*Article*

# Where Freshness Matters in the Control Loop: Mixed Age-of-Information and Event-based Co-design for Multi-loop Networked Control Systems

**Mohammad H. Mamduhi [1,*], Jaya Prakash Champati [2], James Gross [2] and Karl H. Johansson [1]**

1 Division of Decision and Control Systems, School of Electrical and Computer Engineering, KTH Royal Institute of Technology; mamduhi@kth.se (M.H.M.); kallej@kth.se (K.H.J.)
2 Division of Information Science and Engineering, School of Electrical and Computer Engineering, KTH Royal Institute of Technology; jpra@kth.se (J.P.C.); jamesgr@kth.se (J.G.)
* Correspondence: mamduhi@kth.se

**Abstract:** In the design of multi-loop Networked Control Systems (NCSs) wherein each control system is characterized by heterogeneous dynamics and associated with certain set of timing specifications and constraints, appropriate metrics need to be employed for the synthesis of control and networking policies to efficiently respond to the requirements of each control loop. Majority of the design approaches for sampling, scheduling and control policies include either time-based or event-based metrics to perform pertinent actions in response to the changes of the parameters of interest. We specifically focus in this article on Age-of-Information (AoI) as a recently-developed time-based metric and threshold-based triggering function as a generic event-based metric. As the NCS model, we consider multiple heterogeneous stochastic linear control systems that close their feedback loops over a shared-resource communication network. We investigate the co-design across the NCS, and discuss the pros and cons with AoI and ET approaches in terms of asymptotic control performance measured by linear-quadratic Gaussian (LQG) cost functions. In particular, sampling and scheduling policies combining AoI and stochastic event-triggered metrics are proposed. It is argued that pure AoI functions that generate decision variables solely upon minimizing the average age irrespective of control systems dynamics may not be able to improve the overall NCS performance even compared with pure randomized policies. Our theoretical analyses are successfully validated through several simulation scenarios.

**Keywords:** networked control systems; age-of-information; event-triggered sampling; scheduling architecture; resource constraint; asymptotic performance; estimation error

## 0. Introduction

Networked Control Systems (NCSs) generally refer to multiple dynamical systems controlled by possibly remotely located controllers with information exchange supported by a wired or wireless communication infrastructure. The applications of such systems are ranging from smart energy grids, autonomous driving, and industrial production, to healthcare, agriculture, and smart homes [1,2]. The two main layers of a networked system – control and communication – strongly influence each other and face heterogeneous and time-varying conditions, constraints and demands [3]. Hence, the efficient design of networked systems requires novel and integrated strategies that are responsive to the heterogeneity of the control systems and the real-time variations of individual layers, and at the same time possess flexibility and scalabiliby [4–6].

Considering state-of-the-art communication technology, there is a need for novel approaches to modeling, analysis and design of network protocols and control mechanisms capable of jointly supporting information exchange required to make decisions at the right component and at the right time. This is the basic motivation behind employing appropriate utility functions to coordinate the process of data exchange in a network of many dynamical users. Over the last two decades, there have been many attempts from the control and the communication communities to develop, evaluate and improve such utility functions compared to the conventional fixed-period and randomized data coordination approaches. Notions such as Value-of-Information (VoI), [7,8], Age-of-Information (AoI), [9,10], and Event-Triggered (ET) [11,12], are metrics that have been separately shown to be capable of coordinating information distribution taking into account the integrated and coupled context of NCSs. Traditionally, however, two rather distinct paths on addressing the NCS design have been followed: from the communication perspective, the focus mainly resulting in the design approaches that maximize the network throughput or minimize the end-to-end latency and jitter often ignoring the dynamics, requirements and characteristics of the sending and receiving entities and the specific data that are being transmitted [13–15]. From the control perspective, on the other hand, the major goal has been to maximize quality-of-control (QoC), and the communication network is usually abstracted as one or more maximum-rate and delay-negligible transmission channels with enough computation and functional capability to resolve contentions [16,17]. Hence, to fill this research void, it is essential to develop systematic and applicable co-design principles for NCSs that bring both QoC and QoS together by studying novel architectures that take into account the requirements, limitations, and tolerances of both network and control systems.

### *0.1. Contributions*

In this article, the goal is to propose an efficient co-design architecture for heterogeneous NCSs where the influence of both control and network systems are taken into account. Specifically, we study a sampling-scheduling-control co-design problem for stochastic NCSs comprised of multiple heterogeneous linear time-invariant (LTI) control systems. The sampling and control units reside at the control system layer and are designed distributedly, i.e., they are locally installed in every control loop and generate decision variables for their corresponding local control systems. The scheduling unit resides at the network layer and arbitrates the channel access in a centralized fashion, i.e., a unique scheduler coordinates the allocation process of the limited resources among the control loops to avoid contention and consequently data loss. We consider a realistic communication model in that the data packets that are not scheduled for immediate transmissions, if not updated by a newer data sample, are stored in a buffer for possible transmissions in future time instances. If a current sample is not successfully transmitted due to resource limitations, it is not discarded, and remains in the buffer to be either replaced by a newer sample, or transmitted with some delay whenever the communication resource is assigned to it. Therefore, end-to-end delay in our formulation is comprised of an inter-sampling duration induced by the local samplers and a network-induced delay due to the resource limitations. Performance of each local control system is asymptotically measured by the local linear-quadratic Gaussian (LQG) cost function and the overall asymptotic NCS performance is determined by the average sum of their local LQG costs. Note that the performance influenced by the resource constraints and the end-to-end transmission delays.

Motivated by the existing results for the design of control and communication systems, in this article we focus on two celebrated notions of utility metrics: AoI- and ET-based functions. We first discuss if these two design concepts may properly co-exist in a networked control scenario and study where each of them excels in terms of decision making efficiency. We evaluate them based on two crucial aspects: first, which class of policies result in lower local and overall cost values, and second, how much information is required for a policy maker to generate appropriate decisions. The first one, as explained earlier, is evaluated based on asymptotic LQG cost functions, while the second is basically judged based on that a policy maker needs less information, and distributed parts of networked system may not be willing to disclose too much information. Therefore, a desirable and applicable co-design architecture would result in sampling, scheduling and control decisions that jointly induce low local and overall control costs, while they require local or partially accessible information to generate their assigned decision variables at the expense of a viable level of computational complexity.

Under some mild assumptions on the information structures of the policy makers, we first show that the optimal control policy can be obtained independent of the sampling and scheduling policies. In fact, we show that the optimal controllers are of the certainty equivalence (CE) form, which technically means the optimal control inputs are identical as they would be obtained in the absence of the additive stochastic disturbances. This is really helpful as it provides a decomposition opportunity for the cross-layer co-design in the sense that the control law remains fixed for a variety of sampling and scheduling policies within the specified classes that satisfy those assumptions on their information structures. We then propose a joint sampling-scheduling co-design where the local samplers are ET and the centralized scheduler uses AoI-based prioritization for resource management. Considering the asymptotic average LQG cost function as the overall NCS performance metric, we show that the ET function is indeed a more efficient candidate for sampling, compared with its AoI counterpart, in sense of the asymptotic average sum of LQG functions, while AoI performs efficiently for governing the resource allocation process. We compare the performance of the AoI scheduling design with conventional random access resource scheduling and show that the AoI scheduling has the design flexibility to be appropriately adjusted to outperform the pure random access policy.

