0 = M_{N+1}^*. \tag{42}$$

Through direct computation, the partial differential of $\hat{W}_t$ with respect to F_t is given by

$$\begin{aligned}
\partial(\hat{W}_t) = {} & B_t\partial(F_t)(Y_t F_t^{\mathsf{T}} B_t^{\mathsf{T}} + Y_t K_t^{\mathsf{T}}) \\
& + (B_t F_t Y_t + K_t Y_t)(\partial(F_t))^{\mathsf{T}} B_t^{\mathsf{T}},
\end{aligned} \tag{43}$$

and with respect to H_t is

$$\partial(\hat{W}_t) = 0. \tag{44}$$

Using (43)–(44), (40) gives (29)–(31). In addition, (42) gives (32)–(33). The proof is completed. ∎

Remark 3.1: In this paper, the optimal F_{t-1} depends on $T_t^{\{1\}}$ and $T_t^{\{2\}}$, where we have denoted $\mathrm{diag}\{T_t^{\{1\}}, T_t^{\{2\}}\} = H_t$. Thus, F_t cannot be designed separably by solving matrix equation as in Kurtaran and Sivan (1974), Toda and Aoki (1975) and Nayyar et al. (2011). On the other hand, because $\zeta_t^{\{1\}}$ and $\zeta_t^{\{2\}}$ are correlated in this paper, $T_t^{\{1\}}$ and $T_t^{\{2\}}$ cannot be computed directly by recursive equations as in Lamperski and Doyle (2012), Lamperski and Lessard (2015) and Matni and Doyle (2013). Constructing the optimal $T_t^{\{1\}}$, $T_t^{\{2\}}$ and F_{t-1} is a challenging task in this paper. In the following, we construct the optimal $T_t^{\{1\}}$, $T_t^{\{2\}}$ and F_{t-1} jointly via iterative algorithms, as the following subsection presents.

3.3 Iterative algorithms

In order to obtain the optimal gain matrices $F_{0:N-1}$ and $H_{0:N-1}$ numerically, iterative algorithms are exploited in this subsection. According to the proof of Theorem 3.1 and the matrix minimum principle (Athans, 1967), the optimal F_t and H_t are obtained by solving the following optimisation problem:

$$\begin{aligned} &\min \quad \hat{h}_t \\ &\text{subject to} \quad F_t \in \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}, H_t \in \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}, \end{aligned} \tag{45}$$

where

$$\begin{aligned} \hat{h}_t = L_t &+ \mathrm{tr}\left[(A_t + B_t T_t)\Sigma_t (A_t + B_t T_t)^{\mathsf{T}} M_{t+1}^{\mathsf{T}}\right] \\ &+ \mathrm{tr}\left\{(A_t + B_t H_t)\hat{W}_{t-1}(A_t + B_t H_t)^{\mathsf{T}} M_{t+1}^{\mathsf{T}}\right\}. \end{aligned}$$

Here, we solve the optimisation problem (45) by the penalty function method. In the following, we first derive the augmented objective function for the penalty function method.

Gain matrices are given as follows:

$$F = \begin{bmatrix} F^{11} & F^{12} \\ F^{21} & F^{22} \end{bmatrix}, \ H = \begin{bmatrix} H^{11} & H^{12} \\ H^{21} & H^{22} \end{bmatrix},$$

where $F^{11} \in \mathbb{R}^{l_1 \times m_1}$, $F^{12} \in \mathbb{R}^{l_1 \times m_2}$, $F^{21} \in \mathbb{R}^{l_2 \times m_1}$, and $F^{22} \in \mathbb{R}^{l_2 \times m_2}$; $H^{11} \in \mathbb{R}^{l_1 \times n_1}$, $H^{12} \in \mathbb{R}^{l_1 \times n_2}$, $H^{21} \in \mathbb{R}^{l_2 \times n_1}$, and $H^{22} \in \mathbb{R}^{l_2 \times n_2}$. We have that

$$\begin{aligned} \Lambda^1 F \Lambda^2 &= F^{21}, \ \Lambda^3 F \Lambda^4 = F^{12}, \\ \Upsilon^1 H \Upsilon^2 &= H^{21}, \ \Upsilon^3 H \Upsilon^4 = H^{12}, \end{aligned}$$

