

A bi-level approach for the design of event-triggered control systems over a shared network

Adam Molin · Sandra Hirche

Received: date / Accepted: date

Abstract An efficient usage of available resources is a substantial requirement for the successful design of networked control systems. Recent results indicate major benefits of event-based control compared to conventional designs, when resources such as communication, energy, and computation, are sparse. This paper considers multiple entities of heterogeneous control systems whose feedback loops are coupled through a common communication medium. The design of the decentralized event-triggering control system is formulated as an average-cost problem that aims at the minimization of a social cost criterion. A state aggregation technique is used to develop a bi-level design method, which divides into a local average-cost problem within every subsystem and a global resource allocation problem assigning optimal transmission rates to every subsystem. Stability conditions are derived that guarantee stochastic stability of the aggregate system. Under these conditions, it is shown that the design approach is asymptotically optimal as the number of subsystems increases.

Keywords Networked control systems · event-triggered control · stochastic optimal control · resource allocation

1 Introduction

In recent years, networked control systems have become a very active field of research. The main characteristic of such systems is the spatial distribu-

This work was supported in part by the German Research Foundation (DFG) within the Priority Program SPP 1305 "Control Theory of Digitally Networked Dynamical Systems".

A. Molin and S. Hirche
Institute of Automatic Control Engineering, Technische Universität München
Arcisstraße 21, D-80290 München, Germany
Tel.: +49-89-289-25723, Fax: +49-89-289-28340
E-mail: {adam.molin, hirche}@tum.de

tion of system components that communicate over a common digital network to accomplish complex control tasks. An efficient usage of the computational and communication resources is mandatory due to limitations in data bandwidth, energy consumption, and computational power. This has stimulated the research on alternative sampling strategies beyond the conventional periodic sampling scheme in order to utilize resources more efficiently Åström and Bernhardsson (2002); Tabuada (2007); Heemels et al (2008); Henningsson et al (2008); Cervin and Henningsson (2008); Lunze and Lehmann (2010); Rabi et al (2012). These works suggest that the use of event-triggered control may reduce the resource consumption significantly compared to periodic control, while achieving the same level of control performance.

While the majority of results studies event-triggered sampling for single-loop control systems, systems with multiple control loops over a shared network have attained only little attention. Exceptions can be found in the works Cervin and Henningsson (2008); Henningsson and Cervin (2010); Blind and Allgöwer (2011a,b) that analyze event-triggered sampling in multi-loop control systems. Depending on the model that represents the resource-constrained communication medium the authors draw different conclusions. Using carrier sense multiple access schemes with priority or randomized arbitration as proposed in Cervin and Henningsson (2008); Henningsson and Cervin (2010), event-triggered sampling for data transmission enhances the control performance significantly compared with periodic transmission schemes. On the other hand, the results in Blind and Allgöwer (2011a,b) suggest that time-triggered sampling outperforms event-triggered sampling for slotted and unslotted ALOHA transmission schemes. The work in Blind and Allgöwer (2011c) analyzes different protocols with event-triggered sampling and it is shown that event-triggered schemes outperform time-triggered schemes for certain protocols. In contrast to the aforementioned work that focuses on the analysis of event-triggered scheduling over a contention-based medium, the motivation of this paper concerns with the optimal synthesis of decentralized control and event-triggering laws. We consider N subsystems whose feedback loops are closed over a contention-based network. The communication model is adopted to the framework in Cervin and Henningsson (2008); Henningsson and Cervin (2010) and assumes the presence of a random arbitration scheme to resolve collisions. Contrary to the aforementioned work that only considers scalar integrator dynamics, the subsystems may be heterogeneous and are modelled as stochastic linear discrete-time systems of arbitrary order. The synthesis of the event-triggered control system is formulated as an average-cost problem, where the cost function is the sum of the individual control costs within each subsystem. By means of time-triggered sampling each subsystem can be considered as an isolated entity which facilitates further analysis. In contrast to the time-triggered case, event-triggered control systems result in a complex interaction between control and communication complicating the analysis and the design of such systems.

The main contribution of this paper is to tackle this challenge by proposing an approximative formulation of the average-cost problem that allows us

to circumvent the consideration of the interaction of control and communication. In the approximative problem formulation, the communication network is considered to be deterministic assigning each subsystem an average transmission rate. By this approach, the optimization problem can be split into two levels: a local optimal control problem that can be solved by dynamic programming and a global resource allocation problem which is a convex optimization problem. Stability conditions are derived that guarantee stochastic stability of the event-triggered control system. Under these conditions, it is shown that the proposed bi-level approach is asymptotically optimal, when the number of users approaches infinity. Numerical simulations indicate that the performance of the event-triggered control system resulting from the bi-level design approach deviates only slightly from that of the optimal solution for a moderate number of subsystems. The theoretical contributions can be summarized as follows: (i) development of a numerically tractable design method for the decentralized event-triggered control system, (ii) construction of stability conditions that are simple to verify, and (iii) derivation of an asymptotic optimality property of the decentralized event-triggered control system. A preliminary version of this work first appeared in the conference paper Molin and Hirche (2011).

The remainder of the paper is structured as follows. Section 2 describes the considered system model and defines the problem statement. The design of the event-triggered control system and its properties are presented in section 3. Section 4 illustrates the results by numerical simulations.

Notation. In this paper, A^T is the transpose of a matrix A and $\text{tr}[A]$ is the trace operator. The Euclidean vector norm and its induced matrix norm are denoted by $\|\cdot\|_2$. The truncated sequence up to time K of a signal x_k , $k \geq 0$, is denoted by X_K , i.e., $X_K = [x_0, \dots, x_K]$. If not otherwise stated, a variable with superscript i indicates that it belongs to subsystem i . The measure \mathbb{P} denotes the probability measure on the abstract sample space denoted by Ω . The expression F, \mathbb{P} -a.s. denotes that the event F occurs almost surely w.r.t. probability measure \mathbb{P} . The expectation operator is denoted by $\mathbb{E}[\cdot]$ and the conditional expectation is denoted by $\mathbb{E}[\cdot|\cdot]$. The relation $x \sim \mathcal{N}(0, C)$ denotes a Gaussian random variable with zero-mean and covariance matrix C . The expression $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function.