To the best of our knowledge, there is no result available in the literature that considers the co-design of control and communication systems with joint ET and AoI-based policies and compare their joint performance, although, both policies have separately been studied extensively from both control and communication perspectives.

### *0.2. Related Works*

Since the seminal work [18] many results have shown that event-based approach outperforms the conventional time-triggered and periodic schemes in the sense that they are capable of achieving the same control performance with significantly less usage of computation and communication resources [19–22]. The event-based approach is also widely studied in the context of NCSs [23–26], and it is shown that the event-based functions can be employed to efficiently govern the information sampling and scheduling processes taking into account not only the control requirements but also the communication conditions such as resource scarcity and channel properties [27–30].

Many researchers have demonstrated that ET policies preserve stability of NCSs despite updating the controllers less often. In [31], $\mathcal{L}_2$ stability of ET output feedback control is shown in the presence of network-induced delay. Stability of stochastic ET NCSs is also extensively studied, employing appropriate stochastic stability notions such as almost-sure and moment stability, with various sources of randomness such as model uncertainty, sensor noise, and erroneous channels [32–35]. Additionally, event-based medium access control (MAC) and contention resolution (CR) protocols for resource-limited or contention-based communication networks have been proposed in the literature, both in form of centralized and decentralized MAC and CR algorithms [36–39]. Centralized MAC and CR approaches are shown to be capable of fully resolving contentions yet at the expense of not being scalable as they require a huge volume of information exchange, while easily deployable decentralized event-based MAC and CR counterparts can substantially decrease contentions but not fully resolving them.

Design of optimal ET policies for either control and communication systems or cross-layer joint design has been an active area of research. The results suggest that finding global optimal event-based functions is often nontrivial, especially for multi-loop NCSs or more realistic models of communication networks [40–43]. The major difficulty lies behind the tight couplings and inter-layer dependencies between the distributed time-varying parameters of control and network systems, obliging to search for less computationally complex sub-optimal or approximative solutions. Network-induced delays are regarded as major coupling parameters in ET NCS design that depending on the model of sampling and communication network might possess different dynamic characteristics. In fact, delay affects the states of the control systems, and the states themselves affect the decision outcomes of the event-based policies, and those decisions also affect the network-induced delays [44]. Therefore, an optimal co-design needs to keep track of the network-induced delays which might not be feasible for stochastic networked systems.

The AoI metric, proposed in [45], has emerged to quantify the freshness of the received status updates at the estimator and has attracted significant attention from communication and networking communities. The AoI is defined as the time elapsed since the generation of the latest successfully received status update at the estimator. Several authors have studied the problem of minimizing some functions of AoI under different queuing and communication models [46–50]. While the works in [46–48] consider time averaged AoI, the authors in [49] consider minimizing the tail of the AoI, and the authors in [50] consider any non-decreasing and measurable function of AoI. Apart from studying the effects of communication scheduling on AoI, none of the above works consider estimation or control objectives for networked systems. Nonetheless, a general consensus is that, a lower AoI in an NCS may result in a lower estimation and control cost, because having access to fresher state information often improves the performance. However, only a handful of works considered the performance of the solutions proposed for AoI with respect to such costs. The authors in [51] have studied the minimum mean squared error problem with independent and identically distributed (i.i.d.) transmission delays for Wiener process estimation. They have shown that the estimation error is a function of AoI if the sampling decisions are independent of the observed Weiner process; otherwise, the estimation error is not a function of AoI. In [52], we studied a state estimator of a single-loop stochastic LTI system with i.i.d. transmission delays and derived the relation between AoI and the estimation error, assuming that the sampling decisions are independent of the observed states.

There has been an increasing interest recently from the control community to consider AoI utility functions due to their simpler evolution and characteristics compared to ET or VoI metrics. Despite some progress, however, there exist results suggesting that AoI-based approaches with the original linear formulation of AoI, may not be sensitive enough to dynamic changes of control systems and their QoC requirements [53,54]. In [55], various nonlinear functions of AoI are considered to be minimized instead of the conventional average linear AoI and it is shown that these variations of AoI utility functions can be beneficial to improve the control performance. The authors in [56] showed in a recent work that a discounted AoI-dependent monotonic function can be employed to optimally govern wireless network scheduling to maximize control performance over infinite horizon. Despite recent efforts reflected in the literature, there are still many challenges. Specifically, there is no result, to the best of our knowledge, on combined ET and AoI-based co-design across control systems and communication network layers.

### *0.3. Outline*

In the remainder of this article, the NCS model and the problem statement and are described in Section 1. The co-design architecture with CE controllers, sampling and scheduling policies is presented in Section 2. Performance analysis and comparisons with other co-design architectures are presented in Section 3. Simulation results are demonstrated in Section 4 and the concluding remarks are summarized in Section 5.

### *0.4. Notations*

We denote the expectation, conditional expectation, conditional probability, transpose and trace operators by $\mathsf{E}[\cdot]$, $\mathsf{E}[\cdot|\cdot]$, $\mathsf{P}[\cdot|\cdot]$, $[\cdot]^\top$, and $\mathrm{tr}(\cdot)$, respectively. A multivariate Gaussian distributed random vector $X$ with mean vector $\mu$ and covariance matrix $W \succ 0$ is denoted by $X \sim \mathcal{N}(\mu, W)$, where $A \succ B$ denotes $A - B$ is positive definite. The $Q$-weighted squared 2-norm of a column vector $X$ is denoted by $\|X\|_Q^2 \triangleq X^\top Q X$, and $\|X\|_2^2 \triangleq X^\top X$. A time-varying column vector $X_t^i$

