Decentralized Suboptimal Control via Limited Capacity Channels

Alireza Farhadi and N.U. Ahmed

***Abstract*— In this paper decentralized controllers are designed for mean square stability of large scale systems with linear time-invariant distributed subsystems. The subsystems are subject to Gaussian process and measurement noise. For the stability analysis of the system we also consider the effects of noisy and limited capacity communication channels used for exchanging information between subsystems. Hence, the proposed scheme is suitable for controlling networks of Micro Electro Mechanical Systems (MEMS).**

I. Introduction

Development in electronics has given birth to small size embedded systems such as Micro Electro Mechanical Systems (MEMS). These embedded systems, in general, consist of sensors, data processor, communication and actuator. As discussed in [1] distributed parameter systems (which are described by partial differential/difference equations) can be approximated by a large number of interconnected finite dimensional systems. In recent years, technological development in MEMS has made it possible the idea of placing these devices in each interconnected subsystem for efficient control (of distributed parameter systems). As discussed in [1] some examples are: distributed flow control for drag reduction and smart mechanical structures.

Due to limited power supply, the transmission of information from MEMS is subject to short distance, noise, and limited capacity. References [1]-[6] can be viewed as an attempt to address some of the technical issues concerning communication and control of distributed finite dimensional systems equipped with these microscopic embedded systems.

In the present paper we address similar questions by developing a uniform Time Division Multiple Access (TDMA) scheme and use of information theoretic tools for analysis. TDMA scheme is used to avoid collision. The large scale system considered in this paper consists of distributed linear time-invariant partially observed subsystems with Gaussian process and measurement noise. The information is exchanged between subsystems via slow fading Additive White Gaussian Noise (AWGN) channels subject to path loss. Thus, dynamic model and communication channel considered in this paper are the major generalization of the ones addressed in [2]-[6], in which the dynamic system and channel are noiseless. For the linear large scale system, as described above, the quadratic cost functional is used. Encoders, decoders and decentralized controllers are designed for reliable data reconstruction and mean square stability. Encoders and decoders are obtained using Source-Channel matching technique [7]. And the stabilizing controllers are obtained from a suboptimal control solution. In designing the control laws the effects of noisy and limited capacity of transmission are considered. Hence, the proposed scheme is suitable for controlling networks of MEMS.

The paper is organized as follows: In Section II, problem formulation is described. In this section we also present a TDMA scheme. In Section III, we present the dynamic model for the large scale system; and in Section IV we present encoders, decoders and controllers.

A. Farhadi is a Postdoctoral Fellow at the School of Information Technology and Engineering, University of Ottawa, Ontario. E-mail: afarhadi@alumni.uottawa.ca

N. U. Ahmed is a Professor at the School of Information Technology and Engineering, University of Ottawa, Ontario. E-mail: ahmed@site.uottawa.ca
This work is supported by NSERC.

II. Problem Formulation

Throughout the paper we adopt the following notations: Logarithm of base 2 is denoted by $\log(\cdot)$. The transpose of A where A can be either matrix or vector is denoted by $A^{'}$. Euclidean norm with weight R on any finite dimensional space is denote by $||\cdot||_R$. The space of all matrices $A \in \Re^{q\times o}$ is denoted by $M(q\times o)$. The inverse of a square matrix $A \in M(q\times q)$ is denoted by A^{-1}; and $diag(\cdot)$ denotes block diagonal matrix. The covariance of a Random Variable (R.V.) X is denoted by $Cov(X)$. The cross covariance matrix of two R.V.'s X and Y is denoted by $Cov(X,Y)$. The nominal (Gaussian) density function with mean $\bar{x}$ and covariance $\bar{V}$ is denoted by $N(\bar{x},\bar{V})$. Gaussian R. V. X described by the density function $N(\bar{x},\bar{V})$ is denoted by $X\sim N(\bar{x},\bar{V})$.

Dynamic System: Consider a large scale system with M interconnected subsystems. Let $x_t^{(i)} \in \Re^{n_i}$ be the state, $u_t^{(i)} \in \Re^{d_i}$ be the control and $w_t^{(i)} \in \Re^{g_i}$ be the process noise of the ith subsystem ($i \in \{1,2,...,M\}$). Also, let $y_t^{(i)} \in \Re^{m_i}$ be the observation and $v_t^{(i)} \in \Re^{h_i}$ be the measurement noise. Moreover, let the set o_i denote the set of subsystems that can affect the ith subsystem dynamics via their state variables and control signals. In many applications, such as applications involving MEMS, it is more reasonable to assume that the ith subsystem is affected by the neighboring subsystems. In other words, in such applications, o_i is a finite set which includes neighboring subsystems. It is also assumed that each subsystem is linear time-invariant subject to Gaussian process and measurement noise, as described below:

$$\left\{\begin{array}{c} x_{t+1}^{(i)} = A_i x_t^{(i)} + B_i u_t^{(i)} + C_i w_t^{(i)} \\ +\sum_{k\in o_i} D_{ik} x_t^{(k)} + \sum_{k\in o_i} E_{ik} u_t^{(k)}, \\ y_t^{(i)} = F_i x_t^{(i)} + G_i v_t^{(i)}, \quad x_0^{(i)} = \xi^{(i)}, \\ i=1,...,M, \end{array}\right. \tag{1}$$

where $x_t^{(k)} \in \Re^{n_k}$ ($k \in o_i$) is the state and $u_t^{(k)} \in \Re^{d_k}$ is the control signal of the kth subsystem that affect the ith subsystem dynamics. $A_i \in M(n_i \times n_i)$ is the system matrix and $B_i \in M(n_i \times d_i)$ is the control matrix of the ith subsystem. Matrices $D_{ik} \in M(n_i \times n_k)$ and $E_{ik} \in M(n_i \times d_k)$ are interconnection matrices. Moreover, $C_i \in M(n_i \times g_i)$, $F_i \in M(m_i \times n_i)$ and $G_i \in M(m_i \times h_i)$. Gaussian R.V. $\xi^{(i)}$ is described by the density function $N(\bar{x}_0^{(i)}, \bar{V}_0^{(i)})$. Furthermore, $w_t^{(i)}$ i.i.d. $\sim N(0, \Sigma_w^{(i)})$ and $v_t^{(i)}$ i.i.d. $\sim N(0, \Sigma_v^{(i)})$. R.V.'s $\{\xi_0^{(i)}, w_t^{(p)}, v_t^{(q)}\}$ $(i, p, q \in \{1, 2, ..., M\})$ are mutually independent. Also, $\{w_t^{(i)}, w_t^{(j)}\}$ and $\{v_t^{(i)}, v_t^{(j)}\}$ ($i \in \{1, ..., M\}$, $j (\neq i) \in \{1, ..., M\}$) are mutually independent. But R.V.'s $\xi_0^{(i)}$ and $\xi_0^{(j)}$ may be stochastically dependent with known cross covariance matrix $Cov(\xi_0^{(i)}, \xi_0^{(j)})$.