2 Problem formulation

Figure 1 shows the structure of the considered networked control system. It comprises of N independent subsystems whose feedback loops are closed over a shared communication network. The i th subsystem consists of a process \mathcal{P}^i , a controller \mathcal{C}^i , which is implemented at the actuator, and a sensor \mathcal{S}^i . The process \mathcal{P}^i is described by the following difference equation.

$$x_{k+1}^i = A^i x_k^i + B^i u_k^i + w_k^i \quad (1)$$

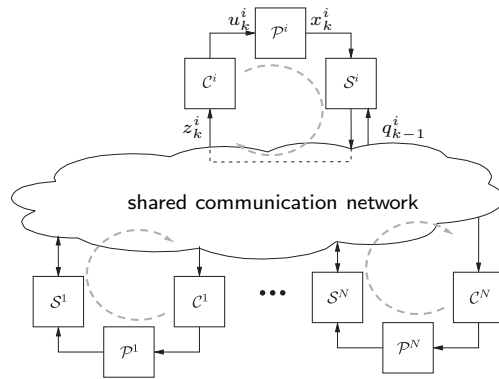


Fig. 1 System model of the networked control system with N control systems closed over a shared communication network with processes $\mathcal{P}^1, \dots, \mathcal{P}^N$, sensors $\mathcal{S}^1, \dots, \mathcal{S}^N$ and controllers $\mathcal{C}^1, \dots, \mathcal{C}^N$.

with $A^i \in \mathbb{R}^{n_i \times n_i}$, $B^i \in \mathbb{R}^{n_i \times d_i}$. The state x_k^i and the control input u_k^i are taking values in \mathbb{R}^{n_i} and in \mathbb{R}^{d_i} , respectively. The noise process w_k^i takes values in \mathbb{R}^{n_i} and is independent and identically distributed with $w_k^i \sim \mathcal{N}(0, C^i)$. The initial state, x_0^i , $i \in \{1, \dots, N\}$ is a random variable with a symmetric distribution around its mean and has a finite second order moment. The statistics of the random variables and the system parameters within a subsystem are known to the controller as well as to sensor station.

At any time k the scheduler \mathcal{S}^i situated at the sensor decides, whether a transmission slot should be requested to transmit the current state of subsystem i to the controller \mathcal{C}^i . Therefore, an event occurs within a subsystem i at time k when a transmission slot is requested. Further, it is assumed that control inputs may not be constant in between of successful transmissions and the controller \mathcal{C}^i may adjust the control inputs based on past updates. Due to bandwidth limitations the number of transmission slots denoted by N_{slot} is constrained and event-triggers must be designed at the sensors that judge the importance of transmitting an update to the corresponding controller. If there are more requests than available transmission slots at time k , then the arbitration mechanism within the communication system selects N_{slot} subsystems that may transmit information. All other subsystems are blocked and are informed instantaneously that their request has been rejected. The arbitration mechanism does not prioritize subsystems, i.e., in case of arbitration, the subsystems are chosen with identical probability. The request for a transmission of the i th subsystem at time k is defined by the variable δ_k^i which takes the following values.

$$\delta_k^i = \begin{cases} 1 & \text{request for transmission} \\ 0 & \text{idle} \end{cases}$$

We represent the arbitration mechanism by the random variable q_k^i taking values

$$q_k^i = \begin{cases} 1 & \text{allow transmission} \\ 0 & \text{block transmission} \end{cases}$$

The probability distribution of $[q_k^1, \dots, q_k^N]$ conditioned on the requests δ_k^i with $i \in \{1, \dots, N\}$ is time-invariant and has the following property.

$$\mathbb{P}[q_k^i = 1 | \delta_k^i, i \in \{1, \dots, N\}] = \begin{cases} 1 & \sum_{i=1}^N \delta_k^i \leq N_{\text{slot}} \\ \frac{N_{\text{slot}}}{\sum_{i=1}^N \delta_k^i} & \text{otherwise} \end{cases} \quad (2)$$

for subsystem i with $\delta_k^i = 1$ and

$$q_k^1(\omega) + \dots + q_k^N(\omega) = N_{\text{slot}}$$

for all sample paths $\omega \in \Omega$ for which $\delta_k^1 + \dots + \delta_k^N \geq N_{\text{slot}}$. The received data at the controller \mathcal{C}^i at time k is denoted by z_k^i and is defined by

$$z_k^i = \begin{cases} x_k^i, & \delta_k^i = 1 \wedge q_k^i = 1 \\ \emptyset, & \text{otherwise} \end{cases} \quad (3)$$

Every subsystem i , $i \in \{1, \dots, N\}$, possesses an individual cost function J^i which is given by the average-cost

$$J^i = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{k=0}^{T-1} x_k^{i,\top} Q_x^i x_k^i + u_k^{i,\top} Q_u^i u_k^i \right]. \quad (4)$$

The weighting matrix Q_x^i is positive definite and Q_u^i is positive semi-definite for each $i \in \{1, \dots, N\}$. We assume that the pair (A^i, B^i) is stabilizable and the pair $(A^i, Q_x^{i,\frac{1}{2}})$ is detectable with $Q_x^i = (Q_x^{i,\frac{1}{2}})^\top Q_x^{i,\frac{1}{2}}$.

The design objective is to design control and scheduling laws that minimize the social cost V that is given by the average of the individual costs, i.e.,

$$V = \frac{1}{N} \sum_{i=1}^N J^i. \quad (5)$$

The control law $\gamma^i = \{\gamma_0^i, \gamma_1^i, \dots\}$ that reflects the behavior of the controller \mathcal{C}^i at subsystem i is described by the mappings γ_k^i , $k \in \{0, 1, \dots\}$. Admissible laws are measurable, causal maps of the available observations until time k denoted by Z_k^i , i.e.,

$$u_k^i = \gamma_k^i(Z_k^i).$$

The set of admissible laws is denoted by Γ . The policy of the scheduler is given by $\pi^i = \{\pi_1^i, \pi_2^i, \dots\}$. The map π_k^i is defined as

$$\delta_k^i = \pi_k^i(X_k^i, Q_{k-1}^i),$$

where π_k^i is measurable with respect to the past observations X_k^i, Q_{k-1}^i . The set of admissible scheduling laws is denoted by Π^{ET} . The scheduling laws in Π^{ET} are only using local information X_k^i, Q_{k-1}^i in order to determine whether a slot for transmission is to be requested. Time-triggered schedulers constitute a special case within Π^{ET} , as the map π_k^i is independent of X_k^i, Q_{k-1}^i for any k , i.e., time-triggered schedulers are elements of $\{0, 1\}^\infty$. Therefore, the set of admissible time-triggered scheduling laws denoted by Π^{TT} can be considered as a subset of Π^{ET} .

