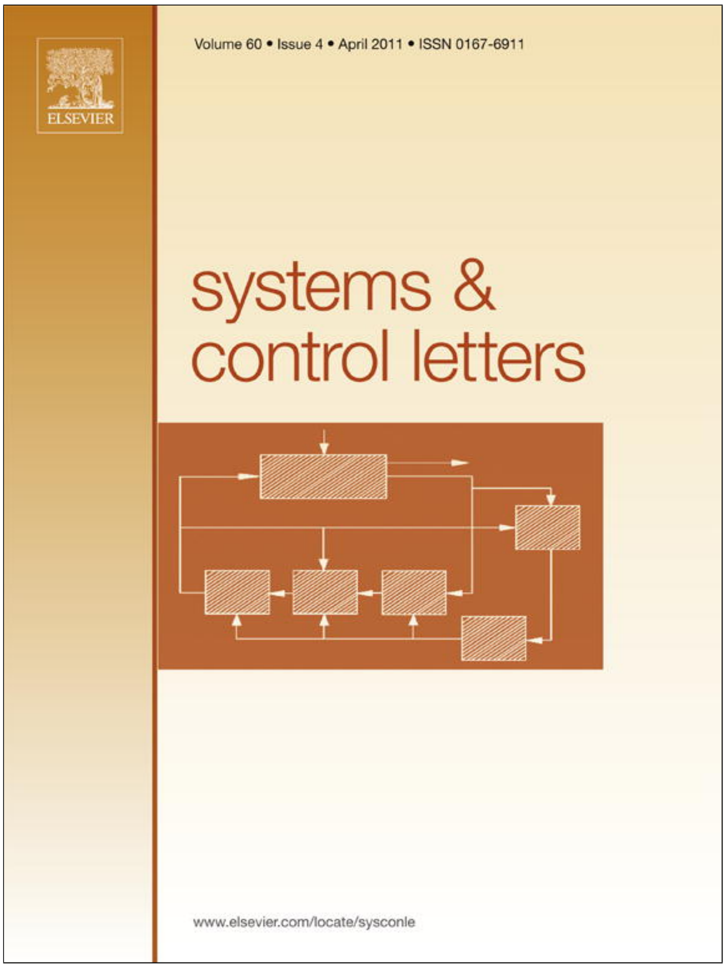
Volume 60 • Issue 4 • April 2011 • ISSN 0167-6911
ELSEVIER
systems & control letters
www.elsevier.com/locate/sysconle

Systems & Control Letters 60 (2011) 285–293

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Suboptimal decentralized control over noisy communication channels

Alireza Farhadi [a,*], N.U. Ahmed [b,1]

[a] INRIA, Grenoble, France

[b] SITE, University of Ottawa, Ottawa, Ontario, Canada

ARTICLE INFO

Article history:
Received 12 May 2009
Received in revised form
17 January 2011
Accepted 9 February 2011
Available online 16 March 2011

Keywords:
Decentralized control
AWGN channel
Linear systems

ABSTRACT

In this paper we present a technique for the design of decentralized controllers for mean square stability of a large scale system with cascaded clusters of subsystems. Each subsystem is linear and time-invariant and both system and measurement are subject to Gaussian noise. For stability analysis of this system we consider the effects of Additive White Gaussian Noise (AWGN) channels used for exchanging information between subsystems.

1. Introduction

In recent years, the development of Micro-Electro Mechanical Systems (MEMS) has made it possible to deploy these small sized embedded devices in distributed parameter systems for efficient control. Some examples are: smart mechanical structures and distributed flow control [1]. Each micro-electro mechanical system consists of sensors, a data processor, a communication unit, and an actuator. MEMS collaborate with each other towards a common goal by exchanging observation and control signals. Due to limited power supply of MEMS, transmission in systems equipped with these devices is limited by communication constraints over short distances. This necessitates the creation of a co-design framework to integrate the control and communication requirements in systems controlled by networks of MEMS. The objective of this paper is to develop such a framework for a large scale system with cascaded clusters of subsystems which is controlled by a network of MEMS. In this system, subsystems are linear time-invariant and each subsystem is controlled by a micro-electro mechanical device attached to it. Subsystems are subject to Gaussian process noise and Gaussian measurement noise and interconnected via Additive White Gaussian Noise (AWGN) channels. For this large scale system we present a decentralized technique for design of controllers, encoders and decoders for mean square stability and reliable data reconstruction. These policies are executed by MEMS.

In contrast with similar works, in Refs. [2–5], optimal stabilizing controllers of Linear Quadratic Gaussian (LQG) team decision problems were presented without considering the effects of communication constraints. In the presence of limited capacity finite alphabet channels, stabilizing controllers were also given in [6–11]. These results are mostly concerned with deterministic systems (e.g., [6–8,10]). In this paper we generalize the results of [2–5] by considering the effects of communication imperfections. We also generalize the results of [6–11] by addressing the stability problem of stochastic dynamic systems over AWGN channels.

The paper is organized as follows: In Section 2, the problem formulation is given. In Section 3, design techniques for decentralized controllers, encoders, and decoders are presented. In Section 4, we compare the rate requirement for stability using the proposed decentralized technique with the minimum rate requirement for stability using the centralized technique of [12]. The paper is concluded in Section 5 with a summary of proposed techniques as developed here.

2. Problem formulation

Throughout, certain conventions are used: Sequences of Random Vectors (R.V.'s) are denoted by $y(T) \equiv (y_0, y_1, \ldots, y_T)$ or $Y(T) \equiv (Y_0, Y_1, \ldots, Y_T)$ for $T \in \mathbf{N}_+ \equiv \{0, 1, 2, \ldots\}$. A logarithm of base 2 is denoted by $\log(\cdot)$ and the Euclidean norm with weight R on any finite dimensional space is denoted by $\|\cdot\|_R$. The space of all matrices $A \in \Re^{q\times o}$ is denoted by $M(q \times o)$ and the transpose of A, where A can be either a matrix or a vector, is denoted by A'. The identity matrix with dimension $M(q \times q)$ is denoted by I_q, the inverse of a square matrix $A \in M(q \times q)$ is denoted by A^{-1}, and $\mathrm{diag}(\cdot)$ denotes a block diagonal matrix. The

* Corresponding author. Tel.: +33 6 58 43 68 36.
E-mail addresses: farhadialireza@yahoo.com (A. Farhadi), ahmed@site.uottawa.ca (N.U. Ahmed).
1 Tel.: +1 613 562 5800x 6222; fax: +1 613 562 5175.

0167-6911/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2011.02.007

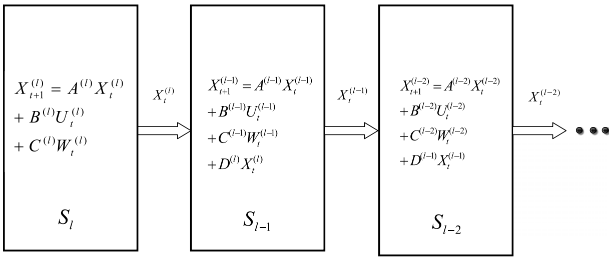

Fig. 1. A cascade system.

covariance of a R.V. X, its expected value and its density function are denoted by $\mathrm{Cov}(X)$, $E[X]$, and f_X, respectively. The joint density function of two R.V.'s X and Y, the conditional density function and the cross covariance function are denoted by $f_{X,Y}, f_{Y|X}$ and $\mathrm{Cov}(X,Y)$, respectively. The Gaussian density function with mean $\bar{x}$ and covariance $\bar{V}$ is denoted by $N(\bar{x},\bar{V})$. Gaussian R.V. X, which is described by the density function $N(\bar{x},\bar{V})$, is denoted by $X \sim N(\bar{x},\bar{V})$.

Now, consider the cascade system of Fig. 1. This system consists of a set of disjoint clusters (of subsystems) $\mathcal{S}_r$, $r=1,2,\ldots,l$, in which no subsystem is contained in more than one cluster. For this system let $X_t^{(r)}$ denote the vector of state variables of all subsystems in cluster $\mathcal{S}_r$ at time t. Similarly, let $U_t^{(r)}$ denote the vector of control signals and $W_t^{(r)}$ denote the vector of process noises. As shown in Fig. 1, cluster $\mathcal{S}_{r+1}$ affects the dynamics of cluster $\mathcal{S}_r$ via state variables; while cluster $\mathcal{S}_r$ does not affect the dynamics of cluster $\mathcal{S}_{r+1}$, and this is true for all r. Cascade systems have been considered by many authors (e.g., [13–15]). This type of system is very common in process control, control of servo-mechanical systems, and assembly lines.

Cluster $\mathcal{S}_r$ is described by the following dynamic model

$$(\mathcal{S}_r):\begin{cases}X_{t+1}^{(r)}=A^{(r)}X_t^{(r)}+B^{(r)}U_t^{(r)}+C^{(r)}W_t^{(r)}+D^{(r+1)}X_t^{(r+1)},\\ r=1,2,\ldots,l,\end{cases}\tag{1}$$

where the matrices $A^{(r)}$, $B^{(r)}$ and $C^{(r)}$ are the system matrices of cluster $\mathcal{S}_r$, and the coupling matrix $D^{(r+1)}$ represents the effect of the state variables of cluster $\mathcal{S}_{r+1}$ on the subsystems of cluster $\mathcal{S}_r$. Note that the R.V.'s $X_0^{(r)}$ and $X_0^{(\acute{r})}$ ($\acute{r}\neq r$) are statistically independent, $D^{(l+1)}=0$, and $B^{(r)}$ is full column rank.

Thus, the entire system is described by the following system of equations:

$$(\mathcal{S}):\{X_{t+1}=AX_t+BU_t+CW_t,\tag{2}$$

where $X_t=\left(X_t^{(1)'}\ \cdots\ X_t^{(l)'}\right)'$ is the state of the full (large scale) system, $U_t=\left(U_t^{(1)'}\ \cdots\ U_t^{(l)'}\right)'$ is the control vector, $W_t=\left(W_t^{(1)'}\ \cdots\ W_t^{(l)'}\right)'$ is the process noise, and the matrices A, B, and C are given by the following blocks:

$$A=\begin{pmatrix}A^{(1)} & D^{(2)} & 0 & 0 & \cdots & 0\\ 0 & A^{(2)} & D^{(3)} & 0 & \cdots & 0\\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & A^{(l)}\end{pmatrix},\tag{3}$$

$$\begin{aligned}B&=\mathrm{diag}(B^{(1)},B^{(2)},\ldots,B^{(l)}),\\ C&=\mathrm{diag}(C^{(1)},C^{(2)},\ldots,C^{(l)}).\end{aligned}\tag{4}$$

Often the coupling between clusters is relatively small. We refer to this case as the weakly cascaded case which is defined as follows:

Definition 2.1 (*Weakly Cascaded System*)**.** The large scale system described by (2)–(4) is said to be weakly cascaded if for every cluster $\mathcal{S}_r$ given by (1) the largest singular value of the coupling matrix $D^{(r+1)}$ is much smaller (e.g., at least thirteen times smaller) than the smallest singular value of the corresponding system matrix $A^{(r)}$.