where

$$\begin{aligned} \Lambda^1 &= \begin{bmatrix} 0_{l_2 \times l_1} & I_{l_2} \end{bmatrix}, \quad \Lambda^2 = \begin{bmatrix} I_{m_1} \\ 0_{m_2 \times m_1} \end{bmatrix}, \\ \Lambda^3 &= \begin{bmatrix} I_{l_1} & 0_{l_1 \times l_2} \end{bmatrix}, \quad \Lambda^4 = \begin{bmatrix} 0_{m_1 \times m_2} \\ I_{m_2} \end{bmatrix}, \\ \Upsilon^1 &= \begin{bmatrix} 0_{l_2 \times l_1} & I_{l_2} \end{bmatrix}, \quad \Upsilon^2 = \begin{bmatrix} I_{n_1} \\ 0_{n_2 \times n_1} \end{bmatrix}, \\ \Upsilon^3 &= \begin{bmatrix} I_{l_1} & 0_{l_1 \times l_2} \end{bmatrix}, \quad \Upsilon^4 = \begin{bmatrix} 0_{n_1 \times n_2} \\ I_{n_2} \end{bmatrix}. \end{aligned}$$

As a result, the following constraint conditions are equivalent:

(1)

$$F_t \in \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}, \ H_t \in \begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}.$$

(2)

$$\begin{aligned} \Lambda^1 F \Lambda^2 = 0, &\quad \Lambda^3 F \Lambda^4 = 0, \\ \Upsilon^1 H \Upsilon^2 = 0, &\quad \Upsilon^3 H \Upsilon^4 = 0. \end{aligned}$$

(3)

$$\begin{aligned} \mathrm{tr}(\Lambda^1 F \Lambda^2 (\Lambda^1 F \Lambda^2)^{\mathsf{T}}) &= 0, \\ \mathrm{tr}(\Lambda^3 F \Lambda^4 (\Lambda^3 F \Lambda^4)^{\mathsf{T}}) &= 0, \\ \mathrm{tr}(\Upsilon^1 H \Upsilon^2 (\Upsilon^1 H \Upsilon^2)^{\mathsf{T}}) &= 0, \\ \mathrm{tr}(\Upsilon^3 H \Upsilon^4 (\Upsilon^3 H \Upsilon^4)^{\mathsf{T}}) &= 0. \end{aligned}$$

Hence, the augmented objective function for the penalty function method can be given by

$$\begin{aligned} &\Psi_t(c, F_{t-1}, H_t) \\ &\quad = \hat{h}_t + c\Big\{ \mathrm{tr}\left[\Lambda^1 F_{t-1} \Lambda^2 (\Lambda^1 F_{t-1} \Lambda^2)^{\mathsf{T}}\right] \\ &\qquad + \mathrm{tr}[\Lambda^3 F_{t-1} \Lambda^4 (\Lambda^3 F_{t-1} \Lambda^4)^{\mathsf{T}}] \\ &\qquad + \mathrm{tr}\left[\Upsilon^1 H_t \Upsilon^2 (\Upsilon^1 H_t \Upsilon^2)^{\mathsf{T}}\right] \\ &\qquad + \mathrm{tr}[\Upsilon^3 H_t \Upsilon^4 (\Upsilon^3 H_t \Upsilon^4)^{\mathsf{T}}]\Big\}, \quad t = 0, \ldots, N, \end{aligned} \tag{46}$$

where c is the penalty parameter; $H_N = 0$ and $T_N = 0$.

Based on the augmented objective function (46), Algorithm 1 is presented to solve the optimisation problem (45) off-line. Based on Theorem 17.1 in Nocedal and Wright (2006), the convergence of Algorithm 1 is guaranteed.