On the other hand, by allowing that the decisions of the schedulers may depend on all measurements $\{X_k^1, Q_{k-1}^1, \dots, X_k^N, Q_{k-1}^N\}$, we obtain a centralized scheduling structure. Let Π^{CEN} denote the set of admissible centralized schedulers. It can be observed that the decentralized event-trigger policies in Π^{ET} is also contained in Π^{CEN} . Hence, we obtain the following relationship among the three scheduling structures.

$$\Pi^{\text{TT}} \subset \Pi^{\text{ET}} \subset \Pi^{\text{CEN}}.$$

It implies that the cost V of the optimal decentralized event-triggered law in Π^{ET} is lower bounded by the optimal centralized scheduler and upper bounded by the minimal costs of the optimal time-triggered scheduler. In the following, we focus on the design of the decentralized scheduler π^i and controller γ^i with the corresponding control law given by

$$\inf_{\substack{\gamma^i \in \Gamma \\ \pi^i \in \Pi^{\text{ET}} \\ i \in \{1, \dots, N\}}} V. \quad (6)$$

The optimal time-triggered and centralized scheduling laws are considered again in section 4.

3 Event-triggered control design

In the following, the aim is to develop a design method for the decentralized event-triggered controller minimizing (6). This section is divided into several subsections. In subsection 3.1, the approximative bi-level formulation of the original optimization problem is addressed. The structural properties of the solution of such problem are studied in subsection 3.2. The resulting closed-loop behavior is discussed in the remaining subsections by giving stability conditions in subsection 3.3 and an asymptotic optimality property in subsection 3.4. In subsection 3.5, the issue of computational complexity of the approximative bi-level approach is addressed.

3.1 Approximative bi-level formulation

Although the coupling between subsystems is solely caused by the resource limitation, determining the optimal event-based control system that solves (6)

is a hard problem. The reason for this is secondarily given by the fact that the number of subsystems might be large, but is rather due to the distributed information pattern. It is shown in Witsenhausen (1968) that optimal stochastic control problems with distributed information pattern are generally hard to solve, even if linear dynamics and quadratic cost functions are considered.

Besides the distributed information pattern, the impact of the bandwidth limitation is another complicating factor. In contrast to time-triggered scheduling schemes, it is in general not possible to guarantee that a transmission request will be approved or not for decentralized event-triggered scheduling with resource constraints. Despite of the non-determinism due to the contention-based communication network, we will observe a significant performance improvement of the event-triggered scheme compared with a time-triggered scheme in section 4.

However, in order to still obtain a systematic approach to find the event-triggered controllers that minimize V in (6), we introduce the following approximation. Thereby, the hard constraint that N_{slot} transmissions are allowed at maximum at each time k is weakened and we require merely that the average number of transmissions per time step is upper bounded by N_{slot} . We define the individual average transmission rate of the i th subsystem by

$$r^i = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{k=0}^{T-1} \delta_k^i \right].$$

Then, the average rate constraint can be stated as

$$\sum_{i=1}^N r^i \leq N_{\text{slot}}.$$

With this rate constraint, the approximative optimization problem can be stated as a bi-level optimization problem. Both levels are coupled through the average transmission rates r^i , $i \in \{1, \dots, N\}$. After assigning an upper bound, \bar{r}^i , on the transmission rate to each subsystem i , the first level of the optimization problem is given by

$$J^{i,*}(\bar{r}^i) = \inf_{\substack{\gamma^i \in \Gamma \\ \pi^i \in \Pi^{\text{ET}} \\ r^i \leq \bar{r}^i}} J^i(\gamma^i, \pi^i) \quad (7)$$

where x_k^i is the state evolving by (1) and it is assumed that in contrast to (3) every request is permitted, i.e.,

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

It should be noted that above optimization problem can be solved locally in each subsystem for a given \bar{r}^i , i.e., assuming a given rate distribution, the optimization problems in (7) for $i \in \{1, \dots, N\}$ are completely decoupled from

each other. The second level of the optimization problem determines finally the optimal transmission rate distribution among the subsystems and is given by

$$\bar{V}^* = \inf_{\substack{\bar{r}^1, \dots, \bar{r}^N \\ \sum_{i=1}^N \bar{r}^i \leq N_{\text{slot}}}} \frac{1}{N} \sum_{i=1}^N J^{i,*}(\bar{r}^i) \quad (8)$$

Thus, the resulting bi-level approach has a hierarchical structure, where the second level can be considered as the global coordinating layer assigning resources to the first layer, where the optimization problem is solved locally in each subsystem.

3.2 Structural properties

The local optimization problem formulated in (8) still has a distributed information pattern because the observations at the controller and at the sensor station may differ. Therefore, standard solution algorithms can not be applied for the constrained optimal control problem in the first level. But instead, existing literature has shown that the optimal solution of (8) exhibits certain structural properties Molin and Hirche (2009); Molin et al (2011); Lipsa and Martins (2011); Xu and Hespanha (2004). These properties allow an efficient calculation of the optimal event-triggered controller in each subsystem. The subsequent theorem sheds light on the form of the optimal control law $\gamma^{i,*}$.

Theorem 1 (Certainty equivalence of $\gamma^{i,*}$, Molin and Hirche (2009))

Let (A^i, B^i) be stabilizable and $(A^i, Q_x^{i,\frac{1}{2}})$ detectable. Then, the form of the optimal control law $\gamma^{i,*}$ for optimization problem (7) is given by

$$u_k^i = \gamma_k^{i,*}(Z_k^i) = -L^i E[x_k^i | Z_k^i], \quad (9)$$

where the linear gain L^i can be calculated by the algebraic Riccati equation

$$\begin{aligned} L^i &= (B^{i,\top} P^i B^i + Q_u^i)^{-1} B^{i,\top} P^i A^i, \\ P^i &= A^{i,\top} (P^i - P^i B^i (B^{i,\top} P^i B^i + Q_u^i)^{-1} B^{i,\top} P^i) A^i + Q_x^i. \end{aligned}$$

□

An optimal controller is said to be *certainty equivalent* when it results from the solution of the corresponding deterministic optimal control problem by substituting all noise variables by its means and by replacing the state x_k^i by its least squares estimate $E[x_k^i | Z_k^i]$.

Theorem 1 enables a number of further simplifications. When taking into account that the distributions of the noise variables are symmetric, then it is shown in Lipsa and Martins (2011) for first-order systems that the optimal scheduling law $\pi^{i,*}$ is a symmetric threshold function of the estimation error. Subsequently, it is conjectured that this is also valid for higher-order systems.

Then, the optimal estimator can be stated similarly to Molin and Hirche (2009) as

$$\mathbb{E}[x_k^i | Z_k^i] = \begin{cases} x_k^i & \delta_k^i = 1 \\ (A^i - B^i L^i) \mathbb{E}[x_{k-1}^i | Z_k^i] & \text{otherwise} \end{cases} \quad (10)$$

with $\mathbb{E}[x_0^i | Z_0^i] = 0$ for $z_0^i = \emptyset$. The first condition in above distinction of cases is extended to $\delta_k^i = 1 \wedge q_k^i = 1$ for the original communication network.