The large scale system of Fig. 1 consists of M subsystems which are distributed in l clusters: $\mathcal{S}_1,\ldots,\mathcal{S}_l$, as described above. Let subsystem s_i be in cluster $\mathcal{S}_r$ ($s_i\in\mathcal{S}_r$). Also, let $x_t^{(i)}\in\Re^{n_i}$ be the state, $u_t^{(i)}\in\Re^{d_i}$ be the control, $w_t^{(i)}\in\Re^{g_i}$ be the process noise, $y_t^{(i)}\in\Re^{m_i}$ be the observation, and $v_t^{(i)}\in\Re^{h_i}$ be the measurement noise of subsystem s_i. Moreover, let O_i contain all subsystems, except subsystem s_i, which are either in cluster $\mathcal{S}_r$ or in cluster $\mathcal{S}_{r+1}$. Similarly, let $\tilde{O}_i$ be the set of all subsystems, except subsystem s_i, which are in cluster $\mathcal{S}_r$ only ($\tilde{O}_i\subseteq O_i$). Subsystem s_i is linear time-invariant and both the system and measurement are subject to Gaussian noise as described below:

$$(s_i):\begin{cases}x_{t+1}^{(i)}=A_ix_t^{(i)}+B_iu_t^{(i)}+C_iw_t^{(i)}+\sum_{k\in O_i}D_{ik}x_t^{(k)}+\sum_{j\in\tilde{O}_i}E_{ij}u_t^{(j)}\\ y_t^{(i)}=F_ix_t^{(i)}+G_iv_t^{(i)},\qquad x_0^{(i)}=\xi_0^{(i)},\end{cases}\tag{5}$$

where $x_t^{(k)}\in\Re^{n_k}$ ($k\in O_i, k\neq i$) is the state of the kth subsystem and $u_t^{(j)}\in\Re^{d_j}$ ($j\in\tilde{O}_i, j\neq i$) is the control signal of the jth subsystem, which affect the ith subsystem dynamic. In (5), $A_i\in M(n_i\times n_i)$, $B_i\in M(n_i\times d_i)$, $C_i\in M(n_i\times g_i)$, $F_i\in M(m_i\times n_i)$ and $G_i\in M(m_i\times h_i)$ are system matrices of subsystem s_i. The matrices $D_{ik}\in M(n_i\times n_k)$ and $E_{ij}\in M(n_i\times d_j)$ are coupling matrices. Furthermore, $\xi_0^{(i)}\sim N(\bar{x}_0^{(i)},\bar{V}_0^{(i)})$, $w_t^{(i)}$ i.i.d. $\sim N(0,\Sigma_w^{(i)})$ and $v_t^{(i)}$ i.i.d. $\sim N(0,\Sigma_v^{(i)})$. The sequences $\{w^{(i)}(t),v^{(i)}(t),w^{(b)}(t),v^{(b)}(t)\}_{t\in\mathbf{N}_+}$, $b(\neq i)\in\{1,2,\ldots,M\}$, are mutually independent. They are also independent of the initial state $\xi_0^{(i)}$. But the R.V.'s $\xi_0^{(i)}$ and $\xi_0^{(j)}$ may be statistically dependent with known cross covariance function $\mathrm{Cov}(\xi_0^{(i)},\xi_0^{(j)})$. Note that the system and coupling matrices, the vector $\bar{x}_0^{(i)}$ and the matrices $\bar{V}_0^{(i)}$, $\Sigma_w^{(i)}$, $\Sigma_v^{(i)}$ and $\mathrm{Cov}(\xi_0^{(i)},\xi_0^{(j)})$ are fixed and known to all subsystems of cluster $\mathcal{S}_r$. Also, note that $u_t^{(i)}$ is the control signal produced by the micro-electro mechanical device responsible for controlling subsystem s_i and $y_t^{(i)}$ is the observation made by the sensors of this device.

Now, consider subsystems s_i and $s_j, j\ (\neq i)$ in cluster $\mathcal{S}_r$. Subsystem s_i broadcasts an encoded observation signal to subsystem s_j. The communication link between these two subsystems is modeled by a multi-input, multi-output AWGN channel with the channel input denoted by $T_t^{(i)}$ and the channel output by $R_t^{(ji)}$. This channel is subject to path loss and input power constraint and is described by

$$\begin{aligned}&R_t^{(ji)}=h^{(ji)}T_t^{(i)}+\zeta_t^{(ji)},\qquad T_t^{(i)}\equiv\mathcal{E}_t^{(ji)}(y_t^{(i)})\in\Re^{p_i},\\ &R_t^{(ji)}\in\Re^{q_i},\qquad E[\|T_t^{(i)}\|^2]\leq P_t^{(i)}<\infty,\end{aligned}\tag{6}$$

where $\zeta_t^{(ji)}\in\Re^{q_i}$ i.i.d. $\sim N(0,\Gamma^{(ji)})$ is the channel noise (which is independent of all initial states, process noises and measurement noises), $h^{(ji)}\equiv 1/(d_{ji})^{a_{ji}}$ is the channel gain described by the line of sight distance d_{ji} and the path loss factor $a_{ji}\in\{0,1,2,\ldots\}$, $\mathcal{E}_t^{(ji)}(\cdot)$ is the encoding function, and $P_t^{(i)}$ is the channel input power constraint. Throughout, it is assumed that the parameters describing the channel, i.e., $P_t^{(i)}$, $\Gamma^{(ji)}$, d_{ji}, a_{ji}, and $h^{(ji)}\neq 0$ are known to subsystems s_i and s_j. Also, the channel noises of different communication links are independent. For simplicity of presentation, it is also assumed that the control signals are exchanged without communication constraints. Fig. 2 illustrates

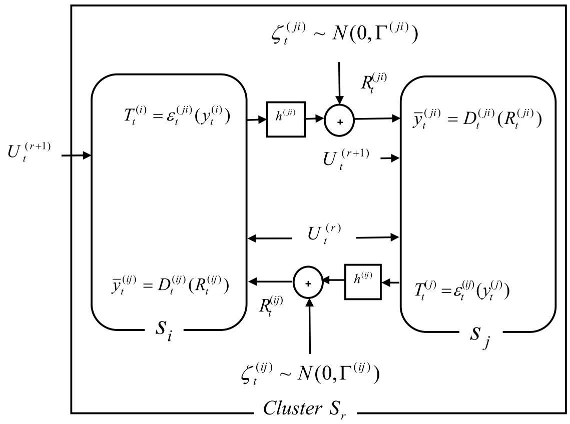

Fig. 2. Flow of information between subsystems s_i and s_j in cluster $\mathcal{S}_r$ at time t.

the flow of information between subsystems s_i and s_j in cluster $\mathcal{S}_r$. Note that in Fig. 2, $\bar{y}_t^{(ji)} \equiv \mathcal{D}_t^{(ji)}(R_t^{(ji)})$ denote the reconstructed version of the observation signal $y_t^{(i)}$ at subsystem s_j for a suitable decoding function $\mathcal{D}_t^{(ji)}(\cdot)$ to be constructed later in the following section. According to Fig. 2, the information available at subsystem s_j at time t is

$$\mathcal{F}_t^{(r,j)} \equiv \sigma\{U^{(r+1)}(t), U^{(r)}(t), y^{(j)}(t), \bar{y}^{(ji)}(t), \forall s_i \in \mathcal{S}_r\ (i \neq j)\}, \quad \forall s_j \in \mathcal{S}_r, r = 1, \dots, l, \tag{7}$$

where $\sigma\{\cdot\}$ denotes the sigma-algebra generated by the sequences as indicated above. As defined earlier, sequences of R.V.'s are denoted by $y(t) \equiv (y_0, y_1, \dots, y_t)$ and $Y(t) \equiv (Y_0, Y_1, \dots, Y_t)$. Therefore, the information pattern (7) includes the present as well as past information. The information structure (7) is a non-classical information pattern in the sense that the action of the controller of subsystem $s_j \in \mathcal{S}_r$ affects the information structure of subsystem $s_e \in \mathcal{S}_{\tilde{r}}, \tilde{r} \in \{1, 2, \dots, r-2\}$, and, in general, there is no way for this subsystem to infer the information available at subsystem s_j. Under this information structure, linear controllers are not generally optimal as shown in [16].

The objective of this paper is to find a control sequence $\{u_t^{(i)}; i = 1, 2, \dots, M,\ t \in \mathbf{N}_+\}$ that stabilizes system (2) in a mean square sense in the presence of AWGN channels and the information pattern (7). That is, $\sup_{t \in \mathbf{N}_+} E[\|x_t^{(i)}\|^2] < \infty,\ \forall i \in \{1, 2, \dots, M\}$.

In this paper stabilizing controllers, subject to the above information pattern, are obtained from a solution to the optimization problem associated to the following quadratic payoff functional

$$J = \sum_{r=1}^{l} J^{(r)}, \qquad J^{(r)} \equiv \lim_{T \to \infty} \frac{1}{T+1} \sum_{t=0}^{T} E[\|X_t^{(r)}\|_{Q_r}^2 + \|U_t^{(r)}\|_{R_r}^2], \quad r = 1, \dots, l, \tag{8}$$

where the weighting matrices $Q_r = Q_r'$ and $R_r = R_r'$ are positive definite. Later we show that the limit in (8) exists for the solution that we find in the next section, in which this solution results in mean square stability.

Optimal controls of linear quadratic Gaussian team decision problems have been presented in several references. In [3] Ho and Chu addressed the case of partially nested information pattern and in [4] Sandell and Athans addressed the case of one step delayed information structure. Also, in [17,18] Bansal and Basar addressed the case of non-classical information pattern when the quadratic cost functional does not include cross product terms between control variables. In all these references, the optimal controller is linear. In the presence of AWGN channels, Yuksel and Tatikonda [19] presented a nonlinear policy (a nonlinear coding scheme and a certainly equivalent controller) for a distributed multi-sensor, single control system. As shown in [19] in distributed multi-sensor systems, for minimizing quadratic payoff functional over Gaussian channels, linear policy (linear coding and controller) is not optimal. It was also illustrated that the upper bound on the cost functional with the proposed nonlinear policy is slightly smaller than the upper bound with the best linear policy. In large scale systems controlled by MEMS which involve components with limited processing capacity, it is important to find simple policies (e.g., linear policies) for exchanging information and control. Therefore, in this paper we use linear policies. Consequently, in view of the above discussion, here controllers are not necessarily optimal.

3. Control through communication channels

In this section we present decentralized controllers, encoders and decoders for mean square stability and reliable communication when the communication links are AWGN channels.

Throughout this section it is assumed that subsystems s_i and s_j in cluster $\mathcal{S}_r$ have overlapped communication range so that they can exchange information. Each subsystem in cluster $\mathcal{S}_r$ broadcasts encoded observation signal to other subsystems. Hence, there is a possibility of collision in the broadcast information. In large scale systems controlled by MEMS, which involve components with limited power supply, it is known that the Time Division Multiple Access (TDMA) scheme [20] is more energy efficient than other protocols for exchanging information without collision. In this paper it is assumed that there exists a suitable TDMA scheme [20] under which information is exchanged.