Algorithm 1: Compute optimal $F_{0:N-1}$ and $H_{0:N-1}$ offline

---

1: Obtain $H_0^* = 0$, and obtain F_{N-1}^* by solving (31).
2: **for** $1 \le t \le N-1$ **do**
3: **Initialisation**: Select an initial penalty parameter c_0, a stopping parameter $\delta > 0$, and a growth parameter $\eta > 1$. Choose the starting point $F_{t-1}[0] = 0$, $H_t[0] = 0$ and formulate the initial augmented objective function $\Psi_t(c[0], F_{t-1}, H_t)$. Set $k = 1$.
4: **Iterative**: Starting from $F_{t-1}[k-1], H_t[k-1]$, use gradient descent algorithm (Algorithm 2) to find the solution of $\min\{\Psi_t(c[k-1], F_{t-1}, H_t)\}$, and the solution is denoted by $F_{t-1}[k], H_t[k]$.
5: **Stopping**: If the following condition is satisfied,

$$
\begin{aligned}
&\mathrm{tr}\left[\Lambda^1 F_{t-1}\Lambda^2(\Lambda^1 F_{t-1}\Lambda^2)^{\mathrm{T}}\right] \\
&\quad + \mathrm{tr}[\Lambda^3 F_{t-1}\Lambda^4(\Lambda^3 F_{t-1}\Lambda^4)^{\mathrm{T}}] \\
&\quad + \mathrm{tr}\left[\tilde{\Lambda}^1 H_t\tilde{\Lambda}^2(\tilde{\Lambda}^1 H_t\tilde{\Lambda}^2)^{\mathrm{T}}\right] \\
&\quad + \mathrm{tr}[\tilde{\Lambda}^3 H_t\tilde{\Lambda}^4(\tilde{\Lambda}^3 H_t\tilde{\Lambda}^4)^{\mathrm{T}}] < \delta,
\end{aligned}
$$

then stop with $F_{t-1}^* = F_{t-1}[k]$, $H_t^* = H_t[k]$.Otherwise, let $c[k] = \eta c[k-1]$, formulate the newaugmented objective function $\Psi_t(c[k], F_{t-1}, H_t)$, let $k = k+1$ and return to Iterative Step.
6: **end for**

In iterative step of Algorithm 1, the gradient descent algorithm is given by Algorithm 2. Algorithm 2 is developed according to the steepest descent algorithm in Snyman (2005).

Algorithm 2: Gradient descent algorithm in Algorithm 1

1: **Initialisation**: Set $F_{t-1}[k][0] = 0$, $H_t[k][0] = 0$, $\Psi_t[k][-1] = \infty$. Select an iteration steplength $\gamma > 0$. Let $i = 0$.
2: **while** $\Psi_t[k][i] < \Psi_t[k][i-1]$ **do**
3: $i = i + 1$,
$F_{t-1}[k][i] = F_{t-1}[k][i-1] - \gamma \dfrac{\nabla_{F_{t-1}}\Psi_t[k]\left(F_{t-1}[k][i-1], H_t[k][i-1]\right)}{||\nabla_{F_{t-1}}\Psi_t[k]\left(F_{t-1}[k][i-1], H_t[k][i-1]\right)||}$,
$H_t[k][i] = H_t[k][i-1] - \gamma \dfrac{\nabla_{H_t}\Psi_t[k]\left(F_{t-1}[k][i-1], H_t[k][i-1]\right)}{||\nabla_{H_t}\Psi_t[k]\left(F_{t-1}[k][i-1], H_t[k][i-1]\right)||}$,
4: **end while**
5: $F_{t-1}[k] = F_{t-1}[k][i]$, $H_t[k] = H_t[k][i]$.

Remark 3.2: In Algorithm 2, the partial gradients of $\Psi_t(c[k], F_{t-1}, H_t)$ with respect to F_{t-1} and H_t are given, respectively, by

$$
\begin{aligned}
&\frac{\partial \Psi_t(c[k], F_{t-1}, H_t)}{\partial F_{t-1}} \\
&\quad = (B^{-\mathrm{T}}\Phi B^- + R^-)F^-Y^- + B^{-\mathrm{T}}\Phi K^- Y^- \\
&\qquad + c[k](\Lambda^{1^{\mathrm{T}}}\Lambda^1 F^-\Lambda^2\Lambda^{2^{\mathrm{T}}} + \Lambda^{3^{\mathrm{T}}}\Lambda^3 F^-\Lambda^4\Lambda^{4^{\mathrm{T}}}), \\
&\frac{\partial \Psi_t(c[k], F_t, H_t)}{\partial H_t} \\
&\quad = (B^{\mathrm{T}} M^+ B + R)H\hat{W}^- + B^{\mathrm{T}} M^+ A\hat{W}^- \\
&\qquad + c[k](\Upsilon^{1^{\mathrm{T}}}\Upsilon^1 H_t\Upsilon^2\Upsilon^{2^{\mathrm{T}}} + \Upsilon^{3^{\mathrm{T}}}\Upsilon^3 H_t\Upsilon^4\Upsilon^{4^{\mathrm{T}}}),
\end{aligned}
$$