By defining the estimation error e_k^i by

$$e_k^i = x_k^i - \mathbb{E}[x_k^i | Z_{k-1}^i],$$

the determination of the optimal scheduling law can be regarded as a constrained Markov decision process Altman (1999). The Markov state $e_k \in \mathbb{R}^{n_i}$ evolves by the time-invariant difference equation

$$e_{k+1}^i = g^i(e_k^i, \delta_k^i, w_k^i) = (1 - \delta_k^i) A^i e_k^i + w_k^i \quad (11)$$

with initial condition $e_0^i = x_0^i - \mathbb{E}[x_0^i]$. Substituting the optimal control law $\gamma^{i,*}$ of Theorem 1 into the costs J^i , we obtain the following stochastic optimal control problem.

$$\inf_{\substack{\pi^i \in \Pi^M \\ r^i \leq \bar{r}^i}} J^{i,\mathcal{S}} \quad (12)$$

with

$$J^{i,\mathcal{S}} = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{k=0}^{T-1} (1 - \delta_k^i) e_k^{i,\top} Q_e^i e_k^i \right]$$

where $Q_e^i = L^{i,\top} (Q_u^i + B^{i,\top} P^i B^i) L^i$. The set Π^M denotes the set of all Markov policies, which is defined as the set of all measurable maps from \mathbb{R}^{n_i} to $\{0, 1\}$. Considering the optimal control law $\gamma^{i,*}$ with the optimal estimator given by (10), we have the following relationship between the original optimization problem (7).

$$J^i(\gamma^{i,*}, \cdot) = J^{i,\mathcal{S}}(\cdot) + \text{tr}[P^i C^i].$$

The expression $\text{tr}[P^i C^i]$ is constant and can therefore be omitted in the optimization problem (12).

In order to simplify the analysis of solving (12), we introduce the following technical assumption.

Assumption 1 *The scheduling policy of subsystem i is $\pi_k^i(\delta_k^i = 1 | e_k^i) = 1$ for $\|e_k^i\|_2 > M^i$ for some arbitrary M^i , $i \in \{1, \dots, N\}$. \square*

Assumption 1 does not impose a severe restriction on the scheduling laws, as M^i can be chosen arbitrarily large for each subsystem. On the one hand, it follows from above assumption that the running cost $(1 - \delta_k^i)e_k^{i,\top} Q_e^i e_k^i$ is uniformly bounded. On the other hand, it has been shown in Xu and Hespanha (2004) that the resulting Markov chain is ergodic, when using an arbitrary Markov policy in Π^M that satisfies Assumption 1. The constrained Markov decision process (12) under Assumption 1 is formulated as an optimization problem without constraints by taking a Lagrangian approach that results in

$$\inf_{\pi^i \in \Pi^M} J^{i,\mathcal{S}} + \lambda r^i. \quad (13)$$

The non-negative weight λ can be regarded as the Lagrange multiplier of the constrained Markov decision process (12), Altman (1999). If λ is fixed, the optimization problem (13) becomes a standard stochastic optimal control problem that can be solved by value iteration as for example shown in Bertsekas (2007). The resulting optimal scheduling policy is a threshold policy of the estimation error e_k^i . Due to the boundedness of the running cost and due to ergodicity that are guaranteed by Assumption 1, the value iteration converges to the optimal solution as shown in Bertsekas (2007). Instead of the direct determination of the optimal solution together with the optimal Lagrange multiplier, we regard above optimization problem as the scalarization approach as in Boyd and Vandenberghe (2004) of the corresponding multi-objective optimization problem with cost vector $[J^{i,\mathcal{S}}, r^i]$.

For any $\lambda \in [0, \infty)$, we obtain a Pareto optimal point in the $J^{i,\mathcal{S}}-r^i$ -plane. It is easy to show that the coordinates $[J^{i,\mathcal{S},*}, r^{i,*}]$ are monotone in λ , i.e., $J^{i,\mathcal{S},*}$ is monotonically increasing in λ and $r^{i,*}$ is monotonically decreasing in λ .

From the continuity of the difference value function in λ that follows from chapter 3.5 in Hernández-Lerma (1989), and the absolute continuity of the stationary distribution of the π^i -controlled Markov chain we can conclude that $[J^{i,\mathcal{S},*}, r^{i,*}]$ is continuous in λ . Therefore, the scalarization approach in (13) yields the desired function $J^{i,*}(\bar{r}^i)$ that results from the set of Pareto optimal points. The function $J^{i,*}(\bar{r}^i)$ is convex and monotonically decreasing in \bar{r}^i . This implies that the global optimization problem in the second level defined by (8) is a convex resource allocation problem, for which many efficient solution algorithms exist, such as in Shakkottai and Srikant (2007).

Figure 2 summarizes the structural properties of the optimal event-triggered controller of the i th subsystem solving optimization problem (8).

3.3 Stochastic stability

The approach developed in the previous subsections ignores that at most N_{slot} transmissions are admitted at each time step. Therefore, it is at first not guaranteed whether the aggregated system with the actual communication constraint will be stochastically stable by using the approximative design approach.

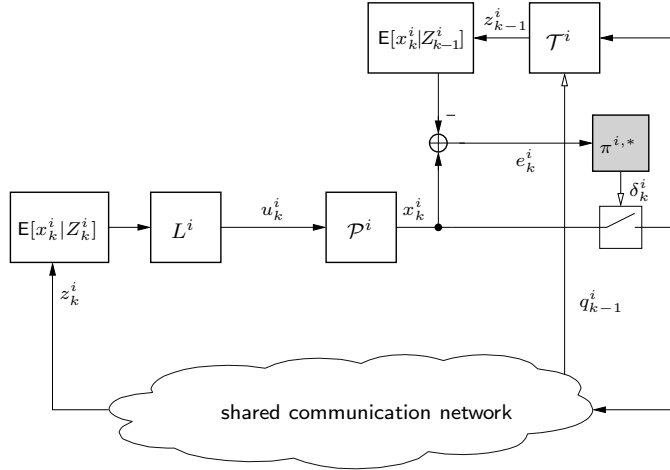


Fig. 2 Block diagram of the optimal decentralized event-triggered control system for subsystem i . Functional block \mathcal{T}^i reconstructs the received signal, z_{k-1}^i , at the controller.

In this paper, a Markov chain is said to be *stochastically stable*, when it is ergodic, and the corresponding stationary distribution has a finite second-order moment.

When there exists a unique stationary distribution to which the Markov chain converges for $k \rightarrow \infty$ for any initial condition, then the Markov chain is said to be *ergodic*.

