Preprints of the 20th World Congress
The International Federation of Automatic Control
Toulouse, France, July 9-14, 2017

Separation in Coupled Event-Triggered Networked Control Systems ⋆

Precious Ugo Abara and Sandra Hirche

Chair of Information-oriented Control, Technische Universität München, Arcisstraße 21, D-80290 München, Germany
{ugoabara, hirche}@tum.de

Abstract: Information constraints, for example due to resource limitations, pose a significant challenge for the control of geographically distributed systems. In this work, we investigate the separation of estimation and control in event-triggered discrete-time stochastic systems which are physically interconnected. The information pattern and resource constraints on the system can undermine the separation property and lead to a problem which is difficult to solve. For the class of problems considered, we investigate the constraints and information pattern between the controllers such that separation can be applied in solving the optimal control problem.

Keywords: Control under communication constraints, multi-agent systems, separation principle, certainty equivalence, networked control systems, cyber-physical systems, event-triggered scheduling.

1. INTRODUCTION

Control of large-scale systems over communication networks is an important research topic with many application domains such as infrastructure systems and smart grids. One of the core challenges in networked cyber-physical systems is the control under resource constraints in the cyber part, e.g. communication constraints due to energy limits of battery-driven wireless sensors. Event-triggered control schemes, see Bernhardsson and Åström (1999), Molin and Hirche (2013) and Heemels et al. (2012), that is, control schemes where information is only transmitted when necessary, have been proven to perform particularly well if resource constraints are present. However, along with the benefits of event-triggered schemes several issues emerge in the analysis and design of such systems that are not present in time-triggered control. For example, the optimal design of control and event-triggered policies of a single loop is a challenging issue as two distributed decision makers are involved (Molin and Hirche (2013)). Even for simple systems with linear dynamics and quadratic cost the optimal solution is hard to find when the information pattern is distributed (Witsenhausen (1968)). What makes these problems challenging is that standard techniques of stochastic optimal control theory are not directly applicable. In the time-triggered sampling approach there is no additional information contained in the timing variables since these are available beforehand. Meanwhile, in the event-triggered strategy particular attention has to be paid to the potential complication in the control design due to the dual effect of control. The control is said to have a dual effect when it affects the system's state evolution and it can probe the system to reduce its state uncertainty, i.e., improve the estimation, which ultimately helps to achieve the control objectives (Bar-Shalom and Tse (1974); Ramesh et al. (2011)). In this paper we will investigate the separation property for a class of interconnected systems under event-triggered control. In particular, we focus on physically coupled plants where the interconnections are represented by a directed acyclic graph (DAG). In addition, each plant can communicate with the controllers over a shared network (networked control systems (NCS)). This kind of interconnections is used to model systems with some kind of hierarchy, e.g. vehicle platoons (Al Alam et al. (2011)) and water distribution networks (Perelman and Ostfeld (2011)).
The main contribution of this paper is the derivation of a necessary and sufficient condition on the event-triggered distributed control law such that the separation property can be guaranteed. Counter-intuitively, we show that if there is too much data sharing in the system the problem becomes hard to solve, and in some cases even infeasible due to the dimension of the network.

The remainder of this paper is organized as follows: In section 2, we introduce the system model and problem statement. The main results are presented in section 3.

Notation

In this paper, the operator $(\cdot)^\top$ denotes the transpose and $tr(\cdot)$ is the trace operator. The Expectation operator is denoted by $\mathrm{E}[\cdot]$ and the conditional expectation is denoted by $\mathrm{E}[\cdot|\cdot]$. The Euclidean norm is denoted by $\|\cdot\|_2$. A state vector is denoted by with superscript i specifying the control loop and the subscript k indicating the the time-step. A sequence given between two time instants t, k, $t > k$ is denoted by $\{s\}_{t:k} = \{s_t, s_{t+1}, \ldots, s_k\}$. The vector $\mathbf{1}^\top = (1, \ldots, 1)$ where the length will be clear from the context.

⋆ This project has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under grant agreement No 674875.

2. PROBLEM STATEMENT AND PRELIMINARIES

Consider a NCS composed of N LTI subsystems which are physically interconnected, see Fig. 1 for illustration. The interconnections are represented through a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ such that node $i \in \mathcal{V}$ for each subsystem $i \in \{1, \ldots, N\}$ and edge $(j, i) \in \mathcal{E}$ if the dynamics of node i is influenced directly by dynamics of node j. In addition to the process $\mathcal{P}^i$, it also consists of a control unit $\mathcal{C}^i$ and an event-based scheduler $\mathcal{S}^i$. The dynamics of the i-th subsystem is given by the stochastic difference equation

$$x_{k+1}^i = A_i x_k^i + B_i u_k^i + \sum_{j \in \mathcal{N}_i} A_{ij} x_k^j + w_k^i \tag{1}$$

where $x_k^i \in \mathbb{R}^{n_i}$ is the state of the subsystem, $A_i \in \mathbb{R}^{n_i \times n_i}$, $A_{ij} \in \mathbb{R}^{n_i \times n_j}$, $B_i \in \mathbb{R}^{n_i \times m_i}$, $u_k^i \in \mathbb{R}^{m_i}$ is the control signal. The initial state x_0^i is a random variable with finite mean and covariance and $w_k^i \in \mathbb{R}^{n_i}$ is a zero-mean i.i.d. Gaussian noise with covariance matrix Σ_w^i which is statistically independent of x_0^i for each k. The set $\mathcal{N}_i \subset \mathcal{V}$ is such that node $j \in \mathcal{N}_i$ if $(j, i) \in \mathcal{E}$, i.e. there is a direct edge $j \to i$. Furthermore, each subsystem $i \in \{1, \ldots, N\}$ has a local cost function J_i defined by

$$J_i = \mathrm{E}\left[\sum_{k=0}^{T-1} {x_k^i}^\top Q_i x_k^i + {u_k^i}^\top R_i u_k^i + {x_T^i}^\top S_i x_T^i + \lambda \delta_k^i\right], \tag{2}$$

where S_i, Q_i are semi-positive definite and R_i is positive definite and communication penalty $\lambda > 0$. We assume that the controller $\mathcal{C}^i$ has no regular access to the systems' states but instead a measurement of the state is transmitted to $\mathcal{C}^i$ in an event-triggered manner. The state transmission depends on the scheduling (event-triggering) δ_k^i of the local event-trigger $\mathcal{S}^i$, $i \in \{1, \ldots, N\}$ and this is defined by

$$\delta_k^i = \begin{cases} 1 & \text{attempt to transmit } x_k^i, \\ 0 & \text{no transmission.} \end{cases} \tag{3}$$