where the time index is omitted here, and the superscript '+', ('−') means that the time index is '$t + 1$', ('$t - 1$'); Φ_t is defined in Theorem 3.1. For simplicity, in Algorithm 2, $\frac{\partial \Psi_t(c[k], F_{t-1}, H_t)}{\partial F_{t-1}}$ is denoted by $\nabla_{F_{t-1}}\Psi_t[k]$, and $\frac{\partial \Psi_t(c[k], F_t, H_t)}{\partial H_t}$ is denoted by $\nabla_{H_t}\Psi_t[k]$.

3.4 The controller and corresponding performance criterion

Based on the results developed in the previous subsections, the designed controller and the corresponding performance criterion are presented in this subsection.

Theorem 3.2: *The optimal solution to problem* (12) *is given by*

$$
u_t^{i*} = \gamma_t^{i*}(\mathcal{E}_t^i) + F_t^{i*} y_t^i, \quad \text{for} \quad i = 1, 2, \tag{47}
$$

where $\gamma_t^{i*}(\mathcal{E}_t^i)$ *is given by* (13) *in Lemma* 3.1. *In* $\gamma_t^{i*}(\mathcal{E}_t^i)$, $T_t^{\{1\}}$ *and* $T_t^{\{2\}}$ *are constructed from* H_t^*. H_t^* *and* F_t^* *are computed by Algorithm* 1 *or Theorem* 3.1. *Moreover, the value of the performance criterion* (11) *achieved by the designed controller is*

$$
\begin{aligned}
J^* = &\sum_{t=0}^{N-1} \mathrm{tr}\left[F_t^{*\mathrm{T}} R_t F_t^* (C_t P_t C_t^{\mathrm{T}} + V_t)\right] \\
&+ \sum_{t=0}^{N} \left\{ \mathrm{tr}\left[(Q_t + T_t^{\mathrm{T}} R_t T_t)\Sigma_t^*\right] \right. \\
&\left. + \ \mathrm{tr}\left[(Q_t + H_t^{*\mathrm{T}} R_t H_t^*)\hat{W}_{t-1}^*\right] + \mathrm{tr}\left[Q_t P_t\right] \right\},
\end{aligned} \tag{48}
$$

where $\Sigma_0^* = 0$, $\hat{W}_{-1}^* = 0$.

Proof: According to Lemma 3.1 and Theorem 3.1, it follows that (47) is the optimal solution to problem (12). We proceed to prove (48). Considering (25), one has that

$$\begin{aligned}\tilde{J}^* &= \sum_{t=0}^{N} \left\{ \hat{x}_t^{*\mathsf{T}}(Q_t + T_t^{\mathsf{T}} R_t T_t)\hat{x}_t^* \right. \\ &\quad \left. + \hat{\omega}_{t-1}^{*\mathsf{T}}(Q_t + H_t^{*\mathsf{T}} R_t H_t^*)\hat{\omega}_{t-1}^* \right\}, \\ &= \sum_{t=0}^{N} \left\{ \operatorname{tr}\left[(Q_t + T_t^{\mathsf{T}} R_t T_t)\Sigma_t^*\right] \right. \\ &\quad \left. + \operatorname{tr}\left[(Q_t + H_t^{*\mathsf{T}} R_t H_t^*)\hat{W}_{t-1}^*\right]\right\}.\end{aligned}$$

The optimal values of the other terms in (21) follow from direct computations. That is,

$$\left[\sum_{t=0}^{N} e_t^{\mathsf{T}} Q_t e_t\right]^* = \sum_{t=0}^{N} \operatorname{tr}\left[Q_t P_t\right],$$

$$\left[\sum_{t=0}^{N-1} \phi_t^{\mathsf{T}} F_t^{*\mathsf{T}} R_t F_t^* \phi_t\right]^* = \sum_{t=0}^{N-1} \operatorname{tr}[F_t^{*\mathsf{T}} R_t F_t^*(C_t P_t C_t^{\mathsf{T}} + V_t)].$$

According to (21), one has that (48) holds. The proof is completed. ■

4. Numerical example

In this section, a numerical example is used to illustrate the proposed method. In particular, the numerical example is used to compare our controller and the optimal controllers in Anderson and Moore (1971) and Kurtaran and Sivan (1974), and exhibit that our controller achieves a good performance.