The following theorem states conditions that guarantee stochastic stability for the aggregated system.

Theorem 2 Let $\gamma^{i,*}$ be the control law given by (9) and let $\pi^i \in \Gamma^M$ be a scheduling law which satisfies Assumption 1. If

$$\frac{N_{slot}}{N} > 1 - \frac{1}{\|A^i\|_2^2} \quad (14)$$

is satisfied for all $i \in \{1, \dots, N\}$, then the Markov chain representing the aggregated system is stochastically stable. \square

Proof The Markov state of the complete system is given by the aggregation of the augmented state variable $[x_k^i, e_k^i]$ of the subsystems $i \in \{1, \dots, N\}$ assuming that the control law $\gamma^{i,*}$ and scheduling law π^i are used.

The evolution of the estimation error e_k^i that has been described in (11) must be extended in order to incorporate the occurrence of a refused request. Therefore, we obtain the following difference equation

$$e_{k+1}^i = (1 - q_k^i \pi^i(e_k^i)) A^i e_k^i + w_k^i \quad (15)$$

with initial condition $e_0^i = x_0^i - E[x_0^i]$ and the conditioned distribution of q_k^i defined by (2). The evolution of the process state, x_k^i , can be rewritten as

$$x_{k+1}^i = (A^i - B^i L^i) x_k^i + (1 - q_k^i \pi^i(e_k^i)) B^i L^i e_k^i + w_k^i. \quad (16)$$

It can be observed from (15) and (16) that the state dynamics $[x_k^i, e_k^i]$ within each subsystem $i \in \{1, \dots, N\}$ has a triangular structure, i.e., the process state x_k^i does not affect the evolution of the estimation error e_k^i .

Furthermore, the evolution defined by (16) can be considered as a stable linear system with additive noise given by $(1 - q_k^i \pi^i(e_k^i))B^i L^i e_k^i + w_k^i$.

Hence, showing that the process $\zeta_k = [e_k^1, \dots, e_k^N]$ is ergodic and has a finite second-order moment implies that the aggregated system state is stochastically stable.

In the following, we study the asymptotic behavior of the Markov chain given by $\zeta_k = [e_k^1, \dots, e_k^N]$ which is described by (11). The stability analysis is based on drift conditions based on Meyn and Tweedie (1996) developed for Markov chains with uncountable state spaces. The drift operator Δ is defined as

$$\Delta h(\zeta_k) = \mathbb{E}[h(\zeta_{k+1})|\zeta_k] - h(\zeta_k), \quad \zeta_k \in \mathbb{R}^{n_1 + \dots + n_N}.$$

where h is a map from $\mathbb{R}^{n_1 + \dots + n_N}$ to \mathbb{R} .

We consider the following Lyapunov candidate

$$h(\zeta_k) = \sum_{i=1}^N \|e_k^i\|_2^2. \quad (17)$$

It follows immediately that the drift for this choice of h is bounded within any compact set. Further, we define the compact set

$$\mathcal{M} = \{\zeta = [e^1, \dots, e^N] \in \mathbb{R}^{n_1 + \dots + n_N} \mid \|e^i\|_2 \leq M^i, i \in \{1, \dots, N\}\}.$$

Due to absolute continuity of the distribution of w_k^i , the Markov chain is ψ -irreducible and aperiodic. Based on these facts, we rely on the following drift condition proposed in Meyn and Tweedie (1996). If the condition

$$\Delta h(\zeta_k) \leq -\epsilon h(\zeta_k), \quad \zeta_k \in \mathbb{R}^{n_1 + \dots + n_N} \setminus \mathcal{O}, \quad (18)$$

where $\epsilon > 0$ and \mathcal{O} is a compact set, is satisfied, then the Markov chain of the aggregate system with state ζ_k is stochastically stable.

In the following, we define the drift of a subsystem as

$$\Delta_i h(\zeta_k) = \mathbb{E}[\|e_k^i\|_2^2 | \zeta_k] - \|e_k^i\|_2^2$$

Due to linearity of the conditional expectation, the definition implies that

$$\Delta h(\zeta_k) = \sum_{i=1}^N \Delta_i h(\zeta_k)$$

Consider two cases for calculating an upper bound on $\Delta_i h(\zeta_k)$. First, let $\|e_k^i\|_2^2 > M^i$. The statistical independence of w_k^i with respect to q_k^i and e_k^i and the fact that $w_k^i \sim \mathcal{N}(0, C^i)$ allows the following simplification.

$$\Delta_i h(\zeta_k) = \mathbb{E}[1 - q_k^i | \zeta_k] \|A^i e_k^i\|_2^2 + \text{tr}[C^i] - \|e_k^i\|_2^2,$$

The expression $\mathbb{E}[1 - q_k^i | \zeta_k]$ describes the average probability that a request of subsystem i is blocked and is upper bounded by $1 - \frac{N_{\text{slot}}}{N}$ because of (2). Therefore, the drift can be upper bounded in the first case by

$$\Delta_i h(\zeta_k) \leq \left(1 - \frac{N_{\text{slot}}}{N}\right) \|A^i\|_2^2 - 1 \|e_k^i\|_2^2 + \text{tr}[C^i]. \quad (19)$$

In the case of $\|e_k^i\|_2^2 \leq M^i$, we have the following uniform bound

$$\Delta_i h(\zeta_k) \leq \| \|A^i\|_2^2 - 1 \| (M^i)^2 + \text{tr}[C^i].$$

By choosing $\mathcal{O} \supset \mathcal{M}$, we ensure that at least one of the subsystems tries to send information over the communication network. This implies that there is at least one subsystem for which we have a Δ_i satisfying the upper bound given above for the first case. Selecting \mathcal{O} accordingly, we can also guarantee that the i th subsystem with the maximal $\|e_k^i\|_2$ is greater than M^i and therefore the bound of the first case applies for this subsystem for all $\zeta \notin \mathcal{O}$. Together with $\sup_i \|e_k^i\|_2^2 \geq \frac{1}{N} \|\zeta_k\|_2^2$, the condition in (14) guarantees that we can find an appropriate ϵ and an appropriate compact set \mathcal{O} such that the condition in (18) is satisfied for all $\zeta \notin \mathcal{O}$. \square

It shall be noted that the stability condition in Theorem 2 can be established separately for each subsystem and does not put constraints on the scheduling behavior of the other subsystems. This implies that stochastic stability can still be guaranteed, even if a malicious subsystem is continuously requesting for transmission.