Whether a transmission attempt is successful or not depends on the current network load, i.e. transmission requests from the other subsystems. Therefore we define the network manager variable q_k^i for every $i \in \{1, \ldots, N\}$ as

$$q_k^i(\delta_k^1, \ldots, \delta_k^N) = \begin{cases} 1 & x_k^i \text{ allow transmission,} \\ 0 & x_k^i \text{ blocked.} \end{cases} \tag{4}$$

We assume that q_k^i is function of all the triggering variables $\delta_k^j, j \in \{\ldots, N\}$ and q_k^i to be independent of q_t^j for all i, j and $k \neq t$. Therefore the shared network introduces additional coupling. In consequence, the state measurement of the physical process $\mathcal{P}^i$ transmitted by the scheduler $\mathcal{S}^i$ to $\mathcal{C}^i$ at time instant k is

$$z_k^i = \begin{cases} x_k^i & \text{if } \delta_k^i = 1 \text{ and } q_k^i = 1, \\ \emptyset & \text{otherwise.} \end{cases}$$

Furthermore, we assume that some of the control units $\mathcal{C}^i$ can exchange information, i.e. are interconnected. Therefore, we introduce the control network's graph $\mathcal{G}^\mathcal{C} = \left(\mathcal{V}^\mathcal{C}, \mathcal{E}^\mathcal{C}\right)$, which has a node $i \in \mathcal{V}^\mathcal{C} = \{1, \ldots, N\}$ for each physical process $\mathcal{P}^i$, and define $\mathcal{N}_i^\mathcal{C} = \{j | (j, i) \in \mathcal{E}^\mathcal{C}\}$ to denote the set of direct incoming neighbors of a controller $\mathcal{C}^i$. We also assume that each control node i has access to the statistics of the primitive random variables and model structure of incoming nodes, i.e. (A_j, B_j, A_{ji}) for every node $j \in \mathcal{N}_i^\mathcal{C}$. Fig 2 illustrates an example of a control network.

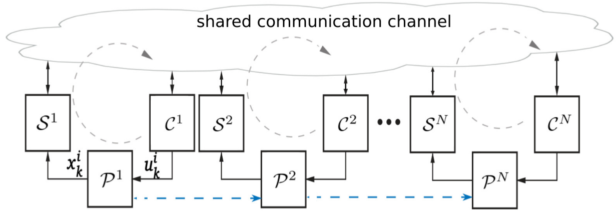

Fig. 1. Example of interconnected NCS. Each control loop is closed over the network, and the dashed interconnections between the processes $\mathcal{P}^i$ represents the physical coupling.

In the following sections we discuss the constraints on the information flow in the set of connections of the control network. We define the causal mapping

$$u_k^i = \gamma_k^i(\mathcal{I}_k^i), \qquad i = 1, \ldots, N. \tag{5}$$

as a measurable function of the information available $\mathcal{I}_k^i$ at controller $\mathcal{C}^i$ at the beginning of each period k. It is defined by

$$\mathcal{I}_k^i = \{\mathcal{I}_{k-1}^i, u_{k-1}^i, z_k^i, \delta_k^i, q_k^i\} \cup_{j \in \mathcal{N}_i^\mathcal{C}} \{\mathcal{I}_{k-1}^j\},$$
$$\mathcal{I}_0^i = \{z_0^i, \delta_0^i, q_0^i\}, \quad \forall i \in \{1, \ldots, N\} \tag{6}$$

Remark 1. In equation (6) we are assuming that neighboring control units are connected through a network without resource constraints. This might be the case for example if we have wireless sensors for which information transmission from the sensors are costly, but from one control unit to another one is cheap. Furthermore, from (3), (4) and (6) we are assuming an instantaneous acknowledgement channel since δ_k^i and q_k^i are both part of the information set $\mathcal{I}_k^i$ at time instant k.

Moreover, we observe that the scheduler and network arbitration mechanism introduces an estimation error in the states of the subsystems defined as

$$\varepsilon_k^i = \begin{cases} 0, & \text{if } \delta_k^i = 1 \text{ and } q_k^i = 1, \\ e_k^i & \text{otherwise.} \end{cases} \tag{7}$$

where e_k^i is the one-step ahead estimation error defined as

$$e_k^i = x_k^i - \mathrm{E}[x_k^i | \mathcal{I}_{k-1}^i]. \tag{8}$$

We assume the causal mapping δ_k^i to be a measurable function of the error

$$\delta_k^i = \psi_k^i(e_k^i), \qquad i = 1, \ldots, N. \tag{9}$$

Remark 2. In general (9) depends on the control law (5). However, we will show that under some assumptions e_k^i is independent of the control law being used.