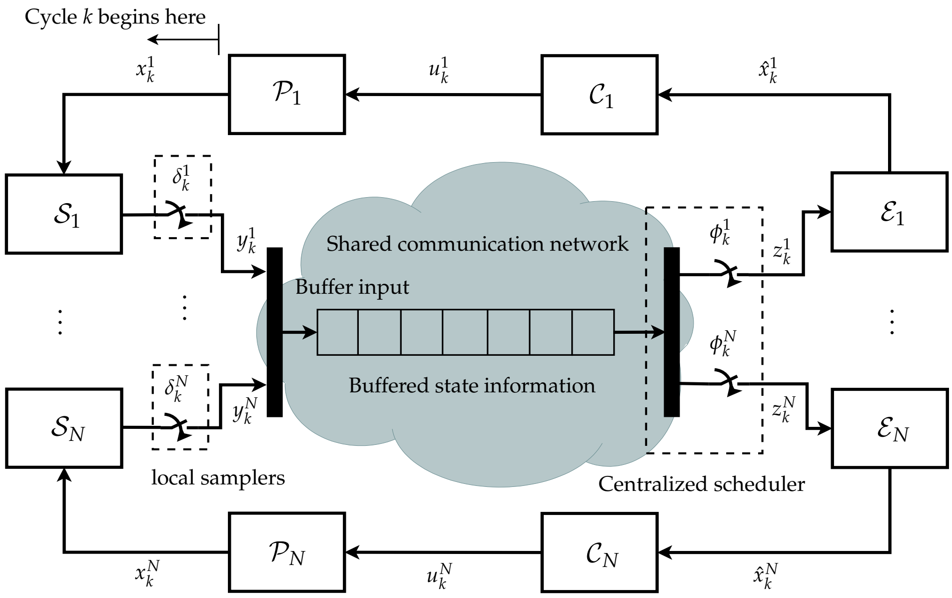


**Figure 1.** Multi-loop NCS with a shared communication network equipped with a data storage buffer.

includes an array of variables belonging to the sub-system indexed by $i$ at time $t$, while we define $X^i_{[t_1,t_2]} \triangleq \{X^i_{t_1}, X^i_{t_1+1}, ..., X^i_{t_2-1}, X^i_{t_2}\}$, and $X^i \triangleq \{X^i_0, X^i_1, ... \}$. For constant matrices, a subscript indicates the corresponding sub-system, and a superscript denotes matrix power. An optimal decision variable/policy $X$ is represented by $X^*$. The set of natural, real, non-negative integer, and non-negative real numbers are denoted by $\mathbb{N}$, $\mathbb{R}$, $\mathbb{N}_0$, and $\mathbb{R}_{\geq 0}$, respectively. For $n$-by-$m$-dimensional real space, we use the notation $\mathbb{R}^{n\times m}$.

## 1. NCS Model and Problem Description

### 1.1. NCS Model

We consider an NCS consisting of $N$ heterogeneous stochastic linear time-invariant (LTI) controlled dynamical processes that synchronously exchange their sensory information with their corresponding controllers via a common resource-limited communication network, see Figure 1. Each process $i \in \mathrm{N} \triangleq \{1, \ldots, N\}$ comprises of a plant $\mathcal{P}_i$, a noisy sensor $\mathcal{S}_i$, and a feedback control unit including a feedback controller $\mathcal{C}_i$ and an estimator $\mathcal{E}_i$. Each process $i \in \mathrm{N}$ is described as follows:

$$x^i_{k+1} = A_i x^i_k + B_i u^i_k + w^i_k, \tag{1}$$
$$y^i_k = x^i_k + v^i_k, \tag{2}$$

where $x^i_k \in \mathbb{R}^{n^i}$, $u^i_k \in \mathbb{R}^{m^i}$ and $y^i_k \in \mathbb{R}^{n^i}$ represent the state vector, control input and sensor measurement of the process $i$ at a time-step $k \in \mathbb{N}_0$, respectively. Constant matrices $A_i \in \mathbb{R}^{n^i \times n^i}$ and $B_i \in \mathbb{R}^{n^i \times m^i}$ describe the system matrix and input matrix, respectively, and we assume that each pair $(A_i, B_i)$ is controllable. To allow for heterogeneity, $A_i$ and $B_i$ matrices may differ for different processes and may also adopt different dimensions. The random processes $w^i_k \in \mathbb{R}^{n^i}$ and $v^i_k \in \mathbb{R}^{n^i}$ are, respectively, the exogenous disturbance acting on the process dynamics and the measurement noise. They are assumed to be Gaussian distributed independent random sequences with mutually i.i.d. realizations $w^i_k \sim \mathcal{N}(0, \Sigma_{w^i})$ and $v^i_k \sim \mathcal{N}(0, \Sigma_{v^i})$, $\forall k$ and $i \in \mathrm{N}$, where $\Sigma_{w^i} \succ 0$ and $\Sigma_{v^i} \succ 0$. The initial states $x^i_0$'s, $i \in \mathrm{N}$, are also presumed to be randomly selected from an arbitrary finite-moment distribution with mean $\mu_{x^i_0}$ and variance $\Sigma_{x^i_0} \succ 0$.

At every time-step $k$, the decision on whether the state measurement $y^i_k$ of sub-system $i$ is sent for transmission is taken by a local sampler $\mathcal{S}_i$ located at the sensor station. The sampling decision is assumed to be the outcome of a local sampling policy $\xi_i : \mathcal{I}_k^i \mapsto \{0,1\}$, where $\mathcal{I}_k^i$ represents the information available at $\mathcal{S}_i$ at time-step $k$ and will be formally defined later. The sampling decision outcome, denoted by the binary-valued variable $\delta_k^i$, is as follows:

$$\delta_k^i = \xi_i(\mathcal{I}_k^i) = \begin{cases} 1, & y_k^i \text{ sent to network for transmission,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

At every time-step $k$, those sub-systems which locally decided to update their corresponding controllers will forward their sensor measurements to the communication network. We assume that the communication network has capacity limitations such that not all $N$ sub-systems can simultaneously close their sensor-to-controller links at a time instance, i.e., if the network capacity at every single time-step is denoted by the constant $c \in \mathbb{N}$, the following resource constraint holds

$$1 \le c < N, \quad \forall\, k \in \mathbb{N}_0. \tag{4}$$

The communication network is assumed to be consisting of a queue to store the received data packets and a scheduling unit that decides which data packets are to be transmitted at each time-step. It should be mentioned that, transmissions of data from sensors to the buffer and from the buffer to the controllers are not subject to communication delay, i.e., if the sampler or scheduler decides on a sample being sent to the buffer or a buffered data sent to the controller, the transmissions are completed instantaneously. The scheduling decision at every time-step $k$ is assumed to be the outcome of a centralized resource allocation policy $\pi : \mathcal{I}_k^s \mapsto \{0,1\} \times \ldots \times \{0,1\} = \{0,1\}^c$, where $\mathcal{I}_k^s$ denotes the information available at the network scheduling unit at time-step $k$ which will be formally defined later, and $c$ is the constant capacity constraint. The scheduling decision associated with sub-system $i$ at time-step $k$ is denoted by the binary variable $\phi_k^i$ and is defined as

$$\phi_k^i = \pi(\mathcal{I}_k^s) = \begin{cases} 1, & \text{send the latest measurement of sub-system } i \text{ in the buffer to } \mathcal{E}_i, \\ 0, & \text{send nothing from sub-system } i \text{ to } \mathcal{E}_i. \end{cases} \tag{5}$$

The network queue buffers at most one data packet from each sub-system at every time instance. Hence, in case a new measurement belonging to a certain sub-system arrives at the queue, the fresher data packet replaces the formerly buffered data of that sub-system. The older data packet will be discarded. Therefore, for each sub-system, there is either no buffered data packet in the queue or there is one which is the latest measurement sent to the network by the local sampler. This means even the freshest data packet of a sub-system in the queue might contain the measurement that corresponds to a previous time-step.