For simplicity of presentation, without loss of generality, from now on we assume that the AWGN channel (6) is single-input, single-output, and the observation signals $y_t^{(i)}$'s are scalars. Furthermore, system (2) consists of three clusters ($l = 3$), in which each cluster includes two subsystems (the general case can be treated similarly). Cluster $\mathcal{S}_1$ includes subsystems s_1, s_2, cluster $\mathcal{S}_2$ includes subsystems s_3, s_4, and cluster $\mathcal{S}_3$ includes subsystems s_5, s_6. It is assumed that subsystem s_3 is the closest subsystem of cluster $\mathcal{S}_2$ to the subsystems of cluster $\mathcal{S}_1$. Similarly, subsystem s_5 is the closest subsystem of cluster $\mathcal{S}_3$ to the subsystems of cluster $\mathcal{S}_2$. Note that there is no role played by the order in which subsystems are considered in the process of obtaining stabilizing controllers which is summarized as follows: (1) For cluster $\mathcal{S}_r$ we choose the control vector $U_t^{(r)}$ such that the coupling effects from other clusters are compensated. Then, (2) stabilizing controllers for cluster $\mathcal{S}_r$ are obtained (independently of other clusters) from a suboptimal control solution for the payoff functional $J^{(r)}$. In obtaining this solution we use linear encoders and decoders. Stabilizing controllers, encoders, and decoders are given next.

3.1. Encoders, decoders, and controllers

In this section we present a methodology for design of encoders, decoders and controllers. In view of the coupling matrix (3), it is convenient to start from cluster $\mathcal{S}_l = \mathcal{S}_3$, which is not affected by other clusters, and then proceed to $\mathcal{S}_2$, and finally $\mathcal{S}_1$.

Design methodology for cluster $\mathcal{S}_3$: Consider cluster $\mathcal{S}_3$, as described by (1), which is reproduced here for convenience of reference.

$$X_{t+1}^{(3)} = A^{(3)} X_t^{(3)} + B^{(3)} U_t^{(3)} + C^{(3)} W_t^{(3)}. \tag{9}$$

Recall that for each $t \geq 0$ the information pattern of subsystem s_i ($i \in \{5, 6\}$) of cluster $\mathcal{S}_3$ is $\mathcal{F}_t^{(3,i)} \equiv \sigma\{U^{(3)}(t), y^{(i)}(t), \bar{y}^{(ij)}(t)\}$

where $\bar{y}_t^{(ij)}$ is the reconstructed version of the observation signal $y_t^{(j)}, j\ (\neq i) \in \{6, 5\}$ at subsystem s_i. For this cluster we use the following encoders and decoders:

Encoders of cluster $\mathcal{S}_3$: Let $\beta_t^{(ji)} \geq 0$ denote the encoding gain. Subsystem s_i produces the mean square state estimate $\hat{x}_t^{(i)} \equiv E[x_t^{(i)}|\mathcal{F}_{t-1}^{(3,i)}]$ using available information. This estimate is used to produce the innovation process $k_t^{(i)} \equiv y_t^{(i)} - F_i\hat{x}_t^{(i)}$ which is used in the encoding function as described below:

$$\mathcal{E}_t^{(ji)}(y_t^{(i)}) \equiv \beta_t^{(ji)} k_t^{(i)}. \tag{10}$$

The message $T_t^{(i)} = \mathcal{E}_t^{(ji)}(y_t^{(i)})$ is then broadcast to other subsystem of cluster $\mathcal{S}_3$.

Decoders of cluster $\mathcal{S}_3$: Subsystem s_j receives $R_t^{(ji)} = h^{(ji)} T_t^{(i)} + \zeta_t^{(ji)}$ through the AWGN channel (6). Let $\gamma_t^{(ji)} \geq 0$ denote the decoding gain and $\bar{k}_t^{(ji)} \equiv (h^{(ji)})^{-1}\gamma_t^{(ji)} R_t^{(ji)}$ be the reconstructed version of the innovation sequence $k_t^{(i)}$ at subsystem s_j. The decoding function for subsystem s_j is then given by

$$\bar{y}_t^{(ji)} = \mathcal{D}_t^{(ji)}(R_t^{(ji)}) \equiv \bar{k}_t^{(ji)} + F_i\hat{x}_t^{(i)}. \tag{11}$$

Thus, the decoder output is the reconstructed version of the observation signal $y_t^{(i)}$ which is denoted by $\bar{y}_t^{(ji)}$. Note that the decoding function (11) involves the state estimate $\hat{x}_t^{(i)}$. As justified later, this estimate is made available at subsystem s_j via the control signal $u_t^{(i)}$. Throughout, it is assumed that the encoding and decoding gains are known to the subsystems of cluster $\mathcal{S}_3$.

Choice of encoding and decoding gains: In this section we choose the encoding gain $\beta_t^{(ji)}$ and the decoding gain $\gamma_t^{(ji)}$ so that we have reliable communication as defined below:

Definition 3.1 (*Reliable Communication*)**.** Let

$$\rho_{ji} \equiv \frac{1}{T+1}\sum_{t=0}^{T} E[\|y_t^{(i)} - \bar{y}_t^{(ji)}\|^2]$$
$$= \frac{1}{T+1}\sum_{t=0}^{T} E[\|k_t^{(i)} - \bar{k}_t^{(ji)}\|^2],$$
$$(k_t^{(i)} = y_t^{(i)} - F_i E[x_t^{(i)}|\mathcal{F}_{t-1}^{(r,i)}]) \tag{12}$$

denote the distortion measure describing the mismatch between the message from subsystem $s_i \in \mathcal{S}_r$ and its reconstructed version at subsystem $s_j \in \mathcal{S}_r$. For any given distortion level $\delta^{(ji)} > 0$, communication from subsystem s_i to subsystem s_j is said to be reliable if $\rho_{ji} \leq \delta^{(ji)}$.

Here, we determine the encoding and decoding gains $\beta_t^{(56)}$ and $\gamma_t^{(56)}$ (respectively) for reliable transmission from subsystem s_6 to subsystem s_5. Encoding and decoding gains for other transmissions are obtained similarly.

Consider encoder (10) and decoder (11). It is easy to verify that the innovation sequence $k_t^{(6)}$ given in (10) is orthogonal and Gaussian with density function $N(0, \Psi_t^{(6)})$, where the variance is given by $\Psi_t^{(6)} \equiv F_6\Theta_t^{(6)}F_6' + G_6\Sigma_v^{(6)}G_6'$. The matrix $\Theta_t^{(6)} \equiv \mathrm{Cov}(x_t^{(6)} - \hat{x}_t^{(6)}|\mathcal{F}_{t-1}^{(3,6)}) \in M(n_6 \times n_6)$ is the second diagonal element of the block matrix $\Xi_t^{(6)} \equiv \mathrm{Cov}(X_t^{(3)} - E[X_t^{(3)}|\mathcal{F}_{t-1}^{(3,6)}]|\mathcal{F}_{t-1}^{(3,6)}) = \begin{pmatrix} \hat{\Theta}_t^{(6)} & \tilde{\Theta}_t^{(6)} \\ \bar{\Theta}_t^{(6)} & \Theta_t^{(6)} \end{pmatrix} \in M\Big((n_5+n_6) \times (n_5+n_6)\Big)$. As shown in the next section $\Xi_t^{(6)}$ is the solution of a filtered Riccati equation and is finite. The finiteness is justified later.

As described in (11) the reconstructed version of the sequence $k^{(6)}(T)$ at subsystem s_5 is denoted by $\bar{k}^{(56)}(T) \equiv \{\bar{k}_t^{(56)},\ t \in [0, T]\}$. Let $I(k^{(6)}(T); \bar{k}^{(56)}(T))$ denote the mutual information between sequences $k^{(6)}(T)$ and $\bar{k}^{(56)}(T)$ [21]. From the standard definition of the rate distortion [22], the associated rate distortion function subject to the distortion measure ρ_{56} and distortion level $\delta^{(56)} > 0$ is given by

$$R^{k,\bar{k}}(T, \delta^{(56)}) \equiv \inf_{\rho_{56} \leq \delta^{(56)}} I(k^{(6)}(T); \bar{k}^{(56)}(T)), \tag{13}$$
$$R^{k,\bar{k}}(\delta^{(56)}) \equiv \lim_{T\to\infty} \frac{1}{T+1} R^{k,\bar{k}}(T, \delta^{(56)}).$$

From [21, Theorem 13.3.3] it follows that when the distortion level specified is less than the infimum of the variance of the innovation process $\Psi_t^{(6)}$, that is $\inf\{\Psi_t^{(6)},\ t \in \mathbf{N}_+\}$, we have $R^{k,\bar{k}}(T, \delta^{(56)}) = \frac{1}{2}\sum_{t=0}^{T} \log \frac{\Psi_t^{(6)}}{\delta^{(56)}}$ and hence

$$R^{k,\bar{k}}(\delta^{(56)}) \equiv \lim_{T\to\infty} \frac{1}{T+1} R^{k,\bar{k}}(T, \delta^{(56)}) = \frac{1}{2}\log \frac{\Psi_\infty^{(6)}}{\delta^{(56)}} \tag{14}$$

provided the limit, $\Psi_\infty^{(6)} = \lim_{t\to\infty} \Psi_t^{(6)}$, exists. The existence is justified later. Following [22, Theorem 4.3.2] one can justify that the conditional density that minimizes the mutual information (13) is given by $f^*_{\bar{k}^{(56)}(T)|k^{(6)}(T)} = \prod_{t=0}^{T} f^*_{\bar{k}_t^{(56)}|k_t^{(6)}}$ where $f^*_{\bar{k}_t^{(56)}|k_t^{(6)}} = N(\eta_t^{(56)} k_t^{(6)}, \eta_t^{(56)}\delta^{(56)}), \eta_t^{(56)} \equiv 1 - \frac{\delta^{(56)}}{\Psi_t^{(6)}}$. On the other hand, the conditional density of the reconstructed message $\bar{k}^{(56)}(T)$, given the transmitted message $k^{(6)}(T)$, is given by $f_{\bar{k}^{(56)}(T)|k^{(6)}(T)} = \prod_{t=0}^{T} f_{\bar{k}_t^{(56)}|k_t^{(6)}}$, where it follows from (10) and (11) that the conditional density $f_{\bar{k}_t^{(56)}|k_t^{(6)}}$ (of the reconstructed message given the message) is given by $f_{\bar{k}_t^{(56)}|k_t^{(6)}} = N(\beta_t^{(56)}\gamma_t^{(56)}k_t^{(6)}, (\gamma_t^{(56)}/h^{(56)})^2 \Gamma^{(56)})$. For reliable communication, we set $f^*_{\bar{k}_t^{(56)}|k_t^{(6)}} = f_{\bar{k}_t^{(56)}|k_t^{(6)}}$ and subsequently determine the encoding and decoding gains $\beta_t^{(56)}$ and $\gamma_t^{(56)}$, as follows:

$$\beta_t^{(56)} = \frac{1}{|h^{(56)}|}\sqrt{\frac{\Gamma^{(56)}\eta_t^{(56)}}{\delta^{(56)}}}, \qquad \gamma_t^{(56)} = |h^{(56)}|\sqrt{\frac{\eta_t^{(56)}\delta^{(56)}}{\Gamma^{(56)}}}. \tag{15}$$

In the following proposition we show that the encoding and decoding gains, as presented above, guarantee reliable communication, as specified in Definition 3.1.