Time Division Multiple Access (TDMA) Scheme: Each subsystem broadcasts information about observation and control signal to neighboring subsystems. Therefore, there is a possibility of collision in the broadcasted information; and in order to avoid such collision we need to employ an appropriate technique.

As it has been discussed in [8], in problems such as MEMS with components having access to limited power supply, TDMA based schemes may be more energy efficient than other protocols. Therefore, we employ a TDMA scheme, as described below, to avoid collision.

Based on the communication range of the broadcasted information about control signals, the large scale system can be decomposed into disjoint groups of subsystems, denoted by groups $1, 2, ..., N_1$, in which subsystems in each group have non-overlapping communication range. Similarly, based on the communication range of the broadcasted information about observation signals, the large scale system can be decomposed into disjoint groups $N_1 + 1, ..., N_2$. Note that for two groups $i \in \{1, 2, ..., N_1\}$ and $j (\neq i) \in \{1, 2, ..., N_1\}$; or $i \in \{N_1 + 1, ..., N_2\}$ and $j (\neq i) \in \{N_1 + 1, ..., N_2\}$, the set $i \cap j$ is empty. But for $i \in \{1, 2, ..., N_1\}$ and $j \in \{N_1 + 1, ..., N_2\}$, the set $i \cap j$ may be non-empty. Subsystems in each group can broadcast information about control or observation signals simultaneously. At the same time the neighboring subsystems are waiting to receive this information; and they will receive the broadcasted information without collision. Following this fact, we divide each time step into N_2 equal size, non-overlapping time slots. In the first N_1 time slots we exchange information about control signals and in the rest of time step we exchange information about observation signals, as described below:

We allocate the first time slot to all subsystems in group 1 to broadcast information about control signals simultaneously. At this time the transmitters of all subsystems in other groups (i.e.,$2, 3, ..., N_1$) are shut down; while the receivers of the neighboring subsystems of the systems in group 1 are on; and they are waiting to receive the broadcasted information. Similarly, we allocate the second time slot to all subsystems in group 2; and we follow this procedure until we allocate the N_1th time slot to all subsystems in group N_1 to broadcast information about control signals simultaneously. We follow similar procedure for time slots $N_1 + 1, ..., N_2$; and we broadcast information about observation signals.

Communication Channel: The communication link from the ith subsystem to the jth neighboring subsystem is modeled by a multi input, multi output AWGN channel with channel input (i.e., transmitter output of the ith subsystem) $T_t^{(i)}$ and channel output (i.e., receiver input of the jth subsystem) $R_t^{(ji)}$. This channel is subject to path loss and slow fading. Depending on the transmitted signal, the channel input is either $T_t^{(i)} = f_t^i(y^{(i)}(t))$ or $T_t^{(i)} = e_t^i(u^{(i)}(t))$, where $f_t^i(\cdot)$ and $e_t^i(\cdot)$ are encoding functions ($f_t^{(i)}(\cdot)$ and $e_t^{(i)}(\cdot)$ are invertible functions and in general they can be nonlinear).

When the information about observation signal is transmitted through the channel (i.e., $T_t^{(i)} = f_t^i(y^{(i)}(t))$) the channel is described by

$$\begin{aligned} &R_t^{(ji)} = h_t^{(ji)} \cdot f_t^{(i)}(y^{(i)}(t)) + \zeta_t^{(ji)}, \\ &T_t^{(i)} = f_t^{(i)}(y^{(i)}(t)) \in \Re^{p_i}, R_t^{(ji)} \in \Re^{q_i},\ \zeta_t^{(ji)} \in \Re^{q_i}, \\ &E||T_t^{(i)}||^2 \leq P_t^{(i)},\ \zeta_t^{(ji)}\ i.i.d.\ \sim N(0, \Gamma^{(ji)}), \end{aligned} \tag{2}$$

in which the channel gain $h_t^{(ji)}$ is given by

$$h_t^{(ji)} \triangleq \alpha_t^{(ji)} / (d_{ji})^{a_{ji}}, \tag{3}$$

where d_{ji} is the line of sight distance and a_{ji} ($a_{ji} \in \{0, 1, 2, ...\}$) is the path loss factor. The random matrix $\alpha_t^{(ji)}$ having components with Rayleigh distribution represents the fading effect. In expression (2) R.V. $\zeta_t^{(ji)}$ represents the additive Gaussian channel noise and $P_t^{(i)}$ denotes the channel input power constraint.

Throughout, it is assumed that the communication channel is subject to slow fading. That is, at any time $t \in \mathbf{N}_+$, the channel gain $h_t^{(ji)}$ is known to both transmitter and receiver. Furthermore, it is assumed that $\lim_{t\to\infty} \alpha_t^{(ji)} = \alpha^{(ji)}$, almost surly, where $\alpha^{(ji)}$ is fixed. The communication model similar to (2) has been used in networks of MEMS and sensor networks, in which the path loss factor $a_{ji} = 3$ has been used (e.g., see [8]).

Objective: Mean square stability means that the state variables have bounded second moment. The objective of this paper is to find control sequence $\{u_t^{(i)}; i = 1, 2, ..., M,\ t \in \mathbf{N}_+\}$ to stabilize all subsystems given by (1) in the mean square sense. That is,

$$\sup_{t \in \mathbf{N}_+} E||x_t^{(i)}||^2 < \infty, \quad \forall i \in \{1, 2, ..., M\}. \tag{4}$$

Information Available at Each subsystem: The information available at the ith subsystem which is used to produce the control signal $u_t^{(i)}$ consists of its past observation and past control signal as well as a noisy version of the past observation and past control signals of the neighboring subsystems.