Consider the global aggregated system

$$x_{k+1} = A x_k + B u_k + w_k \tag{10}$$

where the state $x_k = \left({x_k^1}^\top, \ldots, {x_k^N}^\top\right)^\top \in \mathbb{R}^n$, $w_k = \left({w_k^1}^\top, \ldots, {w_k^N}^\top\right)^\top \in \mathbb{R}^n$ are the stacked vectors. From the DAG assumption on $\mathcal{G}$, $A \in \mathbb{R}^{n \times n}$ can be written as lower block triangular matrix by appropriate permutation

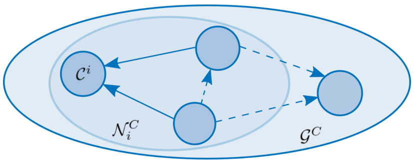

Fig. 2. Direct incoming neighborhoods of $\mathcal{C}^i$ in the control network. The dashed-lines do not have any role.

of indexes and $B \in \mathbb{R}^{n \times m}$ is a block-diagonal matrix and $n = \sum_{i=1}^{N} n_i$, $m = \sum_{i=1}^{N} m_i$. Before defining our minimization problem we need the following concept of team layer.

Definition 1. (Team layer). Consider the set of vertices $\mathcal{V} = \{1, \ldots, N\}$. A team layer is defined as the set of nodes, Υ_k, such that $i \in \Upsilon_k$ if its incoming nodes in the physical layer $\mathcal{N}_i$ and in the control layer $\mathcal{N}_i^{\mathcal{C}}$ are part of previous team layer $\Upsilon_0, \ldots, \Upsilon_{k-1}$. Formally,

$$\Upsilon_0 = \left\{ i \in \{1, \ldots, N\} \middle| \mathcal{N}_i^{\mathcal{C}} \cup \mathcal{N}_i = \emptyset \right\},$$
$$\Upsilon_k = \left\{ i \in \{1, \ldots, N\} \middle| \mathcal{N}_i^{\mathcal{C}} \cup \mathcal{N}_i \subsetneq \bigcup_{j=0,\ldots,k-1} \Upsilon_j \right\}. \quad (11)$$

Note that Υ_0 is the set of all nodes with no incoming direct neighbours in either of the two graphs $\mathcal{G}$ and $\mathcal{G}^{\mathcal{C}}$, i.e., if $i \in \Upsilon_0$ then node i is a leader. Every node in the network is allowed to be part of just one team layer. See Fig. 3 for example. To every team layer we assign the following minimization problem

$$J_{\Upsilon_k} = \min_{u_{\Upsilon_k}} \frac{1}{T} \sum_{i \in \Upsilon_k} J_i, \quad (12)$$

where $\Upsilon_k = \left\{ i_1, \ldots, i_{|\Upsilon_k|} \right\}$ and $\{\cdot\}_{\Upsilon_k} = \left(\cdot_{i_1}, \ldots, \cdot_{i_{|\Upsilon_k|}} \right)$. Moreover, J_{Υ_k} represents the team cost. For each subsystem $i \in \{1, \ldots, N\}$ we assume (A_i, B_i) stabilizable and $(A_i, Q_i^{\frac{1}{2}})$ detectable, with $Q_i = (Q_i^{\frac{1}{2}})^\top Q_i^{\frac{1}{2}}$. We want to minimize the following global cost with communication penalty

$$\min_{\delta, u} J_S = \sum_{k=0}^{M} J_{\Upsilon_k}, \quad (13)$$

where $(M+1)$ is the number of team layers, δ is such that $\delta = \left(\{\delta^1\}_{0:T-1}, \ldots, \{\delta^N\}_{0:T-1} \right)$, similarly u is given by $u = \left(\{u^1\}_{0:T-1}, \ldots, \{u^N\}_{0:T-1} \right)$ and J_S represents the social cost. Let $\zeta(i,j) \in \mathcal{G}^{\mathcal{C}} \cup \mathcal{G}$ be a path from node i to node j and $|\zeta(i,j)| \in \mathbb{N}$ be the length of the path (number of edges). We note that $M = \max \left\{ |\zeta(l,f)|, \ \zeta(l,f) \in \mathcal{G}^{\mathcal{C}} \cup \mathcal{G} \right\}$. Moreover, in the remaining of this paper we will assume $T > M + 1$. From the previous equations it is clear that the controllers (5), as a function of the available information, depend on the event-triggered scheduling law and on the network arbitration mechanism. The graph $\mathcal{G}^{\mathcal{C}}$ represents the information exchange between the neighbouring controllers and can decide on the feasibility of the problem defined in (13). The aim of this first part of the paper of is to give a characterization of the control layer $\mathcal{G}^{\mathcal{C}}$ such that the separation property holds.

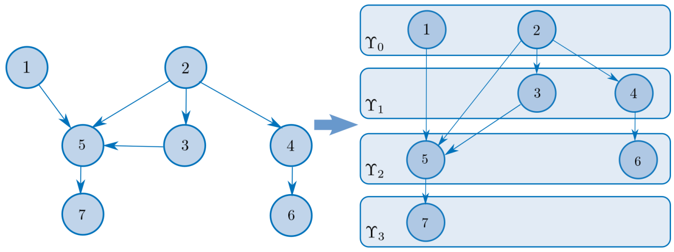

Fig. 3. For simplicity we impose $\mathcal{G}^{\mathcal{C}} = \mathcal{G}$. Based on equation (11), the network is decomposed into 4 layers. Node 1 and 2 are the leaders of the system.

2.1 *Dual effect, Certainty Equivalence and Separation*

We say that the separation property holds if the optimal control can be decomposed into two parts: i) an estimator, which uses the information available defined in (6) to generate the estimation of the states and ii) a state-based controller which generates the u_k^i defined in (5) given the estimate of the state. For this to be achievable there should be no dual effect of control. Let

$$\Omega_0^i = \{x_0^i\}$$
$$\Omega_k^i = \left\{ \Omega_{k-1}^i, w_{\tau_k^i}^i \right\} \cup_{j \in \mathcal{N}_i^{\mathcal{C}}} \Omega_{k-1}^j \quad (14)$$

for $k = 1, \ldots, T$, be the set of all primitive random variables influencing the control unit $\mathcal{C}^i$ at time instant k. The time variable $\tau_k^i < k$ is the latest successful transmission less than k. A formal definition of dual effect, certainty equivalence and separation now follows (Bar-Shalom and Tse (1974); Ramesh et al. (2011)).