When bandwidth is assigned to a certain sub-system, its freshest measurement in the queue will be forwarded to the corresponding control unit. The received state measurement by the control unit of a sub-system $i$ at a time-step $k$, denoted by $z_k^i$, might belong to a previous time $\bar{k} < k$ due to the communication delay imposed by the scheduling unit. Therefore, $z_k^i$ is determined as a function of the scheduling variable, as discussed in the following. Before that, we define the notion of AoI at the control unit in our NCS model, as follows:

**Definition 1.** *AoI at the control side of a sub-system $i \in \mathrm{N}$, at time-step $k \in \mathbb{N}_0$, is defined as $\Delta_k^i = k - \bar{k}^i$, where $y_{\bar{k}^i}^i$ is the latest received measurement by the estimator $\mathcal{E}_i$ up to time $k$, which confirms $\delta_{\bar{k}^i}^i = 1$.*

Assume that at a time-step $k$, $y_{\bar{k}^i}^i = y_{k-\Delta_k^i}^i$ is the freshest measurement of sub-system $i$ in the queue, which ensures $\delta_{k-\Delta_k^i}^i = 1$, and $\delta_{k-\Delta_k^i+1}^i = \ldots = \delta_k^i = 0$, because otherwise, $y_{k-\Delta_k^i}^i$ would have been replaced by a fresher measurement. In addition, this confirms that $\phi_{k-\Delta_k^i}^i = \phi_{k-\Delta_k^i+1}^i = \ldots = \phi_{k-1}^i = 0$, since otherwise, no data belonging to sub-system $i$ would be in the queue at time-step $k$. To conveniently denote this, we use the notation $\phi_k^i(k-\Delta_k^i)=1$ to express the time index of the freshest buffered measurement belonging to sub-system $i$ at time-step $k$ that is scheduled to be transmitted to the estimator $\mathcal{E}_i$. Hence, by $\phi_k^i(k-\Delta_k^i)=1$, we denote that $y^i_{k-\Delta_k^i}$ will be received by $\mathcal{E}_i$ at $k$. If no measurement of sub-system $i$ is scheduled to be transmitted at $k$, we simply write $\phi_k^i=0$. With this notation we declare two essential aspects of the information structure: 1) if a sample is scheduled for transmission, then the estimator knows which time instance the received measurement belongs to, and 2) receiving no measurement update might correspond to having no measurement sample of sub-system $i$ in the queue and not necessarily to resource limitations. It should be noted that if there is no data belonging to a sub-system $i$ buffered at a time-step $k$, then we certainly have $\phi_k^i=0$. In the other words, if the scheduler decides for $\phi_k^i=1$, then there exists exactly one buffered data packet of sub-system $i$ to be sent to its corresponding control unit. Therefore, $\phi_k^i=0$ might correspond to either having no measurement sample of sub-system $i$ in the buffer to forward or having not enough resources to schedule the available sample at that specific time. In the latter case $y^i_{k-\Delta_k^i}$ remains in the queue to be either serviced in future time-steps or replaced by a fresher sampled measurement. Finally, according to Definition 1, the information update at an estimator $\mathcal{E}_i$ can be stated as

$$z_k^i=\begin{cases} y^i_{k-\Delta_k^i} & \text{if} \quad \phi_k^i(k-\Delta_k^i)=1,\ \Delta_k^i\in[0,k], \\ \varnothing & \text{if} \quad \phi_k^i=0. \end{cases} \tag{6}$$

Note that the estimator $\mathcal{E}_i$ receives the current measurement sample $y_k^i$, only if $\phi_k^i(k)=1$, which ensures $\delta_k^i=1$ and $\Delta_k^i=0$. Depending on the information received at the estimator and the state estimate computed, the control input $u_k^i$ is assumed to be generated as the outcome of a causal mapping $\gamma_i:\tilde{\mathcal{I}}_k^i\mapsto\mathbb{R}^{m^i}$, where $\tilde{\mathcal{I}}_k^i$ represents the set of available information at the controller and will be formally defined later.

**Remark 1.** *In the absence of a measurement sample at the control side at a certain time $k$, i.e., if $\phi_k^i=0$, the estimator $\mathcal{E}_i$ may use the information contained in the sampling variable, i.e., knowing the outcome of $\delta_k^i$, and incorporate it in computing $\hat{x}_k^i$. This extra knowledge is known as the side-information contained in the sampling variable. In this article, we do not investigate the impact of the side-information when no measurement update is received by an estimator. As we will see later when we introduce the information structures, we assume that the control unit of a sub-system keeps the history of the sampling variables $\delta^i_{[0,k]}$, however, does not incorporate this side-information in computing $\hat{x}_k^i$ in the absence of a measurement sample. Incorporating side-information results in a nonlinear estimator and possibly non-tractable state estimator design problem, especially for threshold-based sampling policies in the presence of resource limitations. We assume that if no update is received at the estimator at some certain time-steps, then the estimator constructs $\hat{x}_k^i$ in a model-based fashion using the previous estimate $\hat{x}^i_{k-1}$.*

Depending on the sampling and scheduling decision variables $\{\delta_0^i,\dots,\delta_k^i\}$ and $\{\phi_0^i,\dots,\phi_k^i\}$, we can derive the dynamics of the AoI at the estimator $\mathcal{E}_i$. It is straightforward to derive the dynamics of $\Delta_k^i$, as functions of the sampling and scheduling variables:

$$\Delta_k^i=\sum_{t=1}^{k-r}\prod_{l=t}^{k-r}(1-\delta_l^i)+r, \qquad r=\sum_{t=1}^{k}\prod_{l=t}^{k}(1-\phi_l^i). \tag{7}$$

It can be seen from (7) that the AoI at the estimator depends on both sampling and scheduling decision outcomes.

Having the outcomes of the sampling and scheduling policies determined in (3) and (5), we can introduce the information sets $\mathcal{I}_k^i$ and $\mathcal{I}_k^s$, available, respectively, for the local sampler of sub-system $i$ and the centralized scheduling unit, as follows:

$$\mathcal{I}_k^i = \{\mathcal{I}^i_{\text{prim}}, \delta_0^i, \ldots, \delta_{k-1}^i, \phi_0^i, \ldots, \phi_{k-1}^i, z_0^i, \ldots, z_{k-1}^i\}, \tag{8}$$

$$\mathcal{I}_k^s = \cup_{i\in \mathrm{N}}\{\mathcal{I}^i_{\text{prim}}, \mathrm{N}_0^b, \ldots, \mathrm{N}_k^b, \delta_0^i, \ldots, \delta_k^i, \phi_0^i, \ldots, \phi_{k-1}^i, z_0^i, \ldots, z_{k-1}^i\}, \tag{9}$$

where, $\mathcal{I}^i_{\text{prim}} \triangleq \{A_i, B_i, \Sigma_{w^i}, \Sigma_{v^i}, \mu_{x_0^i}, \Sigma_{x_0^i}\}$, and $\mathrm{N}_k^b$ denotes the set of buffered state measurements at time-step $k$. Additionally, we introduce the set of available information for the estimator and controller of sub-system $i$ at time-step $k$:

$$\tilde{\mathcal{I}}_k^i = \mathcal{I}_k^i \cup \{\delta_k^i, \phi_k^i, z_k^i\} \cup \{u_0^i, \ldots, u_{k-1}^i\} = \{\mathcal{I}^i_{\text{prim}}, u_0^i, \ldots, u_{k-1}^i, \delta_0^i, \ldots, \delta_k^i, \phi_0^i, \ldots, \phi_k^i, z_0^i, \ldots, z_k^i\}. \tag{10}$$