It is well known that for a linear system with a quadratic cost functional having positive weighting matrices, the optimal/suboptimal control solution results in the mean square stability. For large scale linear systems, when the controllers have limited access to the observation and control signals of other subsystems, the optimal solution is unknown. Therefore, in addressing the stability question of such systems we look for a suboptimal solution. Hence, the stabilizing controllers for system (1) subject to available information, as indicated above, can be obtained from a suboptimal control solution with the following quadratic payoff functional

$$J = \lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} E\Big[\sum_{i=1}^{M}(||x_t^{(i)}||_{K_i}^2 + ||u_t^{(i)}||_{H_i}^2)\Big], \qquad (5)$$

where $K_i = K_i' \in M(n_i \times n_i)$ is positive semi-definite and $H_i = H_i' \in M(d_i \times d_i)$ is positive definite.

III. DYNAMIC MODEL FOR LARGE SCALE SYSTEM

As discussed in [9] a large scale system can be decomposed into clusters of subsystems, in which each cluster includes subsystems which are strongly coupled. The overall system is then represented as a system of weak interconnection of the resulting clusters.

In this paper we are concerned with the large scale system (1) which can be decomposed into disjoint clusters $\mathcal{S}_r$, $r = 1, 2, ..., l$, where each cluster includes a set of neighboring subsystems which are strongly interconnected. Clusters $\mathcal{S}_r$ and $\mathcal{S}_{r+1}$ weakly affect their dynamics. Cluster $\mathcal{S}_{r+1}$ weakly affects the dynamics of cluster $\mathcal{S}_r$ via its state variables and control signals; and cluster $\mathcal{S}_r$ affects weakly the dynamics of cluster $\mathcal{S}_{r+1}$ via its control signals. We level the full (large scale) system as follows: Cluster $\mathcal{S}_1$ includes subsystems $\{a_1,...,b_1\}$ ($a_1 = 1$, $b_1 \geq a_1$), cluster $\mathcal{S}_2$ includes subsystems $\{a_2,...,b_2\}$ ($a_2 = b_1+1$, $b_2 \geq a_2$), ..., and cluster $\mathcal{S}_l$ includes subsystems $\{a_l, ..., b_l\}$ ($a_l = b_{l-1}+1$, $b_l = M \geq a_l$). Note that subsystem a_{r+1} ($r \in \{1, 2, ..., l-1\}$) is the closest subsystem of cluster $\mathcal{S}_{r+1}$ to the subsystems of cluster $\mathcal{S}_r$.

The information about control and observation signal of each subsystem of cluster $\mathcal{S}_r$ ($r \in \{1, 2, ..., l\}$) is available at other subsystems of this cluster. Moreover, information about those control signals of clusters $\mathcal{S}_r$ and $\mathcal{S}_{r+1}$ which affect their dynamics is available at the subsystems of clusters $\mathcal{S}_{r+1}$ and $\mathcal{S}_r$ which are affected. Furthermore, information about control signal of subsystem a_{r+1} in cluster $\mathcal{S}_{r+1}$ is available at the subsystems of cluster $\mathcal{S}_r$ which are affected by the state variables of cluster $\mathcal{S}_{r+1}$.

Let $X_t^{(r)}$ denote the vector of state variables of all subsystems of cluster $\mathcal{S}_r$ at time t. Similarly, let $U_t^{(r)}$ denote the vector of control signals, $W_t^{(r)}$ the vector of process noises, $Y_t^{(r)}$ the vector of observation signals and $V_t^{(r)}$ the vector of measurement noises of all subsystems of cluster $\mathcal{S}_r$. Then, cluster $\mathcal{S}_r$ is described by the following dynamic model

$$\left\{\begin{array}{c} X_{t+1}^{(r)} = A^{(r)}X_t^{(r)} + B^{(r)}U_t^{(r)} + C^{(r)}W_t^{(r)} + \\ D^{(r+1)}X_t^{(r+1)} + M^{(r+1)}U_t^{(r+1)} + N^{(r-1)}U_t^{(r-1)} \\ Y_t^{(r)} = F^{(r)}X_t^{(r)} + G^{(r)}V_t^{(r)},\ r = 1, 2, ..., l, \end{array}\right. \qquad (6)$$

where matrices $A^{(r)}$, $B^{(r)}$, $C^{(r)}$, $F^{(r)}$, and $G^{(r)}$ represent the interconnection among subsystems of cluster $\mathcal{S}_r$. Interconnection matrices $D^{(r+1)}$ and $M^{(r+1)}$ represent the effect of state variables and control signals of cluster $\mathcal{S}_{r+1}$, respectively, on the subsystems of cluster $\mathcal{S}_r$. Similarly, interconnection matrix $N^{(r-1)}$ represent the effect of the control signals of cluster $\mathcal{S}_{r-1}$. Note that $D^{(l+1)} = 0$, $M^{(l+1)} = 0$, and $N^{(0)} = 0$. Also note that interconnection matrices mostly contain zero components because of the weak interconnection among clusters.

Thus, the overall system is described by the following system of equations.

$$\left\{\begin{array}{c} X_{t+1} = AX_t + BU_t + CW_t, \\ Y_t = FX_t + GV_t, \end{array}\right. \qquad (7)$$

where $X_t = \big(X_t^{(1)'} \ ... \ X_t^{(l)'}\big)'$ is the state of the full (large scale) system, $U_t = (U_t^{(1)'} \ ... \ U_t^{(l)'})'$ is the control vector, $W_t = \big(W_t^{(1)'} ... \ W_t^{(l)'}\big)'$ is the process noise, $V_t = \big(V_t^{(1)'} \ ... \ V_t^{(l)'}\big)'$ is the measurement noise, and $Y_t = \big(Y_t^{(1)'} \ ... \ Y_t^{(l)'}\big)'$ is the observation vector. In (7) $C = diag(C^{(1)} \ C^{(2)} ... \ C^{(l)})$, $F = diag\,(\,F^{(1)} \ F^{(2)} \ ... \ F^{(l)}\,)$ and $G = diag\,(\,G^{(1)} \ G^{(2)} \ ... \ G^{(l)}\,)$. Furthermore, matrices A and B are given by the following block matrices:

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

$$B = \begin{pmatrix} B^{(1)} & M^{(2)} & 0 & 0 & 0 & 0 \\ N^{(1)} & B^{(2)} & M^{(3)} & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & N^{(l-2)} & \cdot & M^{(l)} \\ 0 & 0 & 0 & 0 & N^{(l-1)} & B^{(l)} \end{pmatrix}$$

IV. Control Through Communication Channels With Limited Power

In some applications such as sensor networks and applications involving MEMS with components having access to limited power supply, the power for transmission is limited. Hence, the transmission is subject to limited capacity and noise. Therefore, in such applications, it is important to exchange all or at least some information under minimum capacity (power). This is the subject of study in this section. Here, for simplicity of analysis, we assume information about control sequences is exchanged without communication constraints. But, information about observation sequences is exchanged via AWGN channel (2) subject to limited power.