Proposition 3.2. *Consider encoder* (10) *for subsystem* s_6 *and decoder* (11) *for subsystem* s_5. (i) *The encoding and decoding gains* (15)*, guarantee a reliable communication of the form* $\rho_{56} = \delta^{(56)}$ *by transmission with the capacity* $\mathcal{C}^{(56)} = R^{k,\bar{k}}(\delta^{(56)}) = \frac{1}{2}\log \frac{\Psi_\infty^{(6)}}{\delta^{(56)}} < \infty$ *(bits/time step).* (ii) *By transmission with the capacity* $\mathcal{C}^{(56)} = \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\hat{\delta}^{(56)}} \geq \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}}$ *(where* $\hat{\delta}^{(56)} \leq \delta^{(56)}$*), which is obtained by replacing* $\delta^{(56)}$ *with* $\hat{\delta}^{(56)}$ *in* (15)*, we have a reliable communication of the form* $\rho_{56} = \hat{\delta}^{(56)} \leq \delta^{(56)}$.

Proof. (i) By substituting the encoding and decoding gains $\beta_t^{(56)}$ and $\gamma_t^{(56)}$, respectively, as given by (15), in the distortion measure ρ_{56} as deduced from (12), it follows that, for the specified distortion level $\delta^{(56)}$, we have $\rho_{56} = \delta^{(56)}$. Thus, according to our Definition 3.1, we have reliable communication. Now we must show that we can achieve this reliable communication by transmission with the capacity $\mathcal{C}^{(56)} = R^{k,\bar{k}}(\delta^{(56)}) = \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}} < \infty$. Towards this goal, let us choose the channel input power equal to the power constraint $P_t^{(6)}$ itself, as defined by $P_t^{(6)} = E[\beta_t^{(56)}k_t^{(6)}]^2 < \infty$. By definition the channel capacity $\mathcal{C}^{(56)}$ is given by $\mathcal{C}^{(56)} \equiv \lim_{T\to\infty} \frac{1}{2(T+1)} \sum_{t=0}^{T} E\left[\log\left(1+\frac{(h^{(56)})^2 P_t^{(6)}}{\Gamma^{(56)}}\right)\right]$. Substituting the values of $P_t^{(6)}$ from above and $\beta_t^{(56)}$ from (15) and taking the limit we obtain $\mathcal{C}^{(56)} = \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}}$. As shown in (14), the rate distortion function of the innovation process is given by $R^{k,\bar{k}}(\delta^{(56)}) = \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}}$. Therefore, we have $\mathcal{C}^{(56)} = R^{k,\bar{k}}(\delta^{(56)}) = \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}}$. (ii) It follows similarly. □

In general, for any pair of systems $\{s_i, s_j\}$ in cluster $\mathcal{S}_r$, following the same procedure, the encoding and decoding gains for reliable communication are given by: $\beta_t^{(ji)} = \frac{1}{|h^{(ji)}|}\sqrt{\frac{\Gamma^{(ji)}\eta_t^{(ji)}}{\delta^{(ji)}}}$, $\gamma_t^{(ji)} = |h^{(ji)}|\sqrt{\frac{\eta_t^{(ji)}\delta^{(ji)}}{\Gamma^{(ji)}}}$, where $\eta_t^{(ji)} \equiv 1-\frac{\delta^{(ji)}}{\Psi_t^{(i)}}$, $\Psi_t^{(i)} \equiv \mathrm{Cov}(k_t^{(i)})$, $\delta^{(ji)} < \min_{t\in\mathbf{N}_+}\Psi_t^{(i)}$. Note that by finding solutions to the algebraic Riccati equations associated to the filtered Riccati equations, the channel capacities $\mathcal{C}^{(ji)}$'s are related to the unstable eigenvalues of the system. For the single sensor, single controller system, this relationship was shown in [12].

Control laws for cluster $\mathcal{S}_3$: Next, we find the control signals $u_t^{(5)}$ and $u_t^{(6)}$ for subsystems s_5 and s_6, respectively, by solving two separate centralized problems. Here, we first present the design principle for the control signal $u_t^{(5)}$. The control signal $u_t^{(6)}$ for subsystem s_6 is obtained similarly. Recall that for each $t \geq 0$, the information available at subsystem s_5 is $\mathcal{F}_t^{(3,5)} \equiv \sigma\{\bar{U}^{(3)}(t), y^{(5)}(t), \bar{y}^{(56)}(t)\}$. Therefore, the observation process available at subsystem s_5 is $Y_t^{(5)} \equiv \begin{pmatrix} y_t^{(5)} \\ \bar{y}_t^{(56)} \end{pmatrix}$. This is used along with (11), (15) and (9), as displayed by the dynamic model (16), to construct the control signal $u_t^{(5)}$ which is one of the components of the control $U_t^{(3)}$.

$$\begin{cases} X_{t+1}^{(3)} = A^{(3)}X_t^{(3)} + B^{(3)}U_t^{(3)} + C^{(3)}W_t^{(3)} \\ Y_t^{(5)} = F_t^{(5)}X_t^{(3)} + G_t^{(5)}V_t^{(5)} + H_t^{(5)} + \Phi_t^{(5)}, \end{cases} \tag{16}$$

$$F_t^{(5)} = \mathrm{diag}(F_5, \quad \beta_t^{(56)}\gamma_t^{(56)}F_6),$$

$$G_t^{(5)} = \mathrm{diag}(G_5, \quad \beta_t^{(56)}\gamma_t^{(56)}G_6), \qquad V_t^{(5)} = \begin{pmatrix} v_t^{(5)} \\ v_t^{(6)} \end{pmatrix},$$

$$H_t^{(5)} = \begin{pmatrix} 0 \\ \bar{\zeta}_t^{(56)} \end{pmatrix} (\bar{\zeta}_t^{(56)} = (h^{(56)})^{-1}\gamma_t^{(56)}\zeta_t^{(56)}),$$

$$\Phi_t^{(5)} = \begin{pmatrix} 0 \\ (1-\beta_t^{(56)}\gamma_t^{(56)})F_6\hat{x}_t^{(6)} \end{pmatrix} \qquad (\hat{x}_t^{(6)} \equiv E[x_t^{(6)}|\mathcal{F}_{t-1}^{(3,6)}]).$$

The dynamic model (16) has the following equivalent representation:

$$\begin{cases} X_{t+1}^{(3)} = A^{(3)}X_t^{(3)} + B^{(3)}U_t^{(3)} + C^{(3)}W_t^{(3)} \\ \tilde{Y}_t^{(5)} = F_t^{(5)}X_t^{(3)} + G_t^{(5)}V_t^{(5)} + H_t^{(5)}, \end{cases} \tag{17}$$

where $\tilde{Y}_t^{(5)} \equiv Y_t^{(5)} - \Phi_t^{(5)}$. Note that the vector $\Phi_t^{(5)}$ is known to subsystem s_5. Therefore, this subsystem can treat the vector $\tilde{Y}_t^{(5)} = Y_t^{(5)} - \Phi_t^{(5)}$ as the observation signal. Also, note that in the dynamic model (17), the noise terms $W_t^{(3)}$, $V_t^{(5)}$, and $H_t^{(5)}$ are independent and uncorrelated. The noise term $W_t^{(3)} = \begin{pmatrix} w_t^{(5)} \\ w_t^{(6)} \end{pmatrix}$ includes process noises and the noise term $V_t^{(5)} = \begin{pmatrix} v_t^{(5)} \\ v_t^{(6)} \end{pmatrix}$ includes the measurement noises of subsystems s_5 and s_6. Therefore, following the assumptions made in Section 2, $W_t^{(3)}$ and $V_t^{(5)}$ are independent and uncorrelated. Moreover, the noise term $H_t^{(5)}$ represents the effect of channel noise on the dynamic model (17). Again from the assumptions made in Section 2, $H_t^{(5)}$ is independent and uncorrelated of the noise terms $W_t^{(3)}$ and $V_t^{(5)}$. Now, we use the dynamic model (17) to construct the control signal $u_t^{(5)}$. For the dynamic model (17) we follow the LQG methodology [23] subject to linear policies and the payoff functional $J^{(3)}$; and we find the control vector $U_t^{(3)}$, in which the control signal $u_t^{(5)}$ is one of its components. The solution involves two Riccati equations: the control Riccati equation and filtered Riccati equation. The control Riccati equation is given by

$$\Lambda^{(3)} = A^{(3)'}\Lambda^{(3)}A^{(3)} - A^{(3)'}\Lambda^{(3)}B^{(3)}(B^{(3)'}\Lambda^{(3)}B^{(3)} + R_3)^{-1}B^{(3)'}\Lambda^{(3)}A^{(3)} + Q_3. \tag{18}$$

Under the following two assumptions:

(a1) The pair $(A^{(3)}, B^{(3)})$ is stabilizable

(a2) The pair $\left(Q_3^{\frac{1}{2}}, A^{(3)}\right)$ is detectable, the control Riccati equation (18) has a unique positive semi-definite solution $\Lambda^{(3)}$. This solution is used for constructing the controller gain $\Delta^{(3)}$ which is given by:

$$\Delta^{(3)} = (R_3 + B^{(3)'}\Lambda^{(3)}B^{(3)})^{-1}B^{(3)'}\Lambda^{(3)}A^{(3)}. \tag{19}$$

A stabilizing controller $U_t^{(3)}$ is then given by

$$U_t^{(3)} = \begin{pmatrix} u_t^{(5)} \\ u_t^{(56)} \end{pmatrix} = -\Delta^{(3)}\hat{X}_t^{(3,5)}, \tag{20}$$

where $\hat{X}_t^{(3,5)} \equiv E[X_t^{(3)}|\mathcal{F}_{t-1}^{(3,5)}]$ is the mean square state estimate of the state variable $X_t^{(3)}$ at subsystem s_5. This estimate is given by the solution of the Kalman filter equations which consist of the estimator equation and the error covariance equation as presented below: The estimator equation is given by $\hat{X}_{t+1}^{(3,5)} = (A^{(3)} - L_t^{(5)}F_t^{(5)})\hat{X}_t^{(3,5)} + B^{(3)}\begin{pmatrix} u_t^{(5)} \\ u_t^{(6)} \end{pmatrix} + L_t^{(5)}\tilde{Y}_t^{(5)}$, $\hat{X}_0^{(3,5)} = \begin{pmatrix} \bar{x}_0^{(5)} \\ \bar{x}_0^{(6)} \end{pmatrix}$, where the filter gain $L_t^{(5)}$ is given by $L_t^{(5)} = A^{(3)}\Xi_t^{(5)}F_t^{(5)'}\left(F_t^{(5)}\Xi_t^{(5)}.F_t^{(5)'} + G_t^{(5)}\Sigma_V^{(5)}G_t^{(5)'} + \Upsilon_t^{(5)}\right)^{-1}$, $\Sigma_V^{(5)} \equiv \mathrm{Cov}(V_t^{(5)})$, $\Upsilon_t^{(5)} \equiv \mathrm{Cov}(H_t^{(5)})$. This filter involves the error covariance $\Xi_t^{(5)} \equiv \mathrm{Cov}(X_t^{(3)} - \hat{X}_t^{(3,5)}|\mathcal{F}_{t-1}^{(3,5)})$ which satisfies the following (filter) Riccati equation $\Xi_{t+1}^{(5)} = A^{(3)}\Xi_t^{(5)}A^{(3)'} - A^{(3)}\Xi_t^{(5)}F_t^{(5)'}\left(F_t^{(5)}\Xi_t^{(5)}F_t^{(5)'} + F_t^{(5)}\Sigma_V^{(5)}F_t^{(5)'} + \Upsilon_t^{(5)}\right)^{-1}F_t^{(5)}\Xi_t^{(5)}A^{(3)'} + C^{(3)}\Sigma_W^{(3)}C^{(3)'}$, $\Sigma_W^{(3)} \equiv \mathrm{Cov}(W_t^{(3)})$, $\Xi_0^{(5)} = \mathrm{Cov}\begin{pmatrix} \xi_0^{(5)} \\ \xi_0^{(6)} \end{pmatrix}$.