3.4 Asymptotic optimality

In the following, we focus on the analysis of the approximative bi-level approach developed in subsection 3.1, when the number of subsystems, N , approaches infinity. A design approach is said to be *asymptotically optimal*, when the costs of the solution approaches the optimal costs arbitrarily close for a sufficiently large N . The relevant system parameters of subsystem i are summarized in the 4-tuple $\mathcal{K}^i = \{A^i, B^i, Q_x^i, Q_u^i\}$. In order to compare the control performance between the aggregated systems with increasing N , we scale the communication network accordingly such that the ratio N_{slot}/N stays constant. It is also assumed that there is a finite number of subsystem classes denoted by \bar{N} , i.e., $\mathcal{K}^i \in \{\mathcal{K}^1, \dots, \mathcal{K}^{\bar{N}}\}$ for all $i \in \{1, \dots, N\}$. The number of subsystems in a subsystem class is scaled with increasing N , such that their ratio between each other remains constant. This assumption implies that it suffices to consider subsystems of a subsystem class \mathcal{K}^j separately with a fixed slot assignment of $N_{\text{slot}}^j \leq N_{\text{slot}}$.

We also assume that we have *chaoticity in equilibrium* Graham (2000) which corresponds to the exchange of the limits of time and the number of subsystems, i.e., $\lim_{k \rightarrow \infty} \lim_{N \rightarrow \infty} = \lim_{N \rightarrow \infty} \lim_{k \rightarrow \infty}$. To prove above assumption for the underlying system is out of the scope of this work and shall

be considered in future work. In the following analysis, we therefore consider that the system is in its stationary regime and consider the limit when N approaches ∞ . The subsequent theorem gives a statement about the optimality properties of the approximative design approach.

Theorem 3 *If Assumption 1 and the stability condition in (14) are satisfied, then the solution of the bi-level optimization problem defined by (7) and (8) is asymptotically optimal with respect to the optimization problem given by (6).*

Proof First, note that the optimal cost \bar{V}^* resulting from the second level optimization given by (8) are a lower bound of the original optimization problem. This is because the hard rate constraint to be satisfied at each time step is relaxed by merely restraining the average total transmission rate. In the following, we show that the deviation from the optimal cost \bar{V}^* because of the actual hard rate constraint becomes arbitrarily small for sufficiently large N . As already mentioned above, it suffices to restrict ourselves to a networked control system composed of identical subsystems. It follows from the convexity of the function $J^{i,*}(\bar{r}^i)$ resulting from (7) that the assigned individual average transmission rates are identical. Therefore, the transmission rate is given by $r^i = N_{\text{slot}}/N$ for each subsystem i , $N \in \{1, \dots, N\}$.

Next, we observe that the event of a request of a subsystem relatively to its most recent successful transmission can be regarded as a renewal process. This process is identical with a system without hard rate constraint as the blocking behavior of the arbitration mechanism is removed. Due to Theorem 2, it follows from Assumption 1 and condition (14) that the resulting Markov chain characterizing the aggregate behavior is ergodic for any N . That means that there exists a stationary distribution of the Markov state which implies that the renewal process is aperiodic and recurrent. As all subsystems are identical, the stationary distribution is symmetric with respect to the subsystems.

Therefore, we have

$$\mathbb{P}_{\text{st}}[\delta_k^i = 1 | \text{last transmission successful}] = r_i = \frac{N_{\text{slot}}}{N}.$$

Moreover, we define the probability of requesting transmission assuming the last transmission were not successful as

$$\beta = \mathbb{P}_{\text{st}}[\delta_k^i = 1 | \text{last transmission not successful}].$$

Further, we assume that $\beta > r^i$, which means that it is more likely to transmit when the last transmission has not been successful. This can be easily proved for scalar systems and is also conjectured for higher-order systems.

We define the μ_k^N to be the fraction of the subsystems whose most recent transmission has been successful, i.e.,

$$\mu_k^N = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\text{last transmission successful of } i\text{th subsystem}}$$

Consider the deterministic process

$$\mu_{k+1}^{\infty} = \mu_k^{\infty} - \mu_k^{\infty} r_i + \min\left(\frac{N_{\text{slot}}}{N}, \beta(1 - \mu_k^{\infty}) + \mu_k^{\infty} r_i\right).$$

Due to the law of large numbers, the μ_k^N converges weakly to above process in the stationary regime. We are interested in the fixed points of above process, as we are assuming the process evolving is stationary. It is easy to see that the only fixed point is given by $\mu_k^{\infty} = 1$. Hence, the fraction of the subsystems whose most recent transmission has been successful converges weakly to 1. Therefore, we converge to the optimal cost \bar{V}^* resulting from (8). This implies that the design method described by (7) and (8) is asymptotically optimal. \square

3.5 Computational complexity

For the calculation of the Pareto frontier in the first level of the bi-level approach, we need to solve a dynamic program for different values of λ . This can be accomplished by a sequence of value iterations for subsystems with a moderate state dimension. For higher-order systems, there exist approximative methods, e.g., as developed in Cogill (2009) that reduces the problem to a sequence of semidefinite programs. The solution of the algebraic Riccati equation for the optimal control law in (9) can be computed in polynomial-time. Therefore, it does not represent a computational burden for higher-dimensional systems. It should be noted that the Pareto frontier can be determined offline for each subsystem without having to take the system parameters of the communication network into account.

The solution of the global resource allocation problem with N design parameter in the second level is a convex optimization problem, which can be solved efficiently. Hence, the optimal solution can also be computed for a large number of subsystems.

In contrast to the centralized and time-triggered approach, the event-triggered approach is more flexible with respect to changes during runtime. When subsystems are attached to or detached from the current contention-based communication system, we only have to solve the optimization problem within the second level to adjust the transmission rates.

4 Numerical simulations

In this section, the proposed event-triggered approach is evaluated and compared with the time-triggered and the centralized approach. In order to facilitate the presentation, we restrict our attention to scalar subsystems.

First, suppose we have identical subsystems with parameters

$$\mathcal{K}^i = \mathcal{K} = (1, 1, 1, 0).$$

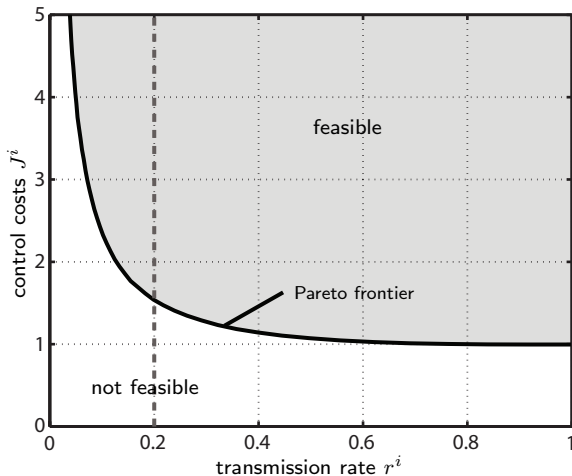


Fig. 3 Pareto frontier of a subsystem and system parameters $\mathcal{K} = (1, 1, 1, 0)$. The vertical line indicates the rate constraint.