We also assume the observation signals $y_t^{(i)}$'s are scalar. Also, the AWGN channel (2) is single input, single output.
In this section, we present a methodology for designing encoders and decoders for reliable transmission of the observation sequences when the capacity (power) used for transmission is minimum. We also present decentralized controllers for mean square stability. Note that the information available at each subsystem is available at the encoder, decoder, and controller of that subsystem.
Consider system (7). In view of the system decomposed as suggested above, the payoff functional (5) for system (7) can now be written as follows

$$J = J^{(1)} + J^{(2)} + ... + J^{(l)}, \tag{8}$$

where for $r = 1, 2, ..., l$

$$J^{(r)} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E(||X_t^{(r)}||^2_{Q_r} + ||U_t^{(r)}||^2_{R_r}),$$
$$Q_r = diag(K_{a_r}...K_{b_r}), R_r = diag(H_{a_r}...H_{b_r}) \tag{9}$$

corresponds to cluster $\mathcal{S}_r$. The steps taken to design decentralized control laws are described below: For each cluster we choose the control vector in an appropriate way such that we compensate the weak interconnection effects from other clusters. For cluster $\mathcal{S}_r$ this compensation is achieved using the information available from clusters $\mathcal{S}_{r-1}$ and $\mathcal{S}_{r+1}$ which is used to regulate the dynamics of cluster $\mathcal{S}_r$. Then, the stabilizing controllers for cluster $\mathcal{S}_r$ are obtained independently by finding a suboptimal control solution for the payoff functional $J^{(r)}$. Finally, the stability of the whole system, when these controllers are enforced, is proved.
According to the interconnection matrices A and B of system (7), it is convenient to start designing stabilizing controllers from cluster $\mathcal{S}_l$ which is affected only by the control vector of cluster $\mathcal{S}_{l-1}$. We then design stabilizing controllers for cluster $\mathcal{S}_{l-1}$, cluster $\mathcal{S}_{l-2}$, ..., till cluster $\mathcal{S}_1$.

A. *Control Law For Cluster $\mathcal{S}_r$ ($r = l, ..., 1$)*

Consider cluster $\mathcal{S}_r$, as described by expression (6). In the followings we present a methodology for designing encoders, decoders, and stabilizing controllers, for mean square stability such that the capacity (power) used for transmission is minimum.
Linear Encoders and Decoders for Subsystems of Cluster $\mathcal{S}_r$: The information about observation signal $y_t^{(i)}$ of subsystem $i \in \{a_r, a_r+1, ..., b_r\}$ in cluster $\mathcal{S}_r$ is transmitted to subsystem $j(\neq i) \in \{a_r, a_r+1, ..., b_r\}$ in this cluster via AWGN channel (2). Let the non-negative scalars $\beta_t^{(ji)}$ and $\gamma_t^{(ji)}$ be the encoding and decoding gain, respectively. Also, let $\mathcal{F}_{t-1}^{(i)}$ denote the available information at subsystem i for each $t \in \mathbf{N}_+$. Subsystem i uses this information and produces the mean square state estimation $\hat{x}_t^{(ii)} \triangleq E[x_t^{(i)}|\mathcal{F}_{t-1}^{(i)}]$. This estimation is used in the encoding function, $T_t^{(ji)} = f_t^{(ji)}(y_t^{(i)})$, where

$$f_t^{(ji)}(y_t^{(i)}) \triangleq \beta_t^{(ji)} k_t^{(i)} \in \Re, \quad k_t^{(i)} \triangleq y_t^{(i)} - F_i \hat{x}_t^{(ii)} \in \Re. \tag{10}$$

The message $T_t^{(ji)}$ is broadcasted via the AWGN channel (2) to the neighboring subsystems and it receives at subsystem j.
Subsystem j receives $\check{k}_t^{(ji)} \triangleq h_t^{(ji)} T_t^{(ji)} + \zeta_t^{(ji)} \in \Re$, where $\check{k}_t^{(ji)}$ is the received signal, $h_t^{(ji)}$ is the channel gain as described by expression (3), and the channel noise $\zeta_t^{(ji)}$ is an i.i.d. sequence with distribution $N(0, \Gamma^{(ji)})$.
The decoding function for this subsystem is $\bar{y}_t^{(ji)} = \bar{f}_t^{(ji)}(\check{k}_t^{(ji)})$, where $\bar{y}_t^{(ji)}$ is the reconstructed version of the observation signal $y_t^{(i)}$ at subsystem j; and

$$\bar{f}_t^{(ji)}(\check{k}_t^{(ji)}) \triangleq \bar{k}_t^{(ji)} + F_i \hat{x}_t^{(ii)} \in \Re,$$
$$\bar{k}_t^{(ji)} \triangleq (h_t^{(ji)})^{-1} \gamma_t^{(ji)} \check{k}_t^{(ji)} \in \Re, \tag{11}$$

where $\bar{k}_t^{(ji)}$ is the reconstructed version of the innovation sequence $k_t^{(i)}$, at sub-system j.
Note that the decoding function (11) involves the state estimation $\hat{x}_t^{(ii)}$. As it will be shown later this estimation is available at subsystem j via the control signal of subsystem i . Also, note that the encoder and decoder, as described above, are causal functions of the source messages.
Control Law for Subsystem j in Cluster $\mathcal{S}_r$: For each $t \geq 0$, in addition of the observation signal $y_t^{(j)}$, a noisy version of the observation signal of subsystem $i(\neq j) \in \{a_r, a_r + 1, ..., b_r\}$ is available at subsystem j. That is, the following observation vector $\bar{Y}_t^{(rj)} \triangleq (\bar{y}_t^{(ja_r)'} \; ... \; y_t^{(j)'} \; ... \; \bar{y}_t^{(jb_r)'})'$ is available.
Let $\bar{Z}_t^{(rj)} \triangleq (z'_{a_r} \; ... \; z'_j \; ...z'_{b_r})'$, where $z_j = y_t^{(j)}$, $z_i = \bar{y}_t^{(ji)} + \gamma_t^{(ji)} \beta_t^{(ji)} F_i \hat{x}_t^{(ii)} - F_i \hat{x}_t^{(ii)}$. Also, let