Note that, with the information about sampling and scheduling variables in (8)-(10) and the expression for the AoI in (7), the sampler $\mathcal{S}_i$ is aware of the sequence $\Delta^i_{[0,k-1]}$, the controller $\mathcal{C}_i$ is aware of $\Delta^i_{[0,k]}$, and the centralized sampler has the knowledge of $\cup_{i\in\mathrm{N}}\{\Delta^i_{[0,k-1]}\}$.

Having the information set $\tilde{\mathcal{I}}_k^i$ introduced, we can construct the state estimate and compute the estimation error at the estimator of sub-system $i$. We denote the state estimate at the estimator of sub-system $i$ at time-step $k$ by $\mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]$, and define the corresponding estimation error as

$$\tilde{e}_k^i = y_k^i - \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]. \tag{11}$$

The dynamics of the estimation error $\tilde{e}_k^i$ can be obtained as

$$\begin{aligned}\tilde{e}_k^i &= y_k^i - \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i] = A_i x_{k-1}^i + B_i u_{k-1}^i + w_{k-1}^i + v_k^i - \mathsf{E}[A_i x_{k-1}^i + B_i u_{k-1}^i + w_{k-1}^i|\tilde{\mathcal{I}}_k^i] \\ &= A_i(x_{k-1}^i - \mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_k^i]) + w_{k-1}^i + v_k^i = A_i(\tilde{e}_{k-1}^i - v_{k-1}^i) + v_k^i + w_{k-1}^i.\end{aligned} \tag{12}$$

Note that, we can write $\mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_k^i] = \mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_{k-1}^i \cup \{\delta_k^i, \phi_k^i, z_k^i, u_{k-1}^i\}]$. Since the evolution of $x_{k-1}^i$ is independent of the parameters $\delta_k^i, \phi_k^i, z_k^i, u_{k-1}^i$, we then have $\mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_k^i] = \mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_{k-1}^i]$, which confirms (12). Assume now that the decision variables $\delta_k^i$ and $\phi_k^i$ are generated and $y^i_{k-\Delta_k^i}$, for any arbitrary $\Delta_k^i \in [\Delta_{k-1}^i + 1, k]$, is the latest received state measurement by the estimator $\mathcal{E}_i$ at time-step $k$, i.e., $\phi_k^i(k-\Delta_k^i) = 1$. Note that the realization of $\Delta_k^i$ is determined by the sampling and scheduling variables $\delta^i_{[\Delta^i_{k-1},k]}$ and $\phi^i_{[\Delta^i_{k-1},k]}$. We can compute the state estimate as

$$\begin{aligned}&\mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i] = \\ &\mathsf{E}[A_i^{\Delta_k^i} x^i_{k-\Delta_k^i} + A_i^{\Delta_k^i-1} B_i u^i_{k-\Delta_k^i} + \ldots + A_i B_i u_{k-2}^i + B_i u_{k-1}^i + A_i^{\Delta_k^i-1} w^i_{k-\Delta_k^i} + \ldots + w_{k-1}^i|\tilde{\mathcal{I}}_k^i] \\ &= A_i^{\Delta_k^i}\,\mathsf{E}[x^i_{k-\Delta_k^i}|y^i_{k-\Delta_k^i}] + A_i^{\Delta_k^i-1} B_i u^i_{k-\Delta_k^i} + \ldots + A_i B_i u_{k-2}^i + B_i u_{k-1}^i,\end{aligned}$$

where $\mathsf{E}[x^i_{k-\Delta^i_k} | y^i_{k-\Delta^i_k}]$ is the Minimum Mean-Square Estimate (MMSE) computed by a Kalman filter at the estimator side $\mathcal{E}_i$ given the received measurement $y^i_{k-\Delta^i_k}$, with the standard Kalman filter equations for a time $t$ at which the measurement sample $y^i_t$ is available, as

$$
\begin{aligned}
\mathsf{E}[x^i_t | y^i_t] &= \hat{x}^{i-}_t + K^i_t \left( y^i_t - \hat{x}^{i-}_t \right), \\
\hat{x}^{i-}_t &= A_i \, \mathsf{E}[x^i_{t-1} | \mathcal{I}^i_{t-1}] + B_i u^i_{t-1}, \\
K^i_t &= P^{i-}_t \left( P^{i-}_t + \Sigma_{v^i} \right)^{-1}, \\
P^{i-}_t &= \mathsf{E}\left[ \left( x^i_t - \hat{x}^{i-}_t \right) \left( x^i_t - \hat{x}^{i-}_t \right)^\top \right] = A_i P^{i-}_{t-1} A_i^\top + \Sigma_{w^i}, \\
P^i_t &= \mathsf{E}\left[ \left( x^i_t - \mathsf{E}[x^i_t|y^i_t] \right) \left( x^i_t - \mathsf{E}[x^i_t|y^i_t] \right)^\top \right] = P^{i-}_t - K^i_t \left( P^{i-}_t + \Sigma_{v^i} \right) K^{i\top}_t,
\end{aligned}
$$

where, $P^{i-}_t$ and $P^i_t$ denote, respectively, the *a priori* and the *a posteriori* estimation error covariances. Therefore, from (11), and using the equivalent expression

$$
y^i_k = A_i^{\Delta^i_k} x^i_{k-\Delta^i_k} + A_i^{\Delta^i_k - 1} B_i u^i_{k-\Delta^i_k} + \ldots + A_i B_i u^i_{k-2} + B_i u^i_{k-1} + A_i^{\Delta^i_k - 1} w^i_{k-\Delta^i_k} + \ldots + w^i_{k-1} + v^i_k,
$$

we conclude that

$$
\begin{aligned}
&\{\tilde{e}^i_k | \phi^i_k(k - \Delta^i_k) = 1\} \\
&= A_i^{\Delta^i_k} \left( x^i_{k-\Delta^i_k} - \mathsf{E}[x^i_{k-\Delta^i_k} | y^i_{k-\Delta^i_k}] \right) + A_i^{\Delta^i_k - 1} w^i_{k-\Delta^i_k} + \ldots + w^i_{k-1} + v^i_k \\
&= A_i^{\Delta^i_k} \left( \tilde{e}^i_{k-\Delta^i_k} - v^i_{k-\Delta^i_k} \right) + v^i_k + \sum_{r=1}^{\Delta^i_k} A_i^{r-1} w^i_{k-r}.
\end{aligned}
\tag{13}
$$

where, $\tilde{e}^i_{k-\Delta^i_k}$ is the MMSE error due to having access to $y^i_{k-\Delta^i_k}$. Otherwise, if $\phi^i_k = 0$, we use the model-based estimation error as in (12), wherein $\tilde{e}^i_{k-1}$ is not necessarily MMSE error.