The communication network has a ratio between available transmission slots and the number of subsystems of $N_{\text{slot}}/N = 0.2$. The Pareto optimal cost region $[J^i, r^i]$ for a subsystem with parameters \mathcal{K} including the rate constraint is drawn in Fig. 3. We observe that J^i is a decreasing and convex function with respect to r^i . For identical subsystems, there is a substantial simplification in the global resource allocation problem performed in the second level as all subsystems attain the same transmission rate, i.e., the optimal transmission rate is given by $r^{i,*} = 0.2$. The optimal cost point is attained at $[J^{i,*}, r^{i,*}] = [1.54, 0.2]$ by an event-triggered scheduling policy $\pi^{i,*}$ that is given by $\delta_k^i = \mathbb{1}_{\{|e_k^i| > 1.7\}}$. The optimal control law gain L^i is given by 1.

Figure 4 compares the cost of the decentralized event-triggered scheme with the optimal time-triggered scheme and the optimal centralized scheduling. It shows the cost per subsystem for various numbers of identical subsystems N with $N_{\text{slot}}/N = 0.2$. The costs for $N \in \{5, 25, 100, 250, 500\}$ are determined by Monte Carlo simulations with a time horizon of $T = 10\,000$. The optimal control law for both the optimal time-triggered scheme and the optimal centralized scheme are given by $u_k^i = -L^i \mathbb{E}[x_k^i | Z_k^i]$ with $L^i = 1$. In the optimal time-triggered scheme, time slots for transmission are assigned successively. Subsystems transmit information periodically with transmission period $\frac{N}{N_{\text{slot}}}$, where we assume that N is a multiple of 5. In the case of identical subsystems, the optimal centralized scheduler selects at each time step k the N_{slot} subsystems with maximum magnitude $|e_k^i|$ whose feedback loop are then closed. Such kind of protocol can be realized by prioritize the medium access through $|e_k^i|$ which has also been done in Walsh et al (2002). It should be noted that this scheduler can be regarded as a lower bound on the performance that can be achieved over the communication networks. In the case of heterogeneous mul-

tidimensional systems, it remains an open problem how to realize centralized schedulers without gathering the state information of all subsystems.

We observe in Fig. 4 that the cost of the optimal decentralized scheduling algorithm approximates this lower bound very closely and outperforms the optimal time-triggered scheme significantly. On the other hand, it can be seen that the costs converge to the asymptotic costs for $N \rightarrow \infty$ very rapidly. Already for $N = 100$, the performance gap is less than 10%.

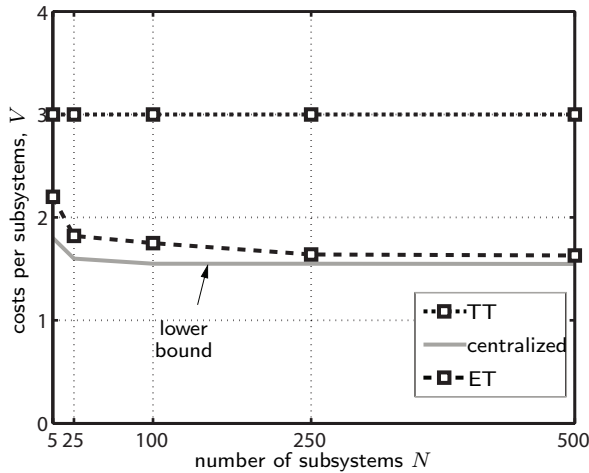


Fig. 4 Numerical comparison of time-triggered (TT), event-triggered (ET) and centralized schemes for a networked control system with homogeneous subsystems $\mathcal{K} = (1, 1, 1, 0)$ and $N_{\text{slot}}/N = 0.2$.

Next, we consider a heterogeneous system, where we have two different kinds of subsystems occurring at the same amount. The system parameters are $\mathcal{K}^1 = (1.25, 1, 1, 0)$ and $\mathcal{K}^2 = (0.75, 1, 1, 0)$ and the communication network has a ratio of $N_{\text{slot}}/N = 0.5$. We note that the stability condition (14) in Theorem 2 is satisfied for the underlying subsystems.

Having obtained the Pareto curves from the first level optimization for both subsystems sketched in Fig. 5, the resource allocation problem given by (8) determines the optimal rate pair. The dashed line in Fig. 5 depicts the mean cost per subsystem V as a function of r^1 for $N = 2$ without collisions. It can be seen that the total cost V is convex with respect to r^1 and it is minimized at the rate pair $[r^1, r^2] = [0.6, 0.4]$ taking a value of 1.07. The optimal control gain is given by $L^i = A^i$ for both subsystems and the scheduling laws are threshold policies, where $\delta_k^1 = \mathbb{1}_{\{|e_k| > 0.5\}}$ for \mathcal{K}^1 and $\delta_k^2 = \mathbb{1}_{\{|e_k^2| > 0.95\}}$ for \mathcal{K}^2 .

Concerning the performance in the presence of the shared network, we consider the mean costs V depicted in Fig. 6 for $N \in \{2, 10, 50, 100, 250, 500\}$. The optimal time-triggered scheme involves a brute-force search over all possible combinations of transmission times. To keep this combinatorial problem numerically tractable, we restricted the admissible transmission scheme to be

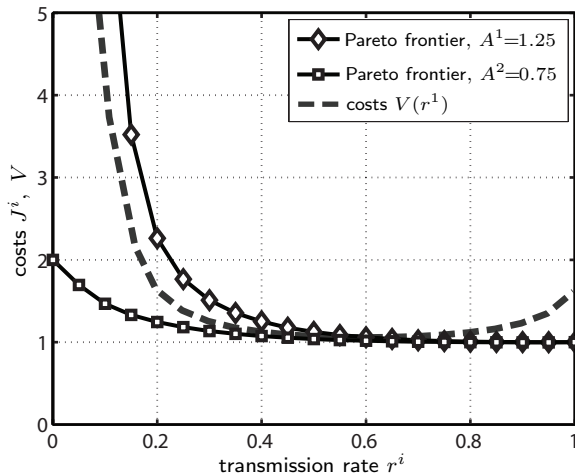


Fig. 5 Solid lines: Pareto frontiers of two different subsystems with system parameters $\mathcal{K}^1 = (1.25, 1, 1, 0)$ and $\mathcal{K}^2 = (0.75, 1, 1, 0)$. Dashed line: Total cost $V(r^1) = \frac{1}{2}(J^1(r^1) + J^2(r^2))$ and constraint $\frac{1}{N}(r^1 + r^2) \leq 0.5$. The optimal rate pair is given at $[r^1, r^2] = [0.6, 0.4]$ with total cost $V = 1.07$ for the two subsystems without collisions.