Since we are interested in the stationary control law, we must introduce assumptions that guarantee the existence of the limit, $\lim_{t\to\infty}\Xi_t^{(5)} = \Xi_\infty^{(5)}$. The following conditions are sufficient for the finiteness of $\Xi_t^{(5)}$ and the existence of the limit, $\lim_{t\to\infty}\Xi_t^{(5)} = \Xi_\infty^{(5)}$ [23, Theorem 6.45].

(a3) The pair $(F_t^{(5)}, A^{(3)})$ is uniformly completely reconstructible [23, Definition 1.22].

(a4) The pair $(A^{(3)}, C^{(3)})$ is controllable.

For subsystem s_6 again we use the dynamic model (9). However, the observation process available at subsystem s_6 is $Y_t^{(6)} \equiv \begin{pmatrix} \bar{y}_t^{(65)} \\ y_t^{(6)} \end{pmatrix}$.

Therefore, we use the following dynamic model to construct the control signal $u_t^{(6)}$:

$$\begin{cases} X_{t+1}^{(3)} = A^{(3)}X_t^{(3)} + B^{(3)}U_t^{(3)} + C^{(3)}W_t^{(3)} \\ \tilde{Y}_t^{(6)} = F_t^{(6)}X_t^{(3)} + G_t^{(6)}V_t^{(6)} + H_t^{(6)}, \end{cases} \tag{21}$$

where $F_t^{(6)} = \mathrm{diag}(\beta_t^{(65)}\gamma_t^{(65)}F_5, F_6)$, $G_t^{(6)} = \mathrm{diag}(\beta_t^{(65)}\gamma_t^{(65)}G_5, G_6)$, $V_t^{(6)} = \begin{pmatrix} v_t^{(5)} \\ v_t^{(6)} \end{pmatrix}$, $H_t^{(6)} = \begin{pmatrix} \bar{\zeta}_t^{(65)} \\ 0 \end{pmatrix}$, $\bar{\zeta}_t^{(65)} = (h^{(65)})^{-1}\gamma_t^{(65)}\zeta_t^{(65)}$ and $\tilde{Y}_t^{(6)} \equiv Y_t^{(6)} - \Phi_t^{(6)} \equiv \begin{pmatrix} \bar{y}_t^{(65)} \\ y_t^{(6)} \end{pmatrix} - \begin{pmatrix} (1-\beta_t^{(65)}\gamma_t^{(65)})F_5\hat{x}_t^{(5)} \\ 0 \end{pmatrix}$, $\hat{x}_t^{(5)} \equiv E[x_t^{(5)}|\mathcal{F}_{t-1}^{(3,5)}]$.

For this system again we use the same reconstructability and controllability assumptions except that here assumption (a3) is replaced by assumption (á3):

(á3) The pair $(F_t^{(6)}, A^{(3)})$ is uniformly completely reconstructible.
Then, for the payoff functional $J^{(3)}$ it follows from the standard LQG results [23] that a stabilizing controller $U_t^{(3)}$ is given by

$$U_t^{(3)} = \begin{pmatrix} u_t^{(65)} \\ u_t^{(6)} \end{pmatrix} = -\Delta^{(3)}\hat{X}_t^{(3,6)}, \tag{22}$$

where the controller gain $\Delta^{(3)}$ is the same as in (19) and $\hat{X}_t^{(3,6)} \equiv E[X_t^{(3)}|\mathcal{F}_{t-1}^{(3,6)}]$, is the mean square estimate of the state variable $X_t^{(3)}$ at subsystem s_6 given the information $\mathcal{F}_{t-1}^{(3,6)} \equiv \sigma\{U^{(3)}(t-1), y^{(6)}(t-1), \bar{y}^{(65)}(t-1)\}$. Following the same procedure as in the case of subsystem s_5, the error covariance, the filter gain and finally control, etc. for subsystem s_6 are determined. The modifications required are as follows: $F_t^{(5)} \to F_t^{(6)}$, $\tilde{Y}_t^{(5)} \to \tilde{Y}_t^{(6)}$, $L_t^{(5)} \to L_t^{(6)}$, $G_t^{(5)} \to G_t^{(6)}$, $\Xi_t^{(5)} \to \Xi_t^{(6)}$, $\Upsilon_t^{(5)} \to \Upsilon_t^{(6)} \equiv \mathrm{Cov}(H_t^{(6)})$ and $\Sigma_V^{(5)} \to \Sigma_V^{(6)} \equiv \mathrm{Cov}(V_t^{(6)})$. Note that the error covariance matrix $\Xi_t^{(6)}$ is the solution of the filtered Riccati equation. Under assumptions (á3) and (a4) the limit, $\lim_{t\to\infty}\Xi_t^{(6)} \equiv \Xi_\infty^{(6)}$, exists.

In order to extract the control signals $u_t^{(5)}$ and $u_t^{(6)}$ from (20) and (22) we can partition the control matrix $\Delta^{(3)}$ as follows: $\Delta^{(3)} = \begin{pmatrix} \Delta_5 \\ \Delta_6 \end{pmatrix}$ where the matrix Δ_5 corresponds to subsystem s_5 and the matrix Δ_6 corresponds to subsystem s_6. Therefore, $u_t^{(5)} = -\Delta_5\hat{X}_t^{(3,5)}$ and $u_t^{(6)} = -\Delta_6\hat{X}_t^{(3,6)}$; and hence

$$U_t^{(3)} = \begin{pmatrix} u_t^{(5)} \\ u_t^{(6)} \end{pmatrix} = \begin{pmatrix} -\Delta_5\hat{X}_t^{(3,5)} \\ -\Delta_6\hat{X}_t^{(3,6)} \end{pmatrix}. \tag{23}$$

Justification of decoding function: To justify the construction of the decoding function as given by (11), here we show how the state estimates are made available at subsystems s_5 and s_6. At each $t \geq 0$, subsystem s_5 knows the control signal $u_t^{(6)} = -\Delta_6\hat{X}_t^{(3,6)}$. Therefore, under the assumption that the matrix Δ_6 has full column rank, subsystem s_5 uses this information to compute $\hat{X}_t^{(3,6)} = -(\Delta_6'\Delta_6)^{-1}\Delta_6'u_t^{(6)}$. This vector has the representation $\hat{X}_t^{(3,6)} \equiv \begin{pmatrix} \hat{x}_t^{(65)} \\ \hat{x}_t^{(6)} \end{pmatrix}$ where $\hat{x}_t^{(65)} \in \Re^{n_5}$ denotes the state estimate of subsystem s_5 at subsystem s_6; and $\hat{x}_t^{(6)} \in \Re^{n_6}$ is the state estimate of subsystem s_6 itself. Therefore, $\hat{x}_t^{(6)}$ is known to subsystem s_5. Similarly, one can justify that the state estimate $\hat{x}_t^{(5)}$ is known at subsystem s_6.

Remark 3.3. From (15) it follows that $\beta_t^{(56)}\gamma_t^{(56)} = \eta_t^{(56)} = 1 - \frac{\delta^{(56)}}{\Psi_t^{(6)}}$ and $\beta_t^{(65)}\gamma_t^{(65)} = \eta_t^{(65)} = 1 - \frac{\delta^{(65)}}{\Psi_t^{(5)}}$. Therefore, when the capacities of communication channels are infinity (i.e., $\delta^{(56)} \approx 0$ and $\delta^{(65)} \approx 0$), $\beta_t^{(56)}\gamma_t^{(56)} = \beta_t^{(65)}\gamma_t^{(65)} = 1$ and $\gamma_t^{(56)} = \gamma_t^{(65)} = 0$. Consequently, from (10) and (11) it follows that $\bar{y}_t^{(56)} = y_t^{(6)}$, $\bar{y}_t^{(65)} = y_t^{(5)}$ and the dynamic models (17) and (21) are reduced to two identical partially observed time invariant Gaussian systems. As a result, two control vectors (20) and (22) are identical and are the optimal controller for systems (17) and (21) with $\delta^{(56)} \approx 0$ and $\delta^{(65)} \approx 0$. Consequently, for this case, the controller, as specified by (23), is the optimal controller. Note that the cost functional $J^{(3)}$ is a continuous function of the distortion levels $\delta^{(56)}$ and $\delta^{(65)}$. Therefore, when the communication channels are of high capacities (i.e., the distortion levels are not zero; but they are small) we would expect that the controller, as specified by (23), to be close to the optimal controller. That is, in the presence of high capacity communication constraints, this controller is only a suboptimal solution for the payoff functional $J^{(3)}$.

Design methodology for cluster $\mathcal{S}_2$: Now, consider cluster $\mathcal{S}_2$, as described by (1), which is reproduced here for convenience of reference.

$$X_{t+1}^{(2)} = A^{(2)}X_t^{(2)} + B^{(2)}U_t^{(2)} + C^{(2)}W_t^{(2)} + D^{(3)}X_t^{(3)}. \tag{24}$$

For each $t \geq 0$, the information available at subsystem s_i ($i \in \{3, 4\}$) of cluster $\mathcal{S}_2$ is $\mathcal{F}_t^{(2,i)} = \sigma\{U^{(3)}(t), U^{(2)}(t), y^{(i)}(t), \bar{y}^{(ij)}(t)\}$, $j(\neq i) \in \{4, 3\}$. Therefore, subsystem s_i can use the control signal $u_t^{(5)} = -\Delta_5\hat{X}_t^{(3,5)}$ to calculate $\hat{X}_t^{(3,5)} = -(\Delta_5'\Delta_5)^{-1}\Delta_5'u_t^{(5)}$. Unlike cluster $\mathcal{S}_3$, cluster $\mathcal{S}_2$ is subject to interaction coming from cluster $\mathcal{S}_3$. Therefore, the controllers of this cluster first compensate the interaction coming from cluster $\mathcal{S}_3$ by using the following control $U_t^{(2)}$:

$$U_t^{(2)} = \tilde{U}_t^{(2)} - B^{(2)'}(B^{(2)}B^{(2)'})^{-1}(D^{(3)}\hat{X}_t^{(3,5)}), \tag{25}$$

where $\tilde{U}_t^{(2)} \equiv \begin{pmatrix} \tilde{u}_t^{(3)} \\ \tilde{u}_t^{(4)} \end{pmatrix}$ and the remaining term compensates the effect of interacting cluster $\mathcal{S}_3$.