### *1.2. Problem Description*

As discussed above, the time of generating a measurement sample and injecting it to the queue is determined by the sampler while the time of delivering that generated sample, if not discarded due to the arrival of a new sample, to the corresponding controller is determined by the network scheduler. Hence, the source-to-destination delay, i.e., the gap between the current time until the time a generated sample is received by the controller, depends on how the local samplers and the centralized scheduler policies are designed. The problem we tackle in this article is the co-design of sampling, scheduling and control policies $\{\xi_i, \pi, \gamma_i\}$. We discuss the optimal control policy, and then consider ET and AoI-based policies for the design of sampling and scheduling policies and study the effects of the combined architecture on the control performance which is correlated with the end-to-end delay. Performance comparisons are made according to the LQG index functions as the asymptotic cost metrics for each local sub-system, denoted by $J_i$:

$$
J_i = \lim_{T \to \infty} \frac{1}{T} \, \mathsf{E}\left[ x^{i\top}_T Q^2_i x^i_T + \sum_{k=0}^{T-1} x^{i\top}_k Q^1_i x^i_k + u^{i\top}_k R_i u^i_k \right],
\tag{14}
$$

where $Q_i^1, Q_i^2 \succeq 0$ and $R_i \succ 0$ are, respectively, the state and control input weight matrices of appropriate dimensions, and we assume each pair $(A_i, \sqrt{Q_i^1})$ is detectable, $\forall i \in \mathrm{N}$. The overall asymptotic NCS performance is measured by the average cost

$$J = \frac{1}{N}\sum_{i=1}^{N} J_i. \tag{15}$$

## 2. NCS Design

In this section, we first study the structural properties of the feedback controllers $\mathcal{C}_i$, $i \in \mathrm{N}$, and show that local control law $\gamma_i(\tilde{\mathcal{I}}_k^i)$ can be designed separately from the local sampling law $\xi_i(\mathcal{I}_k^i)$ and the scheduling law $\pi(\mathcal{I}_k^s)$. Afterwards, we discuss the combined design of the local sampling law and the network scheduling law and discuss which class of ET or AoI-based policies match the corresponding decision maker.

### 2.1. CE Control Law

Let us first make a crucial assumption on the sampling policy $\xi_i(\mathcal{I}_k^i)$:

**Assumption 1.** *The local sampling policies $\xi_i(\mathcal{I}_k^i)$'s are selected from the classes of control-input-independent sampling policies, i.e., $\delta_k^i$, $i \in \mathrm{N}$, are computed independent of the sequence of control inputs $\{u_0^i, \dots, u_{k-1}^i\}$.*

Assumption 1 does not result in a loss of generality w.r.t. the introduced information structure at the sampler, see (8) that indicated $\mathcal{I}_k^i$ does not contain any knowledge of control inputs $\{u_0^i, \dots, u_{k-1}^i\}$. This is crucial for the derivation of the optimal control policies, as will be discussed in Theorem 1.

**Theorem 1.** *Consider an NCS as described in (1)-(6), where each control system is steered at every time-step $k \in \mathbb{N}_0$ by a local sampler $\xi_i(\mathcal{I}_k^i)$ and a local plant controller $\gamma_i(\tilde{\mathcal{I}}_k^i)$ with $\mathcal{I}_k^i$ and $\tilde{\mathcal{I}}_k^i$ given in (8) and (10), respectively. If the local sampling policies are selected according to the Assumption (1), then the optimal control policy in sense of LQG given in (14) is CE, i.e.,*

$$\gamma_i^*(\tilde{\mathcal{I}}_k^i) = L_k^i \,\mathsf{E}\left[x_k^i | \tilde{\mathcal{I}}_k^i\right], \tag{16}$$

*where $L_k^i = -\left(R_i + B_i^\top P_{k+1}^i B_i\right)^{-1} B_i^\top P_{k+1}^i A_i$ is the optimal state feedback control gain.*

**Proof.** See Appendix A. □

**Remark 2.** *Showing that the optimal control law exists over the time horizon $[0, T]$, we can take the limit as $T \to \infty$ which results in having the asymptotic control gain $L_\infty^i = -\left(R_i + B_i^\top P_\infty^i B_i\right)^{-1} B_i^\top P_\infty^i A_i$, with $P_\infty^i = \lim_{k\to\infty} P_k^i$ being the asymptotic a posteriori estimation error covariance. We later show in Section 3.2 that, under appropriate sampling/scheduling co-design, $\forall i \in \mathrm{N}$, $P_\infty^i$ indeed exists and is not unbounded.*

**Remark 3.** *The result of Theorem 1 is in accordance with the existing results on the separation of control and sampling policies w.r.t. the LQG cost function, if the sampling law is independent of the control inputs. In fact, it is discussed in [22,57] that in the presence of control-input-dependent sampling policies, the separation between the sampling and control policies cannot generally be achieved. As it is shown in (18) and (19), the estimation error evolves independent of the control inputs, therefore, the sampling policies are allowed to be function of the estimation error without violating the results of Theorem 1.*

**Remark 4.** *Theorem 1 states that the optimal control law is of certainty equivalence form, however, the optimal control inputs $u_k^{i,*}$ are still computed based on the state estimate $\mathsf{E}\left[x_k^i | \tilde{\mathcal{I}}_k^i\right]$. As shown before, the estimation process depends on the sampling and scheduling policies $\xi_i(\mathcal{I}_k^i)$ and $\pi(\mathcal{I}_k^s)$, hence the sequence of control inputs*

*$\{u_0^{i,*}, \ldots, u_k^{i,*}\}$, $i \in \mathrm{N}$, is only optimal w.r.t. the given sampling and scheduling policies, and the control inputs are globally optimal only if sampling/scheduling policies are optimal. However, under any sampling policy that satisfies Assumption 1 and any scheduling policy, the optimal control law (16) remains CE.*