$$F_t^{(rj)} \triangleq diag(\gamma_t^{(ja_r)} \beta_t^{(ja_r)} F_{a_r} ... F_j ... \gamma_t^{(jb_r)} \beta_t^{(jb_r)} F_{b_r})$$
$$G_t^{(rj)} \triangleq diag(\gamma_t^{(ja_r)} \beta_t^{(ja_r)} G_{a_r} ... G_j ... \gamma_t^{(jb_r)} \beta_t^{(jb_r)} G_{b_r})$$
$$\vartheta_t^{(rj)} \triangleq (\bar{\zeta}_t^{(ja_r)'} \; ... \; \bar{\zeta}_t^{(jj)'} \; ... \; \bar{\zeta}_t^{(jb_r)'})', \; \bar{\zeta}_t^{(jj)} = 0,$$
$$\bar{\zeta}_t^{(ji)} = (h_t^{(ji)})^{-1} \gamma_t^{(ji)} \zeta_t^{(ji)}.$$

The control vector $U_t^{(r)} = (u_t^{(a_r)'} \; ... \; u_t^{(b_r)'})'$ is applied by the subsystems of cluster $\mathcal{S}_r$. Control signal $u_t^{(a_r)}$ is applied by subsystem a_r,..., and the control signal $u_t^{(b_r)}$ is applied by subsystem b_r at time $t \in \mathbf{N}_+$. By choosing $U_t^{(r)}$ in an appropriate way, we can compensate the interconnection effects caused by clusters $\mathcal{S}_{r-1}$ and $\mathcal{S}_{r+1}$. Hence, we can design controllers for cluster $\mathcal{S}_r$ independently of other clusters.
Recall that at time t, those control sequences of clusters $\mathcal{S}_{r-1}$ and $\mathcal{S}_{r+1}$ which affect the subsystems of cluster $\mathcal{S}_r$ via matrices $N^{(r-1)}$ and $M^{(r+1)}$, respectively, are available at the subsystems which are affected. Therefore, subsystems of cluster $\mathcal{S}_r$ can use this information to regulate their dynamics.
The appropriate choice for control signal $U_t^{(r)}$ is given by

$$U_t^{(r)} = \tilde{U}_t^{(r)} + B^{(r)'} (B^{(r)} B^{(r)'})^{-1} \Big(- M^{(r+1)} U_t^{(r+1)}$$
$$- D^{(r+1)} \hat{X}_t^{(r+1 \; a_{r+1})} - N^{(r-1)} U_t^{(r-1)} \Big), \tag{12}$$

where the new control vector $\tilde{U}_t^{(r)} = (\tilde{u}_t^{(a_r)'} \ ... \ \tilde{u}_t^{(b_r)'})'$ is available at all subsystems of cluster $\mathcal{S}_r$. Note that $\hat{X}_t^{(r+1 \ a_{r+1})}$ is the mean square estimation of the state variables of cluster $\mathcal{S}_{r+1}$ produced at subsystem a_{r+1} (in cluster $\mathcal{S}_{r+1}$). This estimation is available at those subsystems of cluster $\mathcal{S}_r$ which are affected by the state variables of cluster $\mathcal{S}_{r+1}$, via the control signal $\tilde{u}_t^{(a_{r+1})} = -\Delta_{a_{r+1}}\hat{X}^{(r+1 \ a_{r+1})}$, where $\Delta_{a_{r+1}}$ is the controller gain (it will be defined shortly). Hence, these subsystems can compute $\hat{X}_t^{(r+1 \ a_{r+1})} = -(\Delta'_{a_{r+1}}\Delta_{a_{r+1}})^{-1}\Delta'_{a_{r+1}}\tilde{u}_t^{(a_{r+1})}$; and therefore can use it in (12). Note that for cluster $\mathcal{S}_l$, we set $\tilde{u}_t^{(a_{l+1})} = 0$, $M^{(l+1)} = 0$ and $D^{(l+1)} = 0$.

Using the state dynamics (6) and by substituting the control vector $U_t^{(r)}$, as given by (12), we have $X_{t+1}^{(r)} = A^{(r)}X_t^{(r)} + B^{(r)}\tilde{U}_t^{(r)} + C^{(r)}W_t^{(r)} + D^{(r+1)}E_t^{(r+1 \ a_{r+1})}, \quad r = l, ..., 1,$ $E_t^{(r+1 \ a_{r+1})} \triangleq X_t^{(r+1)} - \hat{X}_t^{(r+1 \ a_{r+1})}$, where the estimation error $E_t^{(r+1 \ a_{r+1})}$ is an orthogonal Gaussian sequence with distribution $N(0, \Xi_t^{(r+1 \ a_{r+1})})$. Note that for the lth cluster we set $\Xi_t^{(r+1 \ a_{r+1})} = 0$.

Following the information available at the subsystems of cluster $\mathcal{S}_r$, we use the dynamic model (13), as given below, and we design the control law for subsystem j.

$$\begin{cases} X_{t+1}^{(r)} = A^{(r)}X_t^{(r)} + B^{(r)}\tilde{U}_t^{(r)} + C^{(r)}W_t^{(r)} \\ \qquad\qquad + D^{(r+1)}E_t^{(r+1 \ a_{r+1})} \\ \bar{Z}_t^{(rj)} = F_t^{(rj)}X_t^{(r)} + G_t^{(rj)}V_t^{(r)} + \vartheta_t^{(rj)}. \end{cases} \tag{13}$$

Let $\tilde{J}^{(r)}$ denote the payoff functional $J^{(r)}$, as given by expression (9), where the control vector $U_t^{(r)}$ has been replaced by $\tilde{U}_t^{(r)}$. For above system with the payoff functional $\tilde{J}^{(r)}$, we follow LQG methodology [10] and we find the optimal control $\tilde{U}_t^{(r*)} = (\tilde{u}_t^{(a_r*)'} \ ... \ \tilde{u}_t^{(j*)'} \ ... \ \tilde{u}_t^{(b_l*)'})'$, in which just the control signal $\tilde{u}_t^{(j*)}$ is applied at time t.