Consider a linear time-invariant system (1)–(3) given by

$$A_t = \left[\begin{array}{ccc:ccc} -0.1\cdot a & 0.2\cdot a & -0.9 & -0.5 & 0.2 & -0.1 \\ 0.1\cdot a & 0.3\cdot a & 0.6 & 0.2 & 0 & 0.3 \\ 0 & -0.3 & 0.4\cdot b & 0.6\cdot b & 0.8 & -0.1 \\ \hdashline 0 & 0 & 0.2\cdot b & 0.3\cdot b & 0.4 & 0.1 \\ 0.3 & 0.2 & 0.1\cdot b & 0.5\cdot b & 0.7 & 0.8 \\ 0.2 & -0.3 & 0 & 0.1 & 0.2 & -0.1 \end{array}\right],$$

$$B_t = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

$$C_t = \left[\begin{array}{ccc:ccc} 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 \end{array}\right],$$

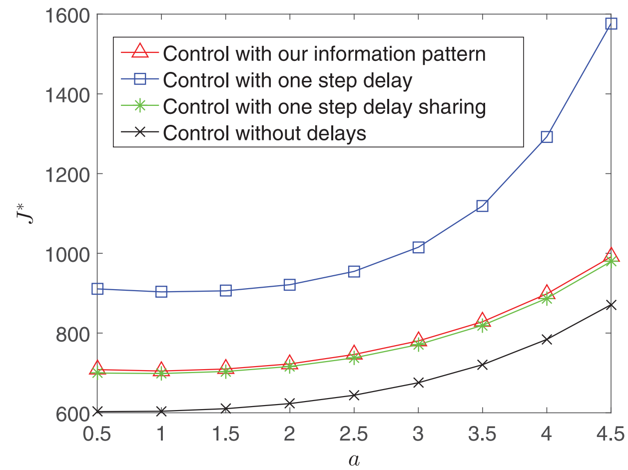

Figure 2. The values of performance criterion with respect to the parameter a, where $b = 1$.

where a and b are parameterised scalars, which will take different values for comparison purpose. The weight matrices in the quadratic performance criterion (11) are chosen to be

$$Q_t = \left[\begin{array}{ccc:ccc} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right], \quad R_t = \left[\begin{array}{ccc:cc} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

The noises ω_t and ν_t are assumed to be Gaussian noises with zero-mean and identity covariance. The time horizon is chosen to be $N = 20$. The gain matrices F_t and $H_t = \operatorname{diag}\{T_t^{\{1\}}, T_t^{\{2\}}\}$ are computed based on Algorithm 1 and Algorithm 2. In Algorithm 1, the stopping rule is chosen to be $\operatorname{tr}\left[\Lambda^1 F_{t-1}\Lambda^2(\Lambda^1 F_{t-1}\Lambda^2)^{\mathsf{T}}\right] + \operatorname{tr}[\Lambda^3 F_{t-1}\Lambda^4(\Lambda^3 F_{t-1}\Lambda^4)^{\mathsf{T}}] + \operatorname{tr}\left[\tilde{\Lambda}^1 H_t \tilde{\Lambda}^2(\tilde{\Lambda}^1 H_t \tilde{\Lambda}^2)^{\mathsf{T}}\right] + \operatorname{tr}[\tilde{\Lambda}^3 H_t \tilde{\Lambda}^4(\tilde{\Lambda}^3 H_t \tilde{\Lambda}^4)^{\mathsf{T}}] < \delta$, where $\delta = 0.0001$. The initial penalty parameter is $c_0 = 2$ and the growth parameter is $\eta = 2$. In Algorithm 2, the iteration step length is $\gamma = 0.0005$. Given the system parameters and algorithm parameters, our controller is designed successfully for two cases: (1) $b = 1$, and a takes the values in $\{0.5, 1, 1.5, \dots, 4.5\}$, respectively; (2) $a = 1$, and b takes the values in $\{0.3, 0.6, 0.9, \dots, 2.7\}$, respectively, where a and b are the scalars in system matrix A.