By substituting the control signal (25) in the dynamic model (24), one obtains the compensated dynamics of cluster $\mathcal{S}_2$ as follows:

$$X_{t+1}^{(2)} = A^{(2)}X_t^{(2)} + B^{(2)}\tilde{U}_t^{(2)} + C^{(2)}W_t^{(2)} + D^{(3)}E_t^{(5)}, \tag{26}$$

where $E_t^{(5)} \equiv X_t^{(3)} - \hat{X}_t^{(3,5)}$ is the estimation error of the state variable $X_t^{(3)}$ at subsystem s_5. It is an orthogonal Gaussian sequence with distribution $E_t^{(5)} \sim N(0, \Xi_t^{(5)})$. Note that the estimation error $E_t^{(5)}$ is independent of the process noise $W_t^{(2)}$. For this cluster we use similar encoding and decoding techniques, as used for cluster $\mathcal{S}_3$. Controllers $\tilde{u}_t^{(3)}$ and $\tilde{u}_t^{(4)}$ are also obtained following a similar procedure by the use of the LQG methodology [23]; except that here we use the compensated dynamic (26) and the payoff functional $\tilde{J}^{(2)}$, which is the cost functional $J^{(2)}$ with the control vector $U_t^{(2)}$ replaced by $\tilde{U}_t^{(2)}$. Controllers $\tilde{u}_t^{(3)}$ and $\tilde{u}_t^{(4)}$ have similar forms as the controllers $u_t^{(5)}$ and $u_t^{(6)}$. That is,

$$\tilde{U}_t^{(2)} = \begin{pmatrix} \tilde{u}_t^{(3)} \\ \tilde{u}_t^{(4)} \end{pmatrix} = \begin{pmatrix} -\Delta_3\hat{X}_t^{(2,3)} \\ -\Delta_4\hat{X}_t^{(2,4)} \end{pmatrix}, \tag{27}$$

where Δ_3 and Δ_4 are the controller gains and $\hat{X}_t^{(2,3)} \equiv E[X_t^{(2)}|\mathcal{F}_{t-1}^{(2,3)}]$ and $\hat{X}_t^{(2,4)} \equiv E[X_t^{(2)}|\mathcal{F}_{t-1}^{(2,4)}]$ are the mean square state estimates of the state variable $X_t^{(2)}$ at subsystems s_3 and s_4, respectively.

Design methodology for cluster $\mathcal{S}_1$: For cluster $\mathcal{S}_1$ we also use similar encoding and decoding techniques. This cluster includes subsystem s_i, $i \in \{1, 2\}$. The information pattern of subsystem s_i for each $t \geq 0$ is $\mathcal{F}_t^{(1,i)} = \sigma\{U^{(2)}(t), \tilde{U}^{(2)}(t), U^{(1)}(t), y^{(i)}(t), \bar{y}^{(ij)}(t)\}$ ($j(\neq i) \in \{2, 1\}$). Controllers for this cluster are obtained following a similar procedure, as described above, by use of the cost functional $\tilde{J}^{(1)}$, which is the payoff functional $J^{(1)}$ with $U_t^{(1)}$ replaced by $\tilde{U}_t^{(1)}(= U_t^{(1)} + B^{(1)'}(B^{(1)}B^{(1)'})^{-1}(D^{(2)}\hat{X}_t^{(2,3)}))$. The controllers of this cluster are given by:

$$U_t^{(1)} = \begin{pmatrix} u_t^{(1)} \\ u_t^{(2)} \end{pmatrix} = \tilde{U}_t^{(1)} - B^{(1)'}(B^{(1)}B^{(1)'})^{-1}(D^{(2)}\hat{X}_t^{(2,3)}),$$

$$\tilde{U}_t^{(1)} = \begin{pmatrix} \tilde{u}_t^{(1)} \\ \tilde{u}_t^{(2)} \end{pmatrix} = \begin{pmatrix} -\Delta_1 \hat{X}_t^{(1,1)} \\ -\Delta_2 \hat{X}_t^{(1,2)} \end{pmatrix}, \quad (28)$$

where Δ_1 and Δ_2 are the controller gains and $\hat{X}_t^{(1,1)} \equiv E[X_t^{(1)}|\mathcal{F}_{t-1}^{(1,1)}]$ and $\hat{X}_t^{(1,2)} \equiv E[X_t^{(1)}|\mathcal{F}_{t-1}^{(1,2)}]$ are the mean square state estimates of the state variable $X_t^{(1)}$ at subsystems s_1 and s_2, respectively.

When the capacities of communication channels are infinity and the coupling matrices $D^{(2)}$ and $D^{(1)}$ are equal to zero, the controllers as specified by (23), (25), (27) and (28) are the optimal solution of the cost functional (8) subject to the dynamic model (2)–(4) with three decoupled clusters. This implies that when this system is weakly cascaded and communication channels are of high capacities, we would expect that the controls, as specified above, to be close to the optimal controller.

Now, we are prepared to consider the main objective of this paper which is the question of stability of the large scale system.

3.2. Stability analysis

It is known that a finite solution to quadratic cost functional with positive definite weighting matrices results in mean square stability of Gaussian systems. Hence, the proposed linear policy, which is a suboptimal solution for quadratic payoff functional, results in mean square stability. In this section we show this result. This is done by proving the mean square stability of each of clusters $\mathcal{S}_3$, $\mathcal{S}_2$ and $\mathcal{S}_1$ in that order and then justifying that this implies the stability of the whole system.

Proposition 3.4 (*Stability of Cluster* $\mathcal{S}_3$). *Consider cluster* $\mathcal{S}_3$ *with all the assumptions following its description. Further, suppose assumptions* (a1)–(a4) *and* (á3) *hold. Then, cluster* $\mathcal{S}_3$ *with the controls as specified by* (23) *is mean square stable.*

Proof. Denote the state estimation error at subsystem s_6 by $E_t^{(6)} \equiv X_t^{(3)} - \hat{X}_t^{(3,6)}$ and the mismatch of the state estimates of cluster $\mathcal{S}_3$ by $E_t^{(5,6)} \equiv \hat{X}_t^{(3,5)} - \hat{X}_t^{(3,6)}$ (as computed at subsystems s_5 and s_6). In order to deduce stability, we introduce an augmented system with corresponding state denoted by $Z_t^{(3)} \equiv \begin{pmatrix} X_t^{(3)} \\ E_t^{(6)} \\ E_t^{(5,6)} \end{pmatrix}$. This is given by

$$Z_{t+1}^{(3)} = S_t^{(3)} Z_t^{(3)} + R_t^{(3)}, \quad (29)$$

where

$$S_t^{(3)} = \begin{pmatrix} A^{(3)} - B^{(3)}\Delta^{(3)} & B^{(3)}\Delta^{(3)} & -B^{(3,5)}\Delta_5 \\ 0 & A^{(3)} - L_t^{(6)}F_t^{(6)} & 0 \\ 0 & L_t^{(5)}F_t^{(5)} - L_t^{(6)}F_t^{(6)} & A^{(3)} - L_t^{(5)}F_t^{(5)} \end{pmatrix},$$

and

$$R_t^{(3)} = \begin{pmatrix} C^{(3)}W_t^{(3)} \\ C^{(3)}W_t^{(3)} - L_t^{(6)}G_t^{(6)}V_t^{(6)} - L_t^{(6)}H_t^{(6)} \\ (L_t^{(5)}G_t^{(5)} - L_t^{(6)}G_t^{(6)})V_t^{(6)} + L_t^{(5)}H_t^{(5)} - L_t^{(6)}H_t^{(6)} \end{pmatrix}.$$

All the entries of the matrices $S_t^{(3)}$ and $R_t^{(3)}$ are described in Sections 2 and 3.1 except the matrix $B^{(3,5)}$ which appears in the 1st row and the 3rd column of the matrix $S_t^{(3)}$. This term is part of the control matrix $B^{(3)}$ of the dynamic model (9) and is given by $B^{(3)} = \begin{pmatrix} B^{(3,5)} & B^{(3,6)} \end{pmatrix}$ where $B^{(3,5)}$ corresponds to the control signal $u_t^{(5)}$ and $B^{(3,6)}$ corresponds to the control signal $u_t^{(6)}$, giving $B^{(3)}U_t^{(3)} = B^{(3,5)}u_t^{(5)} + B^{(3,6)}u_t^{(6)}$. It is evident that the stability of the augmented system (29) is equivalent to the stability of cluster $\mathcal{S}_3$. Hence, it suffices to show that the dynamic model described by (29) is stable in the mean square sense. From the reconstructability and controllability assumptions, (a3)–(á3) and (a4), respectively, we have, $\lim_{t\to\infty} S_t^{(3)} \equiv S^{(3)}$, where in the matrix $S^{(3)}$ we have $L^{(5)} \equiv \lim_{t\to\infty} L_t^{(5)}$, $L^{(6)} \equiv \lim_{t\to\infty} L_t^{(6)}$, $F^{(5)} \equiv \lim_{t\to\infty} F_t^{(5)}$ and $F^{(6)} \equiv \lim_{t\to\infty} F_t^{(6)}$. It is easy to verify that the stability of the matrix $S^{(3)}$ is equivalent to the stability of its diagonal elements. From assumption (a1) it follows that the matrix $A^{(3)} - B^{(3)}\Delta^{(3)}$ is stable. Similarly from assumptions (a3) and (á3) it follows that the matrices $A^{(3)} - L^{(5)}F^{(5)}$ and $A^{(3)} - L^{(6)}F^{(6)}$ are stable. Therefore, the matrix $S^{(3)}$ is stable. This implies that, for sufficiently large enough t, the system matrix $S_t^{(3)}$ is stable. On the other hand, for each $t \geq 0$, the random vector $R_t^{(3)}$ has a finite second moment. Hence, the dynamic model (29) is stable in mean square sense. □

Stability of clusters $\mathcal{S}_2$ *and* $\mathcal{S}_1$: The stability of the remaining clusters $\mathcal{S}_2$ and $\mathcal{S}_1$ is established using a similar procedure. For cluster $\mathcal{S}_2$, again let

$$Z_t^{(2)} \equiv \begin{pmatrix} X_t^{(2)} \\ E_t^{(4)} \\ E_t^{(3,4)} \end{pmatrix}, \qquad E_t^{(4)} \equiv X_t^{(2)} - \hat{X}_t^{(2,4)},$$

$$E_t^{(3,4)} \equiv \hat{X}_t^{(2,3)} - \hat{X}_t^{(2,4)}.$$

Similarly for cluster $\mathcal{S}_1$ let

$$Z_t^{(1)} \equiv \begin{pmatrix} X_t^{(1)} \\ E_t^{(2)} \\ E_t^{(1,2)} \end{pmatrix}, \qquad E_t^{(2)} \equiv X_t^{(1)} - \hat{X}_t^{(1,2)},$$

$$E_t^{(1,2)} \equiv \hat{X}_t^{(1,1)} - \hat{X}_t^{(1,2)}.$$

Then, for these two clusters we have $Z_{t+1}^{(2)} = S_t^{(2)}Z_t^{(2)} + R_t^{(2)}$, and $Z_{t+1}^{(1)} = S_t^{(1)}Z_t^{(1)} + R_t^{(1)}$, where under similar controllability and reconstructability assumptions one can show that the matrices $S_t^{(2)}$ and $S_t^{(1)}$ are stable for sufficiently large enough t. Therefore, it follows from similar arguments that these two clusters are stable.