Definition 2. (Dual Effect). A control signal is said to have no dual effect of order $r \geq 2$ if

$$\mathrm{E}\left[\mathcal{M}_k^r \middle| \mathcal{I}_k^i \right] = \mathrm{E}\left[\mathcal{M}_k^r \middle| \Omega_k^i \right],$$

where $\mathcal{M}_k^r = \mathrm{E}\left[\left(\varepsilon_k^i \right)^r \middle| \mathcal{I}_k^i \right]$ is the r-*th* central moment of x_k^i conditioned on the information available at the controller and the estimation error is defined in (7).

From the definition we note that $\mathcal{M}_k^r$ must not be function of the past controls in order for the control signal to have no dual effect.

Definition 3. (Certainty Equivalence). In a control problem the certainty equivalence property is said to hold if the closed-loop optimal controller has the same form as the deterministic optimal controller, with the states x_k^i replaced by their estimates $\hat{x}_k^i = \mathrm{E}\left[x_k^i | \mathcal{I}_k^i \right]$, i.e,

$$u_k^{i\star} = \begin{cases} \phi_k^i(x_k^1, \ldots, x_k^N) & \text{deterministic} \\ \phi_k^i(\mathrm{E}[x_k^1|\mathcal{I}_k^i], \ldots, \mathrm{E}[x_k^N|\mathcal{I}_k^i]) & \text{stochastic} \end{cases}$$

Definition 4. (Separation property). The closed-loop optimal controller has the separation property if it depends on the data only via the estimates $\hat{x}_k^i$, i.e., $u_k^{i\star} = \eta_k^i(\mathrm{E}[x_k^i|\mathcal{I}_k^i], \ldots, \mathrm{E}[x_k^N|\mathcal{I}_k^N])$, where the function η_k^i can be different from ϕ_k^i obtained in the deterministic case.

It can be noted that separation is a weaker property than certainty equivalence since the latter implies separation but not the other way around.

3. SEPARATION PRINCIPLE

We can now give the following lemma.

Lemma 1. Consider the physically interconnected system in (1), the policies (5) and (9), the information structure (6), the estimation error (7). Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be DAG and let $\mathcal{G}^{\mathcal{C}} = \left(\mathcal{V}^{\mathcal{C}}, \mathcal{E}^{\mathcal{C}}\right)$ be the control graph. Let $M+1$ be the number of team layers and assume $T > M+1$. The estimation error (7) is independent of the control law if and only if $\mathcal{G}^{\mathcal{C}}$ has no conflicting paths, i.e., path $\zeta(i,j) \in \mathcal{G}$ imply that $\zeta(j,i) \notin \mathcal{G}^{\mathcal{C}}$.

Proof. Let us start showing the necessity of the lemma. If the control layer has conflicting paths we show that the network-induced estimation error (7) of a subsystem $i \in \{1, \ldots, N\}$ cannot be written as a function of the primitive random variables of subsystems $j \in \mathcal{N}_i \cup \{i\}$, but the error depends on the controls being applied.
Without loss of generality, since $\mathcal{G}$ is a DAG graph, we can assume l to be a leader and f a follower, i.e., $\mathcal{N}_l = \emptyset$ and $l \in \mathcal{N}_f$. With a slight abuse of notation, let us consider a forced and unforced subsystem l as given by the following equations (Bertsekas (2005))

$$\begin{cases} x^l_{k+1} & = A_l x^l_k + B_l u^l_k + w^l_k \\ z^l_k & = \delta^l_k q^l_k x^l_k \end{cases} \tag{15}$$

and the unforced system is

$$\begin{cases} \bar{x}^l_{k+1} & = A_l \bar{x}^l_k + \bar{w}^l_k \\ \bar{z}^l_k & = \bar{\delta}^l_k \bar{q}^l_k \bar{x}^l_k \end{cases} \tag{16}$$

We consider the evolution of these two systems when they have the same initial conditions and system disturbances, i.e.,

$$\bar{x}^l_0 = x^l_0, \qquad \bar{w}^l_k = w^l_k, \qquad k = 0, \ldots, T-1$$

Linearity allows to rewrite the systems in the following form

$$x^l_k = F^l_k x^l_0 + G^l_k U^{k-1}_l + H^l_k W^{k-1},$$
$$\bar{x}^l_k = F^l_k x^l_0 + H^l_k W^{k-1}.$$

where F^l_k, G^l_k and H^l_k are matrices of appropriate dimensions obtained from A_l and B_l, and $U^k_l = \left(u^l_0, \ldots, u^l_k\right)$. Now from our assumptions, the edge $(l,f) \in \mathcal{E}$ and we also us assume $(f,l) \in \mathcal{E}^C$, i.e., $\mathcal{G}^C$ has a conflicting edge $f \to l$. Since the control layer $\mathcal{G}^C$ has a conflicting edge, the information available at controller l and f at time-step k are

$$\mathcal{I}^l_k = \mathcal{I}^l_{k-1} \cup \{z^l_k\} \cup \{u^l_{k-1}\} \cup \{\mathcal{I}^f_{k-1}\},$$
$$\mathcal{I}^f_k = \mathcal{I}^f_{k-1} \cup \{z^f_k\} \cup \{u^f_{k-1}\} \cup_{i \in \mathcal{N}^C_f} \{\mathcal{I}^i_{k-1}\}, \tag{17}$$

where $\mathcal{I}^l_0 = \{z^l_0, \delta^l_0, q^l_0\}$ and $\mathcal{I}^f_0 = \{z^f_0, \delta^f_0, q^f_0\}$. Clearly $U^{k-1}_l = \mathrm{E}[U^{k-1}_l | \mathcal{I}^l_k]$ since U^{k-1}_l is composed of measurable functions of $\mathcal{I}^l_{k-1} \subset \mathcal{I}^l_k$, from (5), so we obtain the following

$$\mathrm{E}[x^l_k | \mathcal{I}^l_k] = F^l_k \mathrm{E}[x^l_0 | \mathcal{I}^l_k] + G^l_k U^{k-1}_l + H^l_k \mathrm{E}[W^{k-1} | \mathcal{I}^l_k],$$
$$\mathrm{E}[\bar{x}^l_k | \mathcal{I}^l_k] = F^l_k \mathrm{E}[x^l_0 | \mathcal{I}^l_k] + H^l_k \mathrm{E}[W^{k-1} | \mathcal{I}^l_k].$$