Now we compare the performance criterion achieved by our designed controller (47) and the optimal controllers in Anderson and Moore (1971) and Kurtaran and Sivan (1974). The comparison result is given by Figures 2–3. In Figures 2 and 3, the values of the performance criterion with the information pattern $\{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$, $\{y_{0:\,t-1}\}$, $\{y_t^i, y_{0:t-1}\}$ and $\{y_{0:\,t}\}$ are depicted

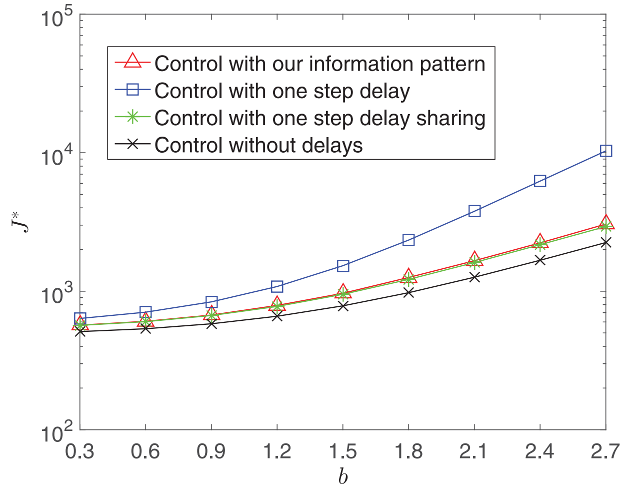

Figure 3. The values of performance criterion with respect to the parameter b, where $a = 1$.

by the solid lines marked by triangle, square, star and cross, respectively.

In theory, optimal controllers using more information should achieve better performance. In Figures 2 and 3, the controller in Anderson and Moore (1971) based on the full information $\{y_{0:\,t}\}$ has the best performance as expected. The controller in Kurtaran and Sivan (1974) based on the information $\{y_t^i, y_{0:t-1}\}$ is worse than the controller with $\{y_{0:\,t}\}$ in Anderson and Moore (1971), since $\{y_{0:t}\} \supsetneqq \{y_t^i, y_{0:t-1}\}$. Our controller is based on the information $\{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$, and has poorer performance than the controller with $\{y_t^i, y_{0:t-1}\}$ in Kurtaran and Sivan (1974). The reason is that the information set $\{y_t^i, y_{0:t-1}\}$ is equivalent to $\{y_t^i, \hat{x}_{0:t}\}$ (see Assertion in Kurtaran & Sivan, 1974) and that $\{y_t^i, \hat{x}_{0:t}\} \supsetneqq \{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$. Our controller outperforms the controller based on $\{y_{0:\,t\,-\,1}\}$ in Anderson and Moore (1971). The information sets $\{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$ and $\{y_{0:\,t\,-\,1}\}$ have not any inclusion relation.

Based on qualitative analysis, our controller is always close to the controller with $\{y_t^i, y_{0:t-1}\}$ in Kurtaran and Sivan (1974). Compared to $\{y_t^i, \hat{x}_{0:t}\}$, our information structure $\{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$ is more practical. Under the information structure $\{y_t^i, \hat{x}_t^i, \hat{x}_{0:t-1}\}$, the estimator of each subsystem only needs to estimate the local subsystem state instead of the global system state in Kurtaran and Sivan (1974). In addition, our controller is almost always much better than the controller with $\{y_{0:\,t\,-\,1}\}$ in Anderson and Moore (1971). This shows that the controller designed in this paper achieves a good performance.

5. Conclusions

This paper studied the output-feedback LQG control problem with a new and more practical information pattern. A general two-player system was considered and a novel controller structure was proposed. Using the method of independence decomposition, the form of the optimal controller was established. The gains of the optimal controller satisfying sparse constraint were constructed by penalty function algorithm. The augmented objective function in penalty function algorithm was minimised by the gradient descent method. In addition, the value of the performance criterion achieved with the designed controller was derived. Under our design, the local estimator only estimates its own subsystem state. This implied that the computational burden of state estimation was reduced compared to the existing related results. Also, the system achieved a good performance under our designed controller, which has been illustrated by the numerical example.

In the future, we will extend this work to general large-scale systems composed of multiple subsystems, the random multiple communication delays and the information pattern arising from a directed connected communication graph.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work in this paper was financially supported by National Natural Science Foundation of China [grant number 61374026].