Now that the control law is characterized, we can derive the dynamics of the estimation error at the sampler, assuming that the local samplers are aware of the control law form in (16). This assumption is essential in the sense that the samplers do not need to have the knowledge of the control inputs $\{u_0^i, \ldots, u_{k-1}^i\}$ to compute the estimation error, and this coincides with the information structure (8). The estimation error at the sampler is defined as

$$e_k^i = y_k^i - \mathsf{E}[x_k^i | \mathcal{I}_k^i]. \tag{17}$$

From (8), and at time-step $k$, the sampler has the knowledge of the latest controller measurement update $z_{k-1}^i$. Let for any arbitrary $\Delta_{k-1}^i \in [0, k-1]$, $y_{k-1-\Delta_{k-1}^i}^i$ be the latest received state measurement by the estimator $\mathcal{E}_i$ at time-step $k-1$, i.e., $\phi_{k-1}^i(k-1-\Delta_{k-1}^i) = 1$. Then, similar to (13), we can compute the estimation error $e_k^i$ as

$$\begin{aligned}
&\{e_k^i | \phi_{k-1}^i(k-1-\Delta_{k-1}^i) = 1\} \\
&\quad = A_i^{\Delta_{k-1}^i+1}\left(x_{k-1-\Delta_{k-1}^i}^i - \mathsf{E}[x_{k-1-\Delta_{k-1}^i}^i | y_{k-1-\Delta_{k-1}^i}^i]\right) + A_i^{\Delta_{k-1}^i} w_{k-1-\Delta_{k-1}^i}^i + \ldots + w_{k-1}^i + v_k^i \\
&\quad = A_i^{\Delta_{k-1}^i+1}\left(\tilde{e}_{k-1-\Delta_{k-1}^i}^i - v_{k-1-\Delta_{k-1}^i}^i\right) + v_k^i + \sum_{r=1}^{\Delta_{k-1}^i+1} A_i^{r-1} w_{k-r}^i.
\end{aligned} \tag{18}$$

If $\phi_{k-1}^i = 0$, the estimation error at the sampler is, similar to (12), computed based on the model parameters, i.e.,

$$\{e_k^i | \phi_{k-1}^i = 0\} = A_i\left(\tilde{e}_{k-1}^i - v_{k-1}^i\right) + v_k^i + w_{k-1}^i. \tag{19}$$

Note the difference between $\tilde{e}_{k-1-\Delta_{k-1}^i}^i$ and $\tilde{e}_{k-1}^i$ in the expressions (18) and (19), where the former is the MMSE error due to having the measurement sample $y_{k-1-\Delta_{k-1}^i}^i$, while the latter is not MMSE as the estimator does not have access to $y_{k-1}^i$ at time-step $k-1$.

**Remark 5.** *Comparing (12) and (19), we conclude that if the estimator $\mathcal{E}_i$ does not receive any state measurement update at time $k-1$, i.e., $\phi_{k-1}^i = 0$, then $e_k^i = \tilde{e}_k^i$. It should, however, be noted that this equality is valid under the assumption that the estimator does not incorporate side information contained in the sampling variables to compute the state estimate.*

### *2.2. Co-design of Sampling and Scheduling Laws*

As the optimal control policy is shown to be CE, we now propose the sampling/scheduling co-design. We specifically focus on two common classes of policies, the ET and AoI utility functions, and study which class of policies is more suitable for sampling and which fits better to govern the scheduling process. Remind that the sampling is performed locally within each sub-system while the scheduler resides in the network layer and is performed in centralized fashion, see Figure 1.

We now introduce the ET and VoI functions used in the rest of this article. For the sampling policy, if the ET threshold-based approach is employed, then a sample of a local sub-system $i \in \mathrm{N}$ is generated and forwarded to the network buffer whenever the square norm of the corresponding sub-system's estimation error exceeds a positive random threshold $r_k^i$, i.e.,

$$\delta_k^{i,\,\mathrm{ET}} = \begin{cases} 1, & \text{if } \|e_k^i\|_2^2 > r_k^i, \\ 0, & \text{if } \|e_k^i\|_2^2 \le r_k^i, \end{cases} \tag{20}$$

where, the binary-valued $\delta_k^{i,\,\mathrm{ET}}$ indicates if a sample is forwarded for transmission or not based on the ET policy. The sequence of i.i.d. real-valued random thresholds $r_k^i \sim \exp(\mu_r^i)$, $k \in \mathbb{N}_0$ are assumed to be exponentially distributed, with $\mu_r^i \in \mathbb{R}_{\geq 0}$ being the rate parameter of the exponential distribution. Random threshold policy is a more general form of the threshold-based policies, hence the presented results in this article are easily extendable for ET deterministic threshold-based approach. Note that, the sampling policy (20) is in accordance with the Assumption 1. Remind that $e_k^i$ denotes the estimation error computed at time $k$ at the sampler side $\mathcal{S}_i$ (not at the controller side $\mathcal{C}_i$).

When AoI policy is employed for sampling, a state sample of a sub-system $i$ is sent to the communication network for transmission whenever the age of the latest received state information at the controller $\mathcal{C}_i$ exceeds a given threshold $\lambda^i \in \mathbb{N}_0$, i.e.,

$$\delta_k^{i,\,\mathrm{AoI}} = \begin{cases} 1, & \text{if} \quad \Delta_{k-1}^i > \lambda^i, \\ 0, & \text{if} \quad \Delta_{k-1}^i \leq \lambda^i. \end{cases} \tag{21}$$

Since age is a discrete variable taking only non-negative integer value, without loss of any generality, the threshold $\lambda^i$ is also assumed to be selected from non-negative integers.

As a comparative scenario, we also consider the periodic sampling, in which each sensor sample is sent for transmission at pre-defined instances of time and the inter-transmission time is determined by the constant time period $T_p \in \mathbb{N}$. Therefore, we have

$$\delta_k^{i,\,\mathrm{P}} = \begin{cases} 1, & \text{if} \quad k = nT_p + i, \quad n \in \mathbb{N} \\ 0, & \text{if} \quad \text{otherwise}. \end{cases} \tag{22}$$

As noticed in the expressions (20)-(22), we use the superscripts "ET", "AoI" and "P" to indicate that the sampling policies are ET, AoI-based, and periodic, respectively.

For the purpose of illustrations and ease of analysis, let us set the communication channel capacity to $c = 1$, i.e., at every time-step $k$ the scheduler allows only one state information to be forwarded to the corresponding controller, (see (4)). We already introduced $\mathrm{N}_k^b$ as the set of all sub-systems that have a state sample in the network buffer at time-step $k$. Note that, this state information might belong to the current time $k$ or to a previous time, hence, the buffered state measurements are not necessarily time-synchronized. For the AoI scheduling, we introduce the highest-age-first policy that in fact minimizes the average age of all sub-systems in $\mathrm{N}_k^b$. For a sub-system $i \in \mathrm{N}_k^b$, this can be expressed as

$$\mathsf{P}[\phi_k^{i,\,\mathrm{AoI}} = 1] = \begin{cases} 1, & \text{if} \quad \Delta_{k-1}^i > \Delta_{k-1}^j, \quad \forall j \in \mathrm{N}_k^b, \; j \neq i \\ \frac{1}{\eta_k}, & \text{if} \quad \underbrace{\Delta_{k-1}^i = \ldots = \Delta_{k-1}^l}_{\eta_k \text{ sub-systems}} > \Delta_{k-1}^j, \quad \forall j \in \mathrm{N}_k^b, \; j \neq i, \ldots, l \\ 0, & \text{if} \quad \exists j \in \mathrm{N}_k^b, \; \Delta_{k-1}^j > \Delta_{k-1}^i \end{cases} \tag{23}$$

where, $\eta_k$ denotes the number of sub-systems in $\mathrm{N}_k^b$ with the highest age at time-step $k$. We also express that if $i \notin \mathrm{N}_k^b$, then $\mathsf{P}[\phi_k^{i,\,\mathrm{AoI}} = 1] = 0$.