This yields

$$\varepsilon^l_k = x^l_k - \mathrm{E}[x^l_k | \mathcal{I}^l_k] = \bar{x}^l_k - \mathrm{E}[\bar{x}^l_k | \mathcal{I}^l_k].$$

Let

$$\Theta^i_0 = \Omega^i_0 \cup \left\{q^i_0, \delta^i_0\right\}$$
$$\Theta^i_k = \Omega^i_k \cup \left\{\Theta^i_{k-1}, q^i_k, \delta^i_k\right\} \cup_{j \in \mathcal{N}^C_i} \Theta^j_{k-1} \tag{18}$$

be the set of all random variables influencing the control unit $\mathcal{C}^i$ at time instant k. When (17) holds the leader's state estimate at the controller is different from the one at the scheduler. This is immediately visible if, e.g., $x^f_{k+1} = A_{fl} x^l_k + w^f_k$ since $\mathcal{I}^f_k$ contains measurements of x^l_k independently of the scheduling policy δ^l_t of the leader, e.g. $\delta^l_t q^l_t = 0$ for $t = 1, \ldots, k-1$. Moreover, since the σ-algebra at the controller l is not contained in the σ-algebra at the scheduler l (Molin and Hirche (2013)) a bijective function between the observations z^l_k and z^f_k does not exist and we can conclude that the estimation error ε^l_k depends on the control law.

For sufficiency of the lemma let us assume the control layer $\mathcal{G}^{\mathcal{C}}$ has no conflicting edges or paths. If this is the case then the information structure at controller l and f at time-step k are

$$\mathcal{I}^l_k = \mathcal{I}^l_{k-1} \cup \{z^l_k\} \cup \{u^l_{k-1}\},$$
$$\mathcal{I}^f_k = \mathcal{I}^f_{k-1} \cup \{z^f_k\} \cup \{u^f_{k-1}\} \cup_{i \in \mathcal{N}^C_f} \{\mathcal{I}^i_{k-1}\}. \tag{19}$$

where $\mathcal{I}^l_0 = \{z^l_0, \delta^l_0, q^l_0\}$ and $\mathcal{I}^f_0 = \{z^f_0, \delta^f_0, q^f_0\}$. When (19) holds then $\sigma(\mathcal{I}^l_k)$ is generated by the random variables in Θ^l_k, with $\mathcal{N}^C_l = \emptyset$. From (9) and from q^i_k being a measurable function of δ^j_k, $j = \{1, \ldots, N\}$, for every k there is a function M_k such that Bertsekas (2005)

$$x^l_k - \mathrm{E}[x^l_k | \mathcal{I}^l_k] = M_k(\Omega^l_k).$$

independent of the control law being used.
The proof is concluded since for every edge $(i,j) \in \mathcal{E}$ we can assume i to be a leader and j a follower and iterate over the control graph. Even if i is not a global leader but just a leader with respect to j this makes no difference since the information available at subsystem j will still contain $\mathcal{I}^i_{k-1}$. The case where a node i has no outgoing edge the conclusion is straightforward.

Remark 3. Notice that $\mathcal{E}^C$ can have more edges than than $\mathcal{E}$ or even $\mathcal{E}^C = \emptyset$ (decentralized control) as long as there are no conflicting edges. Moreover from the proof of Lemma 1 it is straightforward to see that under the same condition the variable e^i_k is independent of the control signal γ^i_k. Thus δ^i_k defined in (8) is also independent of γ^i_k. Furthermore, if the lemma holds, the prediction error e^i_k is not a function of the applied control $u^i_{0:k-1}$ which yields independency at a global level since it holds for every $i \in \{1, \ldots, N\}$.

Remark 4. Under the assumptions of Lemma 1, from the structure of the information available at each $\mathcal{C}^i$ we are assuming that the communication within the control network occurs as least as fast as information travels through the physical layer. This imply that the information struture is partially nested which in turn implies that affine control laws are optimal if the primitive random variables are Gaussian, the observation function linear and the cost quadratic (Ho et al. (1972), Lamperski and Doyle (2012)). This also imply that a person-by-person optimal is a global optimal. Therefore minimizing first with respect to the admissible control policies and then to the triggering policies yields a global optimum, i.e., the controller law given by the following theorems is a dominating class of policies (Molin and Hirche (2013)).

3.1 Certainty Equivalence and Separation Principle

The following is a well-known result of LQG optimization.

Theorem 1. (Bar-Shalom and Tse (1974)). The optimal stochastic control u_k^i for the system (1) and quadratic cost $J_i|_{\lambda=0}$ as defined in (2) has the certainty equivalence property for all $Q_i \geq 0$, $R_i > 0$ if and only if the control has no dual effect of second order, i.e., the expected future uncertainty is not affected by the control with probability one.

We are now able to provide the main result on separation.