Stability of the large scale system: Now, we are ready to prove the stability of the large scale system.

Theorem 3.5. *Consider the large scale system described by* (2)–(4) *consisting of three clusters* $\mathcal{S}_1$, $\mathcal{S}_2$ *and* $\mathcal{S}_3$, *as described earlier. Suppose that the controllability and reconstructability assumptions hold. Then the large scale system with controls as specified by* (23), (25), (27) *and* (28) *is stable in the mean square sense.*

Proof. When these controllers are used the entire (large scale) system appears as follows:

$$\begin{pmatrix} Z_{t+1}^{(1)} \\ Z_{t+1}^{(2)} \\ Z_{t+1}^{(3)} \end{pmatrix} = S_t \begin{pmatrix} Z_t^{(1)} \\ Z_t^{(2)} \\ Z_t^{(3)} \end{pmatrix} + R_t, \quad (30)$$

where

$$S_t = \mathrm{diag}(S_t^{(1)}, S_t^{(2)}, S_t^{(3)}), \qquad R_t = \begin{pmatrix} R_t^{(1)} \\ R_t^{(2)} \\ R_t^{(3)} \end{pmatrix}.$$

As shown above, for sufficiently large enough t, the matrices $S_t^{(1)}$, $S_t^{(2)}$, and $S_t^{(3)}$ are stable; and hence the diagonal matrix S_t is stable. On the other hand, for each $t \geq 0$, the random vector R_t has finite second moment. Hence, the dynamic model (30) is stable in mean square sense. □

When the controls, as specified in Theorem 3.5, are used, from the reconstructability and controllability assumptions (a1)–(a4) and (á3) it follows that the limit, $\lim_{t\to\infty} E[\|Z_t^{(r)}\|_H^2]$, $(r = 1, 2, 3)$, exists for any positive semi-definite matrix H. For $H = \mathrm{diag}(Q_r, 0, 0)$, this implies that the limit, $\lim_{t\to\infty} E[\|X_t^{(r)}\|_{Q_r}^2]$, exists. The above conditions also imply that for the stabilizing controllers specified in Theorem 3.5, the limit, $\lim_{t\to\infty} E[\|U_t^{(r)}\|_{R_r}^2]$, exists. Therefore, the limit, $\lim_{t\to\infty} E[\|X_t^{(r)}\|_{Q_r}^2 + \|U_t^{(r)}\|_{R_r}^2]$, exists and hence the limit in (8) exists.

Now, combining the results of Theorem 3.5 for mean square stability and the results of Section 3.1 (e.g., Proposition 3.2), which are concerned with the rate requirements for reliable communication, we have the following sufficient condition on the capacity of channels for mean square stability.

Theorem 3.6. *Consider the large scale system described by (2)–(4) consisting of three clusters $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$, as described earlier, over AWGN channels (6), with all the assumptions have been made so far for stability and reliable communication. Then, a sufficient condition on the capacity of channels for mean square stability is as follows:*

$$\sum_{r=1}^{l} \sum_{i,j(\neq i)\in\mathcal{S}_r} \mathcal{C}^{(ji)} \geq \mathcal{C}_d \equiv \sum_{r=1}^{l} \sum_{i,j(\neq i)\in\mathcal{S}_r} \frac{1}{2}\log\frac{\Psi_\infty^{(i)}}{\delta^{(ji)}}, \tag{31}$$

where $\mathcal{C}^{(ji)}$ is the capacity of the channel used for transmission from subsystem i to subsystem j, $\delta^{(ji)}$ is the distortion level and $\Psi_\infty^{(i)} = \lim_{t\to\infty}\Psi_t^{(i)}$.

Proof. From condition (31) it follows that it is possible to choose the capacity $\mathcal{C}^{(ji)}$ such that $\mathcal{C}^{(ji)} \geq \frac{1}{2}\log\frac{\Psi_\infty^{(i)}}{\delta^{(ji)}}$, where $i, j(\neq i) \in \mathcal{S}_r$ and $r \in \{1, 2, \ldots, l\}$. This condition guarantees a reliable communication of the form $\rho_{ji} = \hat{\delta}^{(ji)} \leq \delta^{(ji)}$ where $\hat{\delta}^{(ji)}$ is such that $\mathcal{C}^{(ji)} = \frac{1}{2}\log\frac{\Psi_\infty^{(i)}}{\hat{\delta}^{(ji)}}$. Subsequently, by employing the controls as specified by (23), (25), (27) and (28), with $\delta^{(ji)}$ replaced by $\hat{\delta}^{(ji)}$, mean square stability is obtained. □

4. Comparison

In this section, we compare the rate requirement for stability using the decentralized technique of Section 3.1 with the minimum rate requirement for stability using the centralized technique of [12]. For simplicity of comparison, we consider the large scale system of Section 3 with only two clusters $\mathcal{S}_3$ and $\mathcal{S}_2$ ($\mathcal{S}_1$ is omitted). That is, we consider the following system:

$$\begin{cases} X_{t+1} = AX_t + BU_t + CW_t, \\ y_t^{(i)} = F_i x_t^{(i)} + G_i v_t^{(i)}, \quad i = 3, 4, 5, 6, \end{cases} \tag{32}$$

where $X_t = \begin{pmatrix} x_t^{(3)} \\ x_t^{(4)} \\ x_t^{(5)} \\ x_t^{(6)} \end{pmatrix}$, $X_0 \sim N(\bar{X}, \bar{V})$, $U_t = \begin{pmatrix} u_t^{(3)} \\ u_t^{(4)} \\ u_t^{(5)} \\ u_t^{(6)} \end{pmatrix}$, $W_t = \begin{pmatrix} w_t^{(3)} \\ w_t^{(4)} \\ w_t^{(5)} \\ w_t^{(6)} \end{pmatrix}$, $A = \begin{pmatrix} A^{(2)} & D^{(3)} \\ 0 & A^{(3)} \end{pmatrix}$, $B = \begin{pmatrix} B^{(2)} & 0 \\ 0 & B^{(3)} \end{pmatrix}$, $C = \begin{pmatrix} C^{(2)} & 0 \\ 0 & C^{(3)} \end{pmatrix}$, $A^{(2)} = \begin{pmatrix} A_3 & D_{34} \\ D_{43} & A_4 \end{pmatrix}$, $D^{(3)} = \begin{pmatrix} D_{35} & D_{36} \\ D_{45} & D_{46} \end{pmatrix}$, $A^{(3)} = \begin{pmatrix} A_5 & D_{56} \\ D_{65} & A_6 \end{pmatrix}$, $B^{(2)} = \begin{pmatrix} B_3 & E_{34} \\ E_{43} & B_4 \end{pmatrix}$, $B^{(3)} = \begin{pmatrix} B_5 & E_{56} \\ E_{65} & B_6 \end{pmatrix}$, $C^{(2)} = \begin{pmatrix} C_3 & 0 \\ 0 & C_4 \end{pmatrix}$, $C^{(3)} = \begin{pmatrix} C_5 & 0 \\ 0 & C_6 \end{pmatrix}$, $y_t^{(i)} \in \Re$, $w_t^{(i)}$ i.i.d. $\sim N(0, \Sigma_w^{(i)})$ and $v_t^{(i)}$ i.i.d. $\sim N(0, \Sigma_v^{(i)})$. Also, for simplicity, without loss of generality, we assume that the channel gains are equal to one, that is, $h^{(34)} = h^{(43)} = h^{(56)} = h^{(65)} = 1$.

From Theorem 3.6 it follows that using the decentralized technique of Section 3.1, the rate requirement for stability and reliable communication up to the distortion levels $\delta^{(43)}$, $\delta^{(34)}$, $\delta^{(65)}$, and $\delta^{(56)}$ is $\mathcal{C}_d = \frac{1}{2}\log\frac{\Psi_\infty^{(3)}}{\delta^{(43)}} + \frac{1}{2}\log\frac{\Psi_\infty^{(4)}}{\delta^{(34)}} + \frac{1}{2}\log\frac{\Psi_\infty^{(5)}}{\delta^{(65)}} + \frac{1}{2}\log\frac{\Psi_\infty^{(6)}}{\delta^{(56)}}$.

Now, consider the centralized technique of [12] for the system described by (32). Denote by $\mathcal{E}_t(\cdot)$ the centralized encoding law and let $Y_t = \begin{pmatrix} y_t^{(3)} & y_t^{(4)} & y_t^{(5)} & y_t^{(6)} \end{pmatrix}' \in \Re^4$ be the observation vector and U_t be the control signal produced by the centralized controller. Also, denote the channel input by T_t and the output by R_t. Using the centralized technique of [12] the observation vector Y_t is transmitted via the following AWGN channel

$$R_t = T_t + \zeta_t \in \Re^4,$$

$$T_t = \begin{pmatrix} T_t^{(3)} \\ T_t^{(4)} \\ T_t^{(5)} \\ T_t^{(6)} \end{pmatrix} = \mathcal{E}_t(Y(t), U(t-1), R(t-1)) \in \Re^4,$$

$$\zeta_t \text{ i.i.d.} \sim N(0, \mathrm{diag}(\Gamma^{(43)}, \Gamma^{(34)}, \Gamma^{(65)}, \Gamma^{(56)})),$$