periodical for subsystems \mathcal{K}^2 . The optimal periodical transmission scheme is then given by $[\delta_0^1, \delta_1^1, \delta_2^1, \dots] = [1, 1, 0, \dots]$ and $[\delta_0^2, \delta_1^2, \delta_2^2, \dots] = [0, 0, 1, \dots]$ with period 3. A lower bound is given by $V = 1$ assuming no communication constraints on the feedback channels. It should be noted that the lower bound is certainly not tight for the optimal event-triggered scheme and therefore there will always be a gap between the proposed scheme and this bound even in the limiting case $N \rightarrow \infty$. As can be regarded from Fig. 6, this lower bound is approached with a gap of less than 10% for increasing N and the time-triggered scheme is outperformed for every N .

5 Conclusion

This paper shows that decentralized event-triggered control constitutes an attractive design approach for resource constrained networked control systems. The proposed design method manages to establish a compromise between computational complexity and overall performance that circumvents to take the complex behavior of the contention-based network into account. Despite the decreased predictability and a close interaction between control and communication in contrast to time-triggered control schemes, the gain from the proposed event-triggered control scheme is an increased level of flexibility, robustness, and a significant improvement on the control performance for the shown examples.

The major issues to be addressed in future research is the investigation on the online adjustment of the event-trigger according to the network traffic that

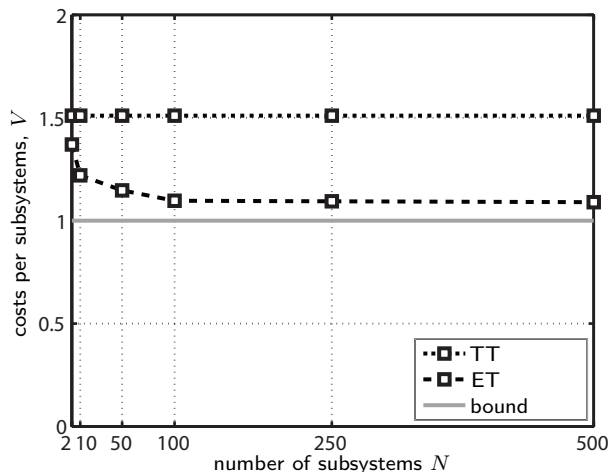


Fig. 6 Numerical comparison of time-triggered (TT) and event-triggered (ET) schemes for a networked control system with heterogeneous subsystems of two classes \mathcal{K}^1 and \mathcal{K}^2 and $N_{\text{slot}}/N = 0.5$.

also leads to a decentralization of the global resource allocation problem as well as the study of more complicated models for the communication network.

References

- Altman E (1999) Constrained Markov Decision Processes. Chapman and Hall/CRC
- Åström K, Bernhardsson B (2002) Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In: Decision and Control, 2002, Proceedings of the 41st IEEE Conference on, vol 2, pp 2011–2016
- Bertsekas DP (2007) Dynamic programming and optimal control. Vol. II. 3rd ed., Athena Scientific, Belmont, MA
- Blind R, Allgöwer F (2011a) Analysis of Networked Event-Based Control with a Shared Communication Medium: Part I - Pure ALOHA. In: IFAC World Congress, pp 10,092–10,097
- Blind R, Allgöwer F (2011b) Analysis of Networked Event-Based Control with a Shared Communication Medium: Part II - Slotted ALOHA. In: IFAC World Congress, pp 8830–8835
- Blind R, Allgöwer F (2011c) On the optimal sending rate for networked control systems with a shared communication medium. In: Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on, pp 4704–4709
- Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge Univ. Press
- Cervin A, Henningsson T (2008) Scheduling of event-triggered controllers on a shared network. In: Decision and Control, 2008. CDC 2008. 47th IEEE Conference on, pp 3601–3606
- Cogill R (2009) Event-based control using quadratic approximate value functions. In: Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on, pp 5883–5888
- Graham C (2000) Chaoticity on path space for a queueing network with selection of the shortest queue among several. Journal of Applied Probability 37(1):198–211
- Heemels W, Sandee J, Van Den Bosch P (2008) Analysis of event-driven controllers for linear systems. International journal of control 81(4):571–590

- Henningsson T, Cervin A (2010) A simple model for the interference between event-based control loops using a shared medium. In: Decision and Control (CDC), 2010 49th IEEE Conference on, pp 3240–3245
- Henningsson T, Johansson E, Cervin A (2008) Sporadic event-based control of first-order linear stochastic systems. *Automatica* 44(11):2890–2895
- Hernández-Lerma O (1989) Adaptive Markov control processes. Springer-Verlag New York
- Lipsa G, Martins N (2011) Remote state estimation with communication costs for first-order LTI systems. *Automatic Control, IEEE Transactions on* 56(9):2013–2025
- Lunze J, Lehmann D (2010) A state-feedback approach to event-based control. *Automatica* 46(1):211–215
- Meyn S, Tweedie R (1996) Markov chains and stochastic stability. Springer London
- Molin A, Hirche S (2009) On LQG joint optimal scheduling and control under communication constraints. In: Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on, pp 5832–5838
- Molin A, Hirche S (2011) Optimal design of decentralized event-triggered controllers for large-scale systems with contention-based communication. In: Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on, pp 4710–4716
- Molin A, Tischer H, Hirche S (2011) Order reduction in optimal event-triggered control design for linear stochastic systems. In: American Control Conference (ACC), 2011, pp 2222–2227
- Rabi M, Moustakides G, Baras J (2012) Adaptive sampling for linear state estimation. *SIAM Journal on Control and Optimization* 50(2):672–702
- Shakkottai S, Srikant R (2007) Network optimization and control. *Foundations and Trends in Networking* 2(3):271–379
- Tabuada P (2007) Event-triggered real-time scheduling of stabilizing control tasks. *Automatic Control, IEEE Transactions on* 52(9):1680–1685
- Walsh G, Ye H, Bushnell L (2002) Stability analysis of networked control systems. *Control Systems Technology, IEEE Transactions on* 10(3):438–446
- Witsenhausen HS (1968) A counterexample in stochastic optimum control. *SIAM Journal on Control* 6(1):131–147
- Xu Y, Hespanha J (2004) Optimal communication logics in networked control systems. In: Decision and Control, 2004. CDC. 43rd IEEE Conference on, vol 4, pp 3527–3532