Theorem 2. Consider the physically interconnected system in (1), the policies (5) and (9), the information structure (6), the estimation error (7). Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be DAG and let $\mathcal{G}^{\mathcal{C}} = \left(\mathcal{V}^{\mathcal{C}}, \mathcal{E}^{\mathcal{C}}\right)$ be the control graph. Let $M+1$ be the number of team layers and assume $T > M+1$. Then the optimal stochastic control for the system (1) and cost J_S as defined in (13) has the separation property for all $Q_i \geq 0$, $R_i > 0$ if and only if the control network $\mathcal{G}^C$ is such that $\mathcal{G}^C$ has no conflicting paths, i.e., path $\zeta(i,j) \in \mathcal{G}$ imply that $\zeta(j,i) \notin \mathcal{G}^{\mathcal{C}}$.

Proof. Under our assumptions, Lemma 1 is valid. Therefore, from *Remark 2* and *Remark 4*, since ψ_k^i is fixed to be a function of the error (7), the output δ_k^i is a random variable that can be described by a function of the primitive random variables and it is independent of the control law γ_k^i. This implies that $\mathrm{E}[\sum_{i=1}^N \sum_{k=0}^{T-1} \delta_k^i]$ is a constant for fixed ψ_k^i. Thus the resulting objective function J_S is purely quadratic and we can apply Theorem 1.

The optimization of the system is performed sequentially, from the top team layer Υ_0 to the bottom layer Υ_M (Siljak (2011)). Once the optimal control laws for Υ_{h-1} is found and implemented, i.e., $u_k^{i\star}$, $i \in \Upsilon_{0:h-1}$, the optimization for Υ_h is carried out until we reach the bottom layer. The process takes advantage of the lower block triangular structure of the system. Starting from layer Υ_0 we want to find the optimal control that minimizes the first term of (13). Since $\mathcal{N}_{\Upsilon_0}^{\mathcal{C}} = \emptyset$, the dynamics of the subsystem at this layer is:

$$x_{k+1}^i = A_i x_k^i + B_i u_k^i + w_k^i,\ i \in \Upsilon_0 . \tag{20}$$

For a general Υ_h the dynamics of the subsystems it contains can be written as

$$x_{k+1}^{\Upsilon_h} = A_{\Upsilon_h} x_k^{\Upsilon_h} + B_{\Upsilon_h} u_k^{\Upsilon_h} + w_k^{\Upsilon_h}. \tag{21}$$

where A_{Υ_h} can is still be assumed to be lower block triangular. The conditional expectation of x_k^i given $\mathcal{I}_k^i$ is

$$\mathrm{E}[x_k^i|\mathcal{I}_k^i] = \begin{cases} x_k^i, & \text{if } q_k^i\delta_k^i = 1, \\ \mathrm{E}[x_k^i|\mathcal{J}_{k-1}^i], & \text{if } q_k^i\delta_k^i = 0. \end{cases}$$

where $\mathcal{J}_{k-1}^i = \mathcal{I}_{k-1}^i \cup_{j\in\mathcal{N}_i} \mathcal{I}_{k-1}^j$. Clearly $\mathrm{E}[x_k^i|\mathcal{J}_{k-1}^i] = \mathrm{E}[x_k^i|\mathcal{I}_{k-1}^i]$ since we have assumed absence of conflicting paths or edges, and u_k^i to be a measurable function of $\mathcal{I}_{k-1}^i$. Furthermore, we can write

$$\begin{aligned} \mathrm{E}[x_{k+1}^i|\mathcal{I}_k^i] &= \mathrm{E}[x_{k+1}^i|\mathcal{J}_{k-1}^i, u_k^i, z_k^i] = \mathrm{E}[x_{k+1}^i|\mathcal{J}_{k-1}^i, u_k^i] + \\ &+ q_k^i\delta_k^i\left(\mathrm{E}[x_{k+1}^i|x_k^{\bar{\mathcal{N}}_i} - \bar{x}_k^{\bar{\mathcal{N}}_i}] - \mathrm{E}[x_{k+1}^i]\right) = \\ &= \mathrm{E}[A_i x_k^i + B_i u_k^i + \textstyle\sum_{j\in\mathcal{N}_i} A_{ij} x_k^j|\mathcal{J}_{k-1}^i, u_k^i] + q_k^i\delta_k^i K_k^{i\star} e_k^{\bar{\mathcal{N}}_i} \end{aligned}$$

where $\bar{\mathcal{N}_i} = \mathcal{N}_i \cup i$, and, with a slight abuse of notation, $x_k^{\bar{\mathcal{N}}_i} = \{x_k^q, q \in \bar{\mathcal{N}_i}\}$. Furthermore, if we define

$$\bar{x}_k = \left(\mathrm{E}[x_k^1|\mathcal{I}_{k-1}^1]^\top, \ldots, \mathrm{E}[x_k^N|\mathcal{I}_{k-1}^N]^\top\right)^\top . \tag{22}$$

we obtain

$$\bar{x}_{k+1}^i = A_i\bar{x}_k^i + B_i u_k + \sum_{j\in\mathcal{N}_i} A_{ij}\bar{x}_k^j + q_k^i\delta_k^i K_k^{i\star} e_k^{\bar{\mathcal{N}}_i} \tag{23}$$

where $K_k^{i\star}$ is the appropriate sparse optimal Kalman-like filter gain since q_k^i and q_t^j are assumed to be independent for all i, j and $k \neq t$ (Sinopoli et al. (2004)). Let $\Upsilon_{0:h} = \cup_{k=0}^h \Upsilon_k$ and $\mathcal{Q}_k^{\Upsilon_h}(\delta_k, q_k) = \mathrm{diag}(\delta_k^i q_k^i I_{n_i})$, I_{n_i} where is the identity matrix of dimension n_i as subsystem $i \in \Upsilon_h$. Under Lemma 1, the estimation error (8) is independent of the control law and such that $\mathrm{E}[e_k^{\Upsilon_h}\bar{x}_k^{\Upsilon_h}] = 0$. Furthermore, since the information pattern is partially nested, the optimal control is linear and has the form (similar to Kurtaran and Sivan (1974))