Here, we shall assume $\gamma_t^{(ji)}$ and $\beta_t^{(ji)}$ asymptotically converge to fixed limits $\gamma^{(ji)}$ and $\beta^{(ji)}$, respectively. Following this assumption we have $\lim_{t\to\infty} F_t^{(rj)} = F^{(rj)} \triangleq diag(\gamma^{(ja_r)}\beta^{(ja_r)}F_{a_r} \ ... \ F_j \ ... \ \gamma^{(jb_r)}\beta^{(jb_r)}F_{b_r})$, $\lim_{t\to\infty} G_t^{(ji)} = G^{(ji)} \triangleq diag(\gamma^{(ja_r)}\beta^{(ja_r)}G_{a_r} \ ... \ G_j \ ... \ \gamma^{(jb_r)}\beta_t^{(jb_r)} .G_{b_r})$. Moreover, $\Sigma_t^{(ji)} \triangleq Cov(\bar{\zeta}_t^{(ji)}) = \Big(\gamma_t^{(ji)}(d_{ji})^{a_{ji}}/\alpha_t^{(ji)}\Big)^2\Gamma^{(ji)}$, in which under assumption that the limit, $\lim_{t\to\infty}\alpha_t^{(ji)} = \alpha^{(ji)}$ exists, which was made earlier, the limit $\Sigma^{(ji)} \triangleq \lim_{t\to\infty}\Sigma_t^{(ji)}$ exists and it is given by $\Sigma^{(ji)} = \Big(\gamma^{(ji)}(d_{ji})^{a_{ji}}/\alpha^{(ji)}\Big)^2\Gamma^{(ji)}$.

The optimal control $\tilde{U}_t^{(r*)}$ is obtained under the following assumptions:

$(\mathbf{a1})$: The pair $(A^{(r)}, B^{(r)})$ is stabilizable.

$(\mathbf{a2})$: The pair $(F^{(rj)}, A^{(r)})$ is detectable.

$(\mathbf{a3})$: The pair $(Q_r^{\frac{1}{2}}, A^{(r)})$ is detectable.

$(\mathbf{a4})$: The pair $(A^{(r)}, (C^{(r)}\Sigma_W^{(r)}C^{(r)'} + D^{(r+1)}\Xi_\infty^{(r+1 \ a_{r+1})}D^{(r+1)'})^{\frac{1}{2}})$ is stabilizable where $\Sigma_W^{(r)} = Cov(W_t^{(r)})$ and $\Xi_\infty^{(r+1 \ a_{r+1})} = \lim_{t\to\infty}\Xi_t^{(r+1 \ a_{r+1})}$.

Under above assumptions, the optimal control which minimizes the payoff functional $\tilde{J}^{(r)}$ follows from the standard LQG results [10]. The optimal control is given by $\tilde{U}_t^{(r*)} = (\tilde{u}_t^{(a_r*)'} \ ... \ \tilde{u}_t^{(j*)'} \ ... \ \tilde{u}_t^{(b_r*)'})' = -\Delta^{(r)}\hat{X}_t^{(rj)}$ where the controller gain $\Delta^{(r)}$ is given by $\Delta^{(r)} = (R_r + B^{(r)'}\Lambda^{(r)}B^{(r)})^{-1}B^{(r)'}\Lambda^{(r)}A^{(r)}$, with $\Lambda^{(r)}$ being the unique positive semi-definite solution of the following Algebraic Riccati equation $\Lambda^{(r)} = A^{(r)'}\Lambda^{(r)}A^{(r)} - A^{(r)'}\Lambda^{(r)}B^{(r)}\Big(B^{(r)'}\Lambda^{(r)}B^{(r)} + R_r\Big)^{-1}B^{(r)'}\Lambda^{(r)}A^{(r)} + Q_r$. Note that the controller gain $\Delta^{(r)}$ has the following representation.

$$\Delta^{(r)} = \begin{pmatrix} \Delta_{a_r} \\ . \\ . \\ . \\ \Delta_j \\ . \\ . \\ . \\ \Delta_{b_r} \end{pmatrix} \in M\Big((d_{a_r} + ... + d_j + ... + d_{b_r}) \times (n_{a_r} + ... + n_j + ... + n_{b_r})\Big),$$
$$\text{where } \Delta_j \in M(d_j \times (n_{a_r} + ... + n_j + ... + n_{b_r})). \tag{14}$$

Therefore, $\tilde{U}_t^{(r*)} = (\tilde{u}_t^{(a_r*)'} \ ... \ \tilde{u}_t^{(j*)'} \ ... \ \tilde{u}_t^{(b_r*)'}) = -\Delta^{(r)}\hat{X}_t^{(rj)} \Leftrightarrow \tilde{u}_t^{(j*)} = -\Delta_j\hat{X}_t^{(rj)}, \quad j \in \{a_r, a_r+1, ..., b_r\}$, where the state estimation $\hat{X}_t^{(rj)}$ is given by the following recursive equation

$$\begin{aligned} &\hat{X}_{t+1}^{(rj)} = A^{(r)}\hat{X}_t^{(rj)} + B^{(r)}(\tilde{u}_t^{(a_r)'} \ ... \ \tilde{u}_t^{(j*)'} \ ... \ \tilde{u}_t^{(b_r)'})' \\ &+ L_t^{(rj)}(\bar{Z}_t^{(rj)} - F_t^{(rj)}\hat{X}_t^{(rj)}), \\ &\hat{X}_0^{(rj)} = (\bar{x}_0^{(a_r)'} \ ... \ \bar{x}_0^{(b_r)'})', \quad \tilde{u}_t^{(j*)} = -\Delta_j\hat{X}_t^{(rj)}, \\ &j \in \{a_r, a_r+1, ..., b_r\}. \end{aligned} \tag{15}$$