For pure random scheduling, we employ the common uniform randomization and we, therefore, have for all $i \in \mathrm{N}$

$$\mathsf{P}[\phi_k^{i,\,\mathrm{R}} = 1] = \begin{cases} \frac{1}{|\mathrm{N}_k^b|}, & \text{if} \quad i \in \mathrm{N}_k^b \\ 0, & \text{if} \quad i \notin \mathrm{N}_k^b \end{cases} \tag{24}$$

where, $|\cdot|$ represents the set cardinality operator and the superscript "R" in (24) stands for random scheduling policy.

| | | Scheduling | | | |
|---|---|---|---|---|---|
| | | ET | AoI | R | P |
| Sampling | ET | | * | * | |
| | AoI | | * | * | |
| | R | | | | |
| | P | | * | * | |

**Table 1.** Considered combinations of sampling/scheduling policies. The combinations designated with * are discussed either analytically or in simulations.

In the following, we analytically compare AoI-based vs. ET design for the decentralized sampling and will show (Section 3.3) that ET threshold-based sampling policy outperforms AoI-based counterpart if thresholds are appropriately designed. We, moreover, show that AoI sampling is in fact a more general form of periodic sampling with two differences, first, the transmission pattern may contain more than one fixed period, and second, the period(s) is a function of the number of sub-systems and the AoI thresholds. For the centralized scheduling process, we employ the AoI-based prioritizing policy of highest-age-first. In comparison with the pure random scheduling policy (Section 3.3), we show that the highest-age-first policy is not necessarily outperforming the pure random scheduling, if heterogeneity of sub-systems is not taken into account. We then propose the highest-age-first prioritization for the unstable sub-systems and show that this AoI-based policy is indeed capable of coordinating the communication resources more efficiently compared to the random scheduling, in sense of lower average sum of estimation errors of all sub-systems. We do not investigate the ET design as an applicable architecture for the scheduling policy since scheduling is a centralized process and decision making based on ET policies requires knowledge of real-time state information from all sub-systems which might not be preferred. It is, however, conjectured that if for certain small-size networked control scenarios ET policy might be favorable to be employed as the centralized scheduler, then it would even outperform AoI-based prioritizing scheduling due to its powerful capability of real-time prioritizing based on the current state of each single sub-system. The sampling/scheduling policy combinations that we address in this article, either analytically or in the simulation results, are summarized in Table 1.

## 3. Performance Analysis of the Joint Design

In this section, we propose two major co-design methodologies for the sampling and scheduling, where in the first method the sampling process is governed by an AoI threshold-based policy introduced in (21) and the scheduling is performed based on highest-age-first policy introduced in (23). In the second co-design the scheduling will be performed similarly based on the highest-age-first policy law in (23), while sampling process is controlled by the ET threshold-based policy shown in (20). We additionally consider periodic sampling policy and random scheduling, introduced in (22) and (24), respectively, as two conventional models for sampling and scheduling and provide comparisons, theoretically or numerically, with the proposed co-designs. For the purpose of brevity, we use the abbreviations "AoI/AoI", "ET/AoI", "ET/R", "AoI/R", "P/R", and "P/AoI" to denote the combined "sampling/scheduling" policy, see Table 1. To avoid confusion, it is worth reminding the difference between the AoI policies for decentralized sampling and centralized scheduling, see (21) and (23).

### *3.1. AoI Sampling and Scheduling Co-design*

In the AoI/AoI co-design architecture, the AoI sampling is performed locally at every sub-system's sensor station according to the threshold-based policy (21), while the AoI scheduling is done in centralized fashion according to the highest-age-first prioritizing policy (23). Assume a NCS is comprised of a set of stable and a set of unstable sub-systems, denoted respectively by $N_s$ and

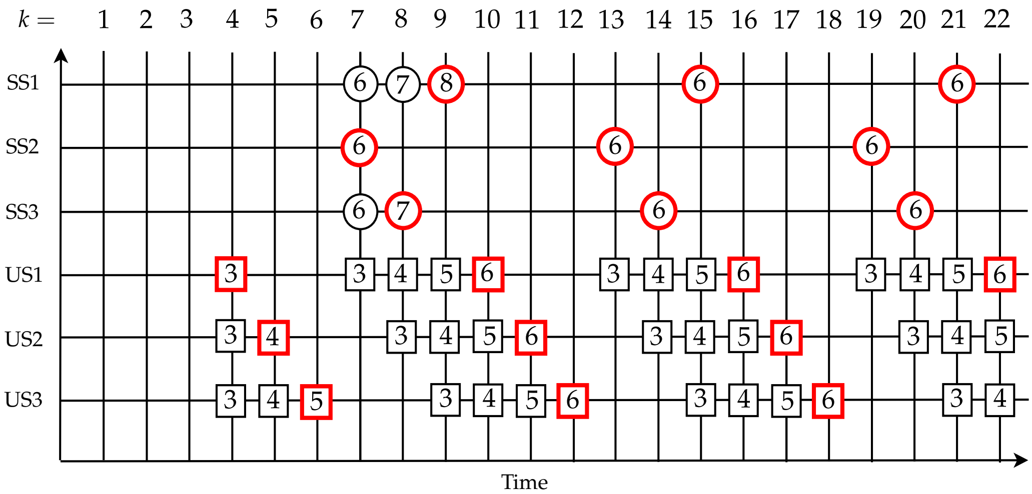


**Figure 2.** Sampling and scheduling patterns for an illustrative heterogeneous NCS of 3 stable (SS1, SS2, SS3) and 3 unstable (US1, US2, US3) sub-systems with AoI/AoI co-design architecture.

$\mathrm{N}_u$, where $\mathrm{N}_s \cup \mathrm{N}_u = \mathrm{N}$, and $N_s = |\mathrm{N}_s|$ and $N_u = |\mathrm{N}_u|$ indicate the number of stable and unstable sub-systems, respectively. Here, we study the asymptotic sampling and transmission patterns for the AoI/AoI co-design for different values of the deterministic thresholds $\lambda^i$, $i \in \mathrm{N}_s$ and $\lambda^j$, $j \in \mathrm{N}_u$.

Let $\lambda^i < N$ and $\lambda^j < N$. It is straightforward to conclude that each sub-system, either stable or unstable, will be scheduled for transmission once in every $N$ time-steps with a fixed unique pattern. Moreover, the sampler of each stable sub-system will send $N - \lambda^i$ number of samples to the buffer in the same cycle of $N$ time-steps, while unstable sub-systems send each $N - \lambda^j$ samples. We demonstrate this pattern for an illustrative example in below, and then summarize the concluding statements in the