$$u_k^\star = F_k^\star z_k + G_k^\star \mathrm{E}[x_k|\mathcal{J}_{k-1}]. \tag{24}$$

where $F_k^\star$ is a block diagonal optimal gain, $G_k^\star$ is lower triangular and z_k^i is assumed to be zero if no transmission occurred. To show this, we define the invertible transformed system input $\bar{u}_k$ as

$$\bar{u}_k^{\Upsilon_h} = u_k^{\Upsilon_h} - \mathcal{Q}_k^{\Upsilon_h}(\delta_k, q_k) F_k^{\Upsilon_h\star} e_k^{\Upsilon_h}, \tag{25}$$

where $F_k^{\Upsilon_h\star} = \mathrm{diag}(F_k^{i_1\star}, \ldots, F_k^{i_{|\Upsilon_h|}\star})$ and $i_j \in \Upsilon_h$. Inserting (25) into (23) gives

$$\begin{cases} \bar{x}_{k+1}^{\Upsilon_h} &= A_{\Upsilon_h}\bar{x}_k^{\Upsilon_h} + B_{\Upsilon_h} u_k^{\Upsilon_h} + v_k^{\Upsilon_h}(\delta_k, q_k) \\ \bar{x}_0^{\Upsilon_h} &= \mathrm{E}[x_0^{\Upsilon_h}] \end{cases} \tag{26}$$

Given a realization of δ_k and q_k, it can be verified that $v_k^{\Upsilon_h}$ is a white zero-mean noise independent of $\bar{x}_k^{\Upsilon_h}$. To simplify our notation, let $\bar{x}_k^h = \bar{x}_k^{\Upsilon_h}$. Now, if we assume to have found and implemented all the control laws up to layer Υ_{h-1}, i.e. $u_k^{i\star}$, $i \in \Upsilon_{0:h-1}$, we note that the global information available at this layer at time instant k is $\mathcal{I}_k^{\Upsilon_h} = \mathcal{I}_k^{\Upsilon_{h-1}} \cup_{j\in\Upsilon_h} \mathcal{I}_k^j$. First, we should note that $\mathrm{E}[x_k^i|\mathcal{I}_k^{\Upsilon_l}] = \mathrm{E}[x_k^i|\mathcal{I}_k^i]$ for every $i \in \Upsilon_l$. This is due to the fact that they are no edges between the subsystems in a same layer. Therefore, the subsystems in Υ_l are decoupled and can apply the standard LQG to each. Using (25) in the argument of our cost functional (12) we have

$$\begin{aligned} &\mathrm{E}\left[\sum_{k=0}^{T-1} x_k^{h\top} Q_{\Upsilon_h} x_k^h + u_k^{h\top} R_{\Upsilon_h} u_k^h + x_T^{h\top} S_{T,\Upsilon_h} x_T^h + \lambda \mathbf{1}^\top \delta_k^h\right] \\ &= \mathrm{E}\left[\sum_{k=0}^{T-1}\left(\bar{x}_k^{h\top} Q_{\Upsilon_h}\bar{x}_k^h + \bar{u}_k^{h\top} R_{\Upsilon_h}\bar{u}_k^h\right) + \bar{x}_T^{h\top} S_{T,\Upsilon_h}\bar{x}_T^h\right] + \\ &+ \mathrm{E}\left[\sum_{k=0}^{T-1} e_k^{h\top} Q_{\Upsilon_h} e_k^h + e_{T,h}^\top Q_{T,\Upsilon_h} e_{T,h} + \lambda\mathbf{1}^\top\delta_k^{\Upsilon_h}\right] + \\ &+ \mathrm{E}\left[\sum_{k=0}^{T-1} e_k^{h\top} F_k^{\Upsilon_h\star\top}\mathcal{Q}_k^{\Upsilon_h}(\delta_k, q_k) R_{\Upsilon_h} F_k^{\Upsilon_h\star} e_k^h\right] + \\ &+ 2\mathrm{E}\left[\sum_{k=0}^{T-1} e_k^{h\top}\left(F_k^{\Upsilon_h\star}\right)^\top \mathcal{Q}_k^{\Upsilon_h}(\delta_k, q_k) R\bar{u}_k^h + e_k^{h\top} Q_{\Upsilon_h}\bar{x}_k^h\right]. \end{aligned}$$

It can be observed that $\bar{u}_k^h$ is a linear function of $\mathcal{I}_k^{\Upsilon_h}$ and so independent of e_k^h. Furthermore, with a fixed triggering policy, all the expectations involving e_k, δ_k are constants and our initial minimization problem (12) is equivalent to

$$\min_{\bar{u}^{\Upsilon_h}} \frac{1}{T}\mathrm{E}\left[\sum_{k=0}^{T-1}\left(\bar{x}_k^{h\top} Q_h \bar{x}_k^h + \bar{u}_k^{h\top} R_h \bar{u}_k^h\right) + \bar{x}_T^{h\top} S_h \bar{x}_T^h\right]. \tag{27}$$