In above expression the estimation gain $L_t^{(rj)}$ is given by $L_t^{(rj)} = A^{(r)}\Xi_t^{(rj)}F_t^{(rj)'}\Big(F_t^{(rj)}\Xi_t^{(rj)}F_t^{(rj)'} + G^{(rj)}\Sigma_V^{(r)}G^{(rj)'} + \Upsilon_t^{(rj)}\Big)^{-1}$, $\Upsilon_t^{(rj)} \triangleq Cov(\vartheta_t^{(rj)}) = diag(\Sigma_t^{(ja_r)} \ ...\Sigma_t^{(jj)} \ ... \ \Sigma_t^{(jb_r)})$, $\Sigma_t^{(jj)} = 0$, with the covariance of the estimation error $\Xi_t^{(rj)}$ being the solution of the following Riccati equation $\Xi_{t+1}^{(rj)} = A^{(r)}\Xi_t^{(rj)}A^{(r)'} - A^{(r)}\Xi_t^{(rj)}F_t^{(rj)'}\Big(F_t^{(rj)}\Xi_t^{(rj)}F_t^{(rj)'} + F_t^{(rj)}\Sigma_V^{(r)}F_t^{(rj)'} + \Upsilon_t^{(rj)}\Big)^{-1}F_t^{(rj)}\Xi_t^{(rj)}A^{(r)'} + C^{(r)}\Sigma_W^{(r)}C^{(r)'} + D^{(r+1)}\Xi_t^{(r+1 \ a_{r+1})}D^{(r+1)'}$, $\Xi_0^{(rj)} = Cov(\xi_0^{(a_r)'} \ ... \ \xi_0^{(b_r)'})'$, $\Sigma_V^{(r)} = Cov(V_t^{(r)})$. Note that, as we discussed earlier, just the control signal $\tilde{u}_t^{(j*)} = -\Delta_j\hat{X}_t^{(rj)}$ is applied at time t. Also, note that the state estimation $\hat{x}_t^{(ii)}$ is available at subsystem j via $\tilde{u}_t^{(i)} = -\Delta_i\hat{X}_t^{(ri)}$. Subsystem j uses this information and computes $\hat{X}_t^{(ri)} = (\hat{x}_t^{(ia_r)'} \ ... \ \hat{x}_t^{(ii)'} \ ...\hat{x}_t^{(ib_r)'})' = -(\Delta'_i\Delta_i)^{-1}\Delta'_i\tilde{u}_t^{(i)}$.

Selection of the Encoding and Decoding Gain $\beta_t^{(ji)}$ and $\gamma_t^{(ji)}$: Consider the encoding and decoding functions (10) and (11), respectively. The encoding and decoding gain $\beta_t^{(ji)}$ and $\gamma_t^{(ji)}$ are chosen such that i) $E||y_t^{(i)} - \bar{y}_t^{(ji)}||^2 = E||k_t^{(i)} - \bar{k}_t^{(ji)}||^2 = D_v^{(ji)}$ where $D_v^{(ji)} \geq 0$ is the desired distortion level (it is an auxiliary parameter). ii) The capacity used for transmission is minimum.

Solution to above problem is obtained using Source-Channel matching technique [7]. Applying this technique we have [11] $\beta_t^{(ji)} = \frac{(d_{ji})^{a_{ji}}}{\alpha_t^{(ji)}} \sqrt{\frac{\Gamma^{(ji)}\eta_t^{(ji)}}{D_v^{(ji)}}}, \gamma_t^{(ji)} = \frac{\alpha_t^{(ji)}}{(d_{ji})^{a_{ji}}} \sqrt{\frac{\eta_t^{(ji)} D_v^{(ji)}}{\Gamma^{(ji)}}}$, $\eta_t^{(ji)} \triangleq 1 - \frac{D_v^{(ji)}}{\Psi_t^{(i)}}$, where $\Psi_t^{(i)} \triangleq F_i \Theta_t^{(i)} F_i' + G_i \Sigma_v^{(i)} G_i'$ ($i \in \{a_r, a_r+1, ..., b_r\}$), in which $\Theta_t^{(i)} \in M(n_i \times n_i)$ is the $(i - a_r + 1)$th matrix on the diagonal of the block matrix $\Xi_t^{(ri)} \in M\Big((n_{a_r} + ... + n_i + ... + n_{b_r}) \times (n_{a_r} + ... + n_i + ... + n_{b_r})\Big)$, as described above. Note that under the stabilizability and detectability assumptions $(\mathbf{a2})$ and $(\mathbf{a4})$ we have $\lim_{t\to\infty} \Xi_t^{(ri)} = \Xi_\infty^{(ri)}$, where $\Xi_\infty^{(ri)}$ is the unique positive semi-definite solution of the following Algebraic Riccati equation $\Xi_\infty^{(rj)} = A^{(r)} \Xi_\infty^{(rj)} A^{(r)'} - A^{(r)} \Xi_\infty^{(rj)} F^{(rj)'} \Big(F^{(rj)} \Xi_\infty^{(rj)} F^{(rj)'} + F^{(rj)} \Sigma_V^{(r)} F^{(rj)'} + \Upsilon^{(rj)} \Big)^{-1} F^{(rj)} \Xi_\infty^{(rj)} A^{(r)'} + C^{(r)} \Sigma_W^{(r)} C^{(r)'} + D^{(r+1)} \Xi_\infty^{(r+1\ a_{r+1})} D^{(r+1)'}$, $\Upsilon^{(rj)} = \lim_{t\to\infty} \Upsilon_t^{(rj)} = diag(\Sigma^{(ja_r)} ... \Sigma^{(jj)} ... \Sigma^{(jb_r)})$, $\Sigma^{(jj)} = 0$. Following above selection for $\beta_t^{(ji)}$ and $\gamma_t^{(ji)}$, we have [11] $E||y_t^{(i)} - \bar{y}_t^{(ji)}||^2 = E||k_t^{(i)} - \bar{k}_t^{(ji)}||^2 = D_v^{(ji)}$ and $\mathcal{C}_y^{(ji)} = \mathcal{C}_k^{(ji)} = R_{SRD}^{y,\bar{y}}(D_v^{(ji)}) = R_{SRD}^{k,\bar{k}}(D_v^{(ji)}) = R^{k,\bar{k}}(D_v^{ji}) = \frac{1}{2}\log\frac{\Psi_\infty^{(i)}}{D_v^{(ji)}}$, where $\mathcal{C}_y^{(ji)}$ and $\mathcal{C}_k^{(ji)}$ are the capacity used for transmission of sequence $\{y_t^{(i)}, t \in \mathbf{N}_+\}$ and $\{k_t^{(i)}, t \in \mathbf{N}_+\}$, respectively, from subsystem i to j, $R_{SRD}^{y,\bar{y}}(D_v^{(ji)})$ and $R_{SRD}^{k,\bar{k}}(D_v^{(ji)})$ are sequential rate distortion [12], and $R^{k,\bar{k}}(D_v^{(ji)})$ is the rate distortion function [13] with single letter mean square distortion measure. Note that $\Psi_\infty^{(i)} = \lim_{t\to\infty} \Psi_t^{(i)}$ and the capacity used for reliable data reconstruction up to the distortion level $D_v^{(ji)}$, as described above, is minimum.