$$E[(T_t^{(i)})^2] \leq P_t^{(i)}, \quad i = 3, 4, 5, 6, \tag{33}$$

to the centralized controller. As described in [12], the centralized encoder first produces the Gaussian innovation sequence K_t. Then, it orthogonalizes K_t by applying the matrix E_t' where E_t is the unitary matrix that diagonalizes the symmetric matrix $\Psi_t = \mathrm{Cov}(K_t)$ (i.e., $\Sigma_t = E_t'\Psi_t E_t$ where $\Sigma_t = \mathrm{diag}(\lambda_t^{(3)}, \lambda_t^{(4)}, \lambda_t^{(5)}, \lambda_t^{(6)})$). Finally, the encoder applies the diagonal encoding gain $\mathcal{A}_t \in M(4 \times 4)$ to the vector $E_t'K_t$ and produces $T_t = \mathcal{A}_t E_t' K_t$. The encoding gain $\mathcal{A}_t$ is defined such that the source-channel matching principle [24] holds for the orthogonal Gaussian source message $E_t'K_t = \begin{pmatrix} g_t^{(3)} & g_t^{(4)} & g_t^{(5)} & g_t^{(6)} \end{pmatrix}'$ and the Gaussian channel (33), giving a reliable communication up to the distortion levels $\delta^{(43)}, \delta^{(34)}, \delta^{(65)}$, and $\delta^{(56)}$ for messages $g_t^{(3)}, g_t^{(4)}, g_t^{(5)}$, and $g_t^{(6)}$, respectively. Note that the symmetric matrix Ψ_t is given by $\Psi_t = F\Pi_t F' + G\Sigma_V G'$, where $F = \mathrm{diag}(F_3, F_4, F_5, F_6)$, $G = \mathrm{diag}(G_3, G_4, G_5, G_6)$, $\Sigma_V = \mathrm{diag}(\Sigma_v^{(3)}, \Sigma_v^{(4)}, \Sigma_v^{(5)}, \Sigma_v^{(6)})$, and Π_t is the solution of the following Riccati equation: $\Pi_{t+1} = A\Pi_t A' - A\Pi_t F'\left(F\Pi_t F' + G\Sigma_V G' + (E_t\mathcal{A}_t^{-1})W_c(E_t\mathcal{A}_t^{-1})'\right)^{-1} F\Pi_t A' + C\Sigma_W C'$, $\Pi_0 = \mathrm{Cov}(X_0)$, $W_c = \mathrm{diag}(\Gamma^{(43)}, \Gamma^{(34)}, \Gamma^{(65)}, \Gamma^{(56)})$, $\Sigma_W = \mathrm{diag}(\Sigma_w^{(3)}, \Sigma_w^{(4)}, \Sigma_w^{(5)}, \Sigma_w^{(6)})$. Following source-channel matching principles, $\mathcal{A}_t$ is given by: $\mathcal{A}_t = \mathrm{diag}\left(\sqrt{\frac{\eta_t^{(43)}\Gamma^{(43)}}{\delta^{(43)}}}, \sqrt{\frac{\eta_t^{(34)}\Gamma^{(34)}}{\delta^{(34)}}}, \sqrt{\frac{\eta_t^{(65)}\Gamma^{(65)}}{\delta^{(65)}}}, \sqrt{\frac{\eta_t^{(56)}\Gamma^{(56)}}{\delta^{(56)}}}\right)$, where $\eta_t^{(43)} = 1 - \frac{\delta^{(43)}}{\lambda_t^{(3)}}$, $\eta_t^{(34)} = 1 - \frac{\delta^{(34)}}{\lambda_t^{(4)}}$, $\eta_t^{(65)} = 1 - \frac{\delta^{(65)}}{\lambda_t^{(5)}}$, and $\eta_t^{(56)} = 1 - \frac{\delta^{(56)}}{\lambda_t^{(6)}}$. Note that the power constraints $P_t^{(i)}$'s are specific to the encoder and are given by $P_t^{(3)} = \frac{\eta_t^{(43)}\Gamma^{(43)}}{\delta^{(43)}}\lambda_t^{(3)}$, $P_t^{(4)} = \frac{\eta_t^{(34)}\Gamma^{(34)}}{\delta^{(34)}}\lambda_t^{(4)}$, $P_t^{(5)} = \frac{\eta_t^{(65)}\Gamma^{(65)}}{\delta^{(65)}}\lambda_t^{(5)}$, and $P_t^{(6)} = \frac{\eta_t^{(56)}\Gamma^{(56)}}{\delta^{(56)}}\lambda_t^{(6)}$.

At the receiver, the centralized decoder first multiples the received signal R_t by $E_t\mathcal{B}_t$ and produces $\tilde{K}_t = E_t\mathcal{B}_t R_t$, where $\tilde{K}_t$ is the reconstructed version of the innovation process K_t and $\mathcal{B}_t$ is the decoding gain. It is obtained from the source-channel matching principle and is given by $\mathcal{B}_t = \mathrm{diag}$

$\left(\sqrt{\frac{\eta_t^{(43)}\delta^{(43)}}{\Gamma^{(43)}}}, \sqrt{\frac{\eta_t^{(34)}\delta^{(34)}}{\Gamma^{(34)}}}, \sqrt{\frac{\eta_t^{(65)}\delta^{(65)}}{\Gamma^{(65)}}}, \sqrt{\frac{\eta_t^{(56)}\delta^{(56)}}{\Gamma^{(56)}}}\right)$. The decoder uses the reconstructed process $\tilde{K}_t$ and by use of the Kalman filter produces $\hat{X}_t$ which is the state estimate of the state of (32). This estimation is then used in the construction of the optimal centralized controller given by $U_t = -\Delta\hat{X}_t$, where Δ is the controller gain, as specified in [12].

Let $\lambda^{(i)} = \lim_{t\to\infty} \lambda_t^{(i)}$ $(i = 3, 4, 5, 6)$. From the analysis of [12] it follows that under some detectability and stabilizability assumptions, the minimum rate requirement for stability and reliable communication using the above centralized technique is, $\mathcal{C}_c = \frac{1}{2}\log\frac{\lambda^{(3)}}{\delta^{(43)}} + \frac{1}{2}\log\frac{\lambda^{(4)}}{\delta^{(34)}} + \frac{1}{2}\log\frac{\lambda^{(5)}}{\delta^{(65)}} + \frac{1}{2}\log\frac{\lambda^{(6)}}{\delta^{(56)}}$.

For illustration we now present the following numerical results.

Numerical results: For illustration let us pick the values as $A^{(2)} = \begin{pmatrix} 1.2 & 0.5 \\ 0.5 & -1.2 \end{pmatrix}$, $A^{(3)} = \begin{pmatrix} 2 & 0.1 \\ 1 & -3 \end{pmatrix}$, $W_c = C = \Sigma_W = \Sigma_V = I_4$, $\delta^{(34)} = \delta^{(43)} = \delta^{(56)} = \delta^{(65)} = 0.12$ and $X_0 \sim N(0, I_4)$. We compare the rate requirements for two cases: (i) Strongly cascaded case and (ii) Weakly cascaded case.

For the first case (strongly cascaded case) let us pick the coupling matrix as $D^{(3)} = \begin{pmatrix} 16.9 & 0 \\ 0 & 16.9 \end{pmatrix}$ where its smallest singular value is 16.9. It is thirteen times larger than the largest singular value of the system matrix $A^{(2)}$ which is 1.3. For this case it can be verified that the rate requirement for stability using the decentralized technique is $\mathcal{C}_d = 35.41$ (bits/time step). It is almost two times bigger than the minimum rate $\mathcal{C}_c = 14.55$ (bits/time step) required for stability using the centralized version. For the second case (weakly cascaded case) let us pick the coupling matrix as $D^{(3)} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$ where its largest singular value is 0.1. It is thirteen times smaller than the smallest singular value of the system matrix $A^{(2)}$ which is 1.3. For this case it can be verified that the rate requirement for stability using the decentralized technique is $\mathcal{C}_d = 10.91$ (bits/time step). This rate is slightly bigger than the minimum rate $\mathcal{C}_c = 10.83$ (bits/time step) required for stability using the centralized version. This suggests that for the weakly cascaded systems, the rate requirement for stability using the proposed decentralized technique is slightly bigger than the minimum rate required for stability using the centralized version.

Note that by increasing the number of clusters and/or the strength of couplings, things get worse in the sense that the rate requirement for stability is going to be bigger than the minimum required rate.

5. Conclusion

In this paper we considered large scale systems with cascaded clusters of linear subsystems, in which the observation signals are exchanged between subsystems via AWGN channels. It was shown that the linear policies stabilize these systems. That was shown by finding a suboptimal solution for quadratic payoff functional which results in mean square stability. A sufficient condition on the capacity of channels for mean square stability was presented. Using an example it was shown that, for the weakly cascaded systems, the rate requirement for stability using the proposed scheme is slightly bigger than the minimum required rate for stability using the centralized version. For the future we would like to extend the results to more complicated systems.

Acknowledgements

A. Farhadi is a Post Doctoral Fellow at INRIA (the French national institute for research in computer science and control), Grenoble, France. This paper is a result of a research work done in Canada. N.U. Ahmed is a Professor at SITE, the University of Ottawa, Canada. This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant no. A7109.

References

[1] M. Gad-el-Hak, The MEMS Handbook: MEMS Applications, CRC Press, 2006.
[2] B.Z. Kurtaran, R. Sivan, Linear-quadratic Gaussian control with one-step-delay sharing pattern, IEEE Transactions on Automatic Control (1974) 571–574.
[3] Y.C. Ho, K.C. Chu, Team decision theory and information structures in optimal control problems-part I, IEEE Transactions on Automatic Control 17 (1) (1972) 15–22.
[4] N.R. Sandell, M. Athans, Solution of some nonclassical LQG stochastic decision problem, IEEE Transactions on Automatic Control 19 (2) (1974) 108–116.
[5] K.C. Chu, Team decision theory and information structures in optimal control problems-part II, IEEE Transactions on Automatic Control 17 (1) (1972) 22–28.
[6] S. Tatikonda, Some scaling properties of large distributed control systems, in: Proceedings of 42nd IEEE Conference on Decision and Control, Hawaii, 2003, pp. 3142–3147.
[7] A.S. Matveev, A.V. Savkin, Stabilization of multi sensors networked control systems with communication constraints, in: 5th Asian Control Conference, 2004, pp. 1905–1913.
[8] G.N. Nair, R.J. Evans, P.E. Caines, Stabilising decentralised linear systems under data rate constraints, in: Proceedings of 43rd IEEE Conference on Decision and Control, Atlandis, Paradise Island, Bahamas, 2004, pp. 3992–3997.
[9] K. Li, J. Baillieul, Problems in decentralized sensor-actuator networks, in: Proceedings of the 44th IEEE Conference on Decision and Control, Seville, 2005, pp. 3207–3212.
[10] A.S. Matveev, A.V. Savkin, Decentralized stabilization of linear system via limited capacity communication networks, in: Proceedings of the 44th IEEE Conference on Decision and Control, Seville, 2005, pp. 1155–1161.
[11] S. Yuksel, T. Basar, Information theoretic study of the signaling problem in decentralized stabilization of noisy linear systems, in: Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, 2007, pp. 328–333.
[12] C.D. Charalambous, A. Farhadi, LQG optimality and separation principle for general discrete time partially observed stochastic systems over finite capacity communication channels, Automatica 44 (12) (2008) 3181–3188.
[13] Y. Zhang, S. Li, Optimization design for a class of multi-input nonlinear cascade systems: backstepping approach, in: 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China, 2004, pp. 1723–1728.
[14] A. Chaillet, D. Angeli, Integral input to state stable systems in cascade, Systems and Control Letters 57 (2008) 519–527.
[15] P. Seibert, R. Suarez, Global stabilization of nonlinear cascade systems, Systems and Control Letters 14 (4) (1990) 347–352.
[16] H.S. Witsenhausen, A counterexample in stochastic optimum control, SIAM Journal on Control 6 (1) (1968) 131–147.
[17] R. Bansal, T. Basar, Stochastic teams with nonclassical information revisited: when is an affine law optimal? IEEE Transactions on Automatic Control 32 (6) (1987) 554–559.
[18] R. Bansal, T. Basar, Solutions to a class of linear – quadratic – Gaussian, LQG, stochastic team problems with nonclassical information, Systems and Control Letters 9 (1987) 125–130.
[19] S. Yuksel, S. Tatikonda, A counterexample in distributed optimal sensing and control, IEEE Transactions on Automatic Control 54 (4) (2009) 841–844.
[20] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.
[21] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley and Sons, 1991.
[22] T. Berger, Rate Distortion Theory: A Mathematical Basis for Data Compression, Prentice-Hall, 1971.
[23] H. Kwakernnak, R. Sivan, Linear Optimal Control Systems, John Wiley and Sons, 1972.
[24] M. Gastpar, B. Rimoldi, M. Vetterli, To code, or not to code: lossy source-channel communication revisited, IEEE Transactions on Information Theory 49 (5) (2003) 1147–1158.