It is important to note that after the implementation of u^{Υ_i}, $i = 0, \dots, h-1$, we have to consider the augmented state

$$X_{k+1}^h = \begin{bmatrix} \star & 0 \\ \star & A_{\Upsilon_h} \end{bmatrix} \left[\begin{array}{c} \bar{x}_k^{\Upsilon_0} \\ \vdots \\ \bar{x}_k^{\Upsilon_{h-1}} \\ \hline \bar{x}_k^{\Upsilon_h} \end{array}\right] + \begin{bmatrix} 0 \\ B_{\Upsilon_h} \end{bmatrix} u_k^{\Upsilon_h} + \tilde{v}_k \qquad (28)$$

where $\tilde{v}_k$ is a white zero-mean noise independent of X_k^h. The minimization problem in (27) can be written equivalently in terms of X_k^h with new weighting matrix $\bar{Q}_h = \begin{bmatrix} 0 & 0 \\ 0 & Q_h \end{bmatrix}$. We can solve it as an optimal control problem with global information yielding:

$$\bar{u}_k^{\Upsilon_h} = -H_k^{h\star} X_k^h$$

where $H_k^{h\star}$ is the optimal gain matrix. With the inverse transformation of (25) and iterating this procedure for all layers, we find that the optimal control law for (13) is given by

$$\begin{aligned} u_k^\star &= \mathcal{Q}_k(\delta_k, q_k) F_k^\star x_k - \left(H_k^{h\star} + \mathcal{Q}_k(\delta_k, q_k) F_k^\star\right) \bar{x}_k \\ &= -L_k^{1\star} \hat{x}_k - L_k^{2\star} \mathrm{E}[x_k | \mathcal{J}_{k-1}]. \end{aligned} \qquad (29)$$

where $\hat{x}_k = \left(\mathrm{E}[x_k^1 | \mathcal{I}_k^1]^\top, \dots, \mathrm{E}[x_k^N | \mathcal{I}_k^N]^\top\right)^\top$ and $\mathcal{J}_{k-1} = \cup_{j \in \mathcal{V}^\mathcal{C}} \mathcal{I}_{k-1}^j$. Furthermore, $L_k^{1\star}$ is a diagonal matrix function of q_k and δ_k, while $L_k^{2\star}$ is lower block triangular, not necessary the same structure of A. We see that we can design the estimates and the gains independently, i.e, we have separation.

Example Consider a system described by the physical layer graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and control layer graph $\mathcal{G}^\mathcal{C} = (\mathcal{V}^\mathcal{C}, \mathcal{E}^\mathcal{C})$, where

- $\mathcal{V}^\mathcal{C} = \mathcal{V} = \{1, 2, 3, 4, 5\}$,
- $\mathcal{E} = \{(1, 2), (2, 3), (3, 4), (3, 5)\}$,
- $\mathcal{E}^\mathcal{C} = \{(1, 2), (3, 4), (3, 5), (3, 1)\}$.

Moreover, each plant is assumed to be scalar and satisfy all properties of (1). The interconnections are described by

$$\begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \\ x_{k+1}^5 \end{bmatrix} = \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \star & 0.5 & 0 \\ 0 & 0 & 0 & \star & 0.5 \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \\ x_k^3 \\ x_k^4 \\ x_k^5 \end{bmatrix} + B u_k + w_k$$

We note that $(3, 1) \in \mathcal{E}^\mathcal{C}$, i.e., a conflicting path exist since $\zeta(1, 3) \in \mathcal{G}$. Recalling (6), we note that the control units

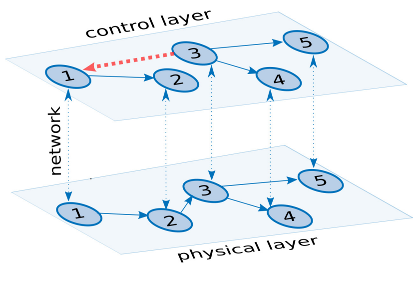

Fig. 4. Design of control layer: The red edge represents a bad link as it violates the assumptions of Lemma 1.

$\mathcal{C}^1$ and $\mathcal{C}^3$ are such that $\mathcal{I}_{k-1}^3 \subset \mathcal{I}_k^1$. Therefore, Lemma 1 is violated and there is no separation. However, if $(3, 1)$ is removed from the control graph the information sharing is such that separation holds. This is illustrated in Fig. 4.

4. CONCLUSION

In this paper we consider the problem of minimizing a LTI system with information and communication constraints. In consequence of the information pattern in the control network, the design of the optimal event-triggered controllers is a hard problem because of dual effect and not classical information pattern. However, we have shown that we can easily solve the problem if the scheduler is error based and partially nestedness between the decision makers is considered, i.e., both controllers and schedulers. Therefore, the optimal controller can be obtained by separating the design of the control gains and the filter.

REFERENCES

Al Alam, A., Gattami, A., and Johansson, K.H. (2011). Suboptimal decentralized controller design for chain structures: Applications to vehicle formations. In *2011 50th IEEE CDC-ECC*, 6894–6900. IEEE.

Bar-Shalom, Y. and Tse, E. (1974). Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Transactions on Automatic Control*, 19(5), 494–500.

Bernhardsson, B. and Åström, K.J. (1999). Comparison of periodic and event based sampling for first-order stochastic systems. In *14th IFAC world congress*.

Bertsekas, D.P. (2005). *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific Belmont, MA.

Heemels, W., Johansson, K.H., and Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *2012 IEEE 51st IEEE CDC*, 3270–3285.

Ho, Y.C. et al. (1972). Team decision theory and information structures in optimal control problems–part i. *IEEE transactions on automatic control*, 17(1), 15–22.

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

Lamperski, A. and Doyle, J.C. (2012). Dynamic programming solutions for decentralized state-feedback lqg problems with communication delays. In *ACC, 2012*.

Molin, A. and Hirche, S. (2013). On the optimality of certainty equivalence for event-triggered control systems. *IEEE TAC*, 58(2), 470–474.

Perelman, L. and Ostfeld, A. (2011). Water-distribution systems simplifications through clustering. *Journal of Water Resources Planning and Management*, 138(3), 218–229.

Ramesh, C., Sandberg, H., Bao, L., and Johansson, K.H. (2011). On the dual effect in state-based scheduling of networked control systems. In *American Control Conference (ACC), 2011*.

Siljak, D.D. (2011). *Decentralized control of complex systems*. Courier Corporation.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., and Sastry, S.S. (2004). Kalman filtering with intermittent observations. *IEEE TAC*.

Witsenhausen, H.S. (1968). A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1).