B. Stability Analysis

The stability of subsystems of cluster $\mathcal{S}_r$, when the controllers as indicated above are enforced, is shown in the following lemma.

Lemma 4.1: **(Stability of Clusters $\mathcal{S}_r$, $r = l, ..., 1$):** Consider cluster $\mathcal{S}_r$ and suppose assumptions $(\mathbf{a1})$-$(\mathbf{a4})$ hold. Also, let the distortion level $D_v^{(ji)}$ ($j \in \{a_r, a_r+1, ..., b_r\}$, $i (\neq j) \in \{a_r, a_r+1, ..., b_r\}$, $r = l, ..., 1$) be sufficiently small such that the following condition holds: $L_t^{(r\ j+1)} - L_t^{(rj)} \approx 0$. Then, the subsystems of cluster $\mathcal{S}_r$ are stable in the mean square sense when the controllers , as indicated by expression (12) with $\tilde{u}_t^{(j*)} = -\Delta_j \hat{X}_t^{(rj)}$ and $\hat{X}_t^{(rj)}$ as given by (15), are enforced.

Proof: It follows by employing a similar methodology as used in [14]. The complete proof can be found in [11].

Note that from the expression for the estimation gain $L_t^{(rj)}$ follows that the estimation gains $L_t^{(rj)}$ and $L_t^{(r\ j+1)}$ are similar. The only difference comes from covariance matrices $\Upsilon_t^{(rj)} = diag(\Sigma_t^{(ja_r)} ... \Sigma_t^{(jj)} ... \Sigma_t^{(jb_r)})$, ($\Sigma_t^{(jj)} = 0$) and $\Upsilon_t^{(r\ j+1)} = diag(\Sigma_t^{(j+1\ a_r)} ... \Sigma_t^{(j+1\ j+1)} ... \Sigma_t^{(j+1\ b_r)})$, ($\Sigma_t^{(j+1\ j+1)} = 0$). But, $\Sigma_t^{(ji)} = (1 - \frac{D_v^{(ji)}}{\Psi_t^{(i)}}) D_v^{(ji)}$; and therefore when the distortion level $D_v^{(ji)}$ is sufficiently small, we have $L_t^{(r\ j+1)} - L_t^{(rj)} \approx 0$.

Next, in the following theorem, under assumptions of Lemma 4.1, the stability of the whole system is proved.

Theorem 4.2: **(Stability of Full System):** Consider the large scale system (7) and suppose assumptions **(a1)-(a4)** hold. Also, let the distortion level $D_v^{(ji)}$ ($j \in \{a_r, a_r + 1, ..., b_r\}$, $i (\neq j) \in \{a_r, a_r+1, ..., b_r\}$, $r = 1, 2, ..., l$) be sufficiently small (as described in Lemma 4.1). Then, subject to the control laws as indicated in Lemma 4.1, the subsystems of the large scale system (7) given by (1) are stable in the mean square sense.

Proof: It follows from stability of clusters $\mathcal{S}_r$. The complete proof can be found in [11].

Note that the results of this paper can be extended to also account for the cases where the information about control signals is transmitted with finite power.

REFERENCES

[1] B. Bamieh, F. Paganini, and M. A. Dahleh (2002), Distributed control of spatially invariant systems, *IEEE Transactions on Automatic Control*, 47(7), pp. 1091-1107.

[2] S. Tatikonda (2003). Some scaling properties of large distributed control systems. *Proceedings of 42nd IEEE Conference on Decision and Control*, Hawaii, pp. 3142-3147.

[3] G. N. Nair, R. J. Evans and P. E. Caines (2004). Stabilising decentralised linear systems under data rate constraints. *Proceedings of 43rd IEEE Conference on Decision and Control*, Atlandis, Paradise Island, Bahamas, pp. 3992-3997.

[4] K. Li and J. Baillieul (2005). Problems in decentralized sensor-actuator networks. *Proceedings of 44th IEEE Conference on Decision and Control*, Seville, pp. 3207-3212.

[5] Alexey S. Matveev and Andrey V. Savkin (2005). Decentralized stabilization of linear system via limited capacity communication networks. *Proceedings of 44th IEEE Conference on Decision and Control*, Seville, pp. 1155-1161.

[6] S. Yuksel and T. Basar (2007). Information theoretic study of the signaling problem in decentralized stabilization of noisy linear systems. *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, pp. 328-333.

[7] M. Gastper, B. Rimoldi and M. Vetterli (2003). To code, or not to code: lossy source-channel communication revisited. *IEEE Transactions on Information Theory*, 49(5), pp. 1147-1158.

[8] J. C. F. Li, S. Dey and J. Evans (2008). Maximal lifetime power and rate distortion allocation for wireless sensor systems with data distortion constraints. *IEEE Trans. on Signal Processing*, 56(5), pp. 2076-2090.

[9] M. E. Sezer and D. D. Siljak (1986). Nested ϵ-decomposition and clustering of complex systems. *Automatica*, 22(3), pp. 321-331.

[10] P. E. Caines (1988). Linear stochastic systems. *John Wiley and Sons*.

[11] A. Farhadi and N. U. Ahmed (2009). Stability and reliable data reconstruction of distributed control and communication systems subject to capacity constraints. Submitted to *Automatica*, available on line at www.site.uottawa.ca/∼afarhadi

[12] S. Tatikonda, A. Sahai, and S. Mitter (2004). Stochastic linear control over a communication channel. *IEEE Transactions on Automatic Control*, 49(9), pp. 1549-1561.

[13] Toby Berger (1971). Rate distortion theory: A mathematical basis for data compression. *Prentice-Hall*.

[14] K. Shoarinejad, I. Kanellakopoulos, J. D. Wolfe and J. L. Speyer (1999). A two-station decentralized LQG problem with noisy communication. *Proceedings of the 38th Conference on Decision and Control*, Arizona, pp. 4953-4958.