References

Aldeen, M., Saha, S., Alpcan, T., & Evans, R. (2015). New online voltage stability margins and risk assessment for multi-bus smart power grids. *International Journal of Control, 88*(7), 1338–1352.

Anderson, B. D., & Moore, J. B. (1971). *Linear optimal control.* Englewood Cliffs, NJ: Prentice-Hall.

Åström, K. J. (2012). *Introduction to stochastic control theory.* New York, NY: Courier Corporation.

Athans, M. (1967). The matrix minimum principle. *Information and Control, 11*(5), 592–606.

Blondel, V. D., & Tsitsiklis, J. N. (2000). A survey of computational complexity results in systems and control. *Automatica, 36*(9), 1249–1274.

Duan, Z., Wang, J., & Huang, L. (2007). Special decentralized control problems in discrete-time interconnected systems composed of two subsystems. *Systems & Control Letters, 56*(3), 206–214.

Esfahani, N., & Khorasani, K. (2016). A distributed model predictive control (MPC) fault reconfiguration strategy for formation flying satellites. *International Journal of Control, 89*(5), 960–983.

Fax, J. A., & Murray, R. M. (2004). Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control, 49*(9), 1465–1476.

Ho, Y. C., & Chu, K.-H. (1972). Team decision theory and information structures in optimal control problems–Part I. *IEEE Transactions on Automatic Control, 17*(1), 15–22.

Jin, X., & Ray, A. (2014). Navigation of autonomous vehicles for oil spill cleaning in dynamic and uncertain environments. *International Journal of Control, 87*(4), 787–801.

Kurtaran, B.-Z., & Sivan, R. (1974). Linear-quadratic-gaussian control with one-step-delay sharing pattern. *IEEE Transactions on Automatic Control, 19*(5), 571–574.

Lamperski, A., & Doyle, J. C. (2012). Dynamic programming solutions for decentralized state-feedback LQG problems with communication delays. In *Proceedings of American control conference* (pp. 6322–6327). Montreal: IEEE.

Lamperski, A., & Lessard, L. (2015). Optimal decentralized state-feedback control with sparsity and delays. *Automatica, 58*, 143–151.

Lessard, L., & Lall, S. (2014). An algebraic approach to the control of decentralized systems. *IEEE Transactions Control of Network Systems, 1*(4), 308–317.

Lessard, L., & Lall, S. (2015). Optimal control of two-player systems with output feedback. *IEEE Transactions on Automatic Control, 60*(8), 2129–2144.

Matni, N., & Doyle, J. C. (2013). Optimal distributed LQG state feedback with varying communication delay. In *Proceedings of IEEE conference on decision and control* (pp. 5890–5896). Florence: IEEE.

Nayyar, A., Mahajan, A., & Teneketzis, D. (2011). Optimal control strategies in delayed sharing information structures. *IEEE Transactions on Automatic Control, 56*(7), 1606–1620.

Nocedal, J., & Wright, S. (2006). *Numerical optimization*. New York, NY: Springer Science & Business Media.

Papadimitriou, C. H., & Tsitsiklis, J. (1986). Intractable problems in control theory. *SIAM Journal on Control and Optimization, 24*(4), 639–654.

Qi, X., Salapaka, M. V., Voulgaris, P. G., & Khammash, M. (2004). Structured optimal and robust control with multiple criteria: A convex solution. *IEEE Transactions on Automatic Control, 49*(10), 1623–1640.

Rotkowitz, M., & Lall, S. (2006). A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control, 51*(2), 274–286.

Sandell, N., Varaiya, P., Athans, M., & Safonov, M. (1978). Survey of decentralized control methods for large scale systems. *IEEE Transactions on Automatic Control, 23*(2), 108–128.

Snyman, J. (2005). *Practical mathematical optimization: An introduction to basic optimization theory and classical and new qradient-based algorithms*. New York, NY: Springer Science & Business Media.

Toda, M., & Aoki, M. (1975). Second-guessing technique for stochastic linear regulator problems with delayed information sharing. *IEEE Transactions on Automatic Control, 20*(2), 260–262.

Zheng, Y., & Feng, G. (2011). Stabilisation of second-order LTI switched positive systems. *International Journal of Control, 84*(8), 1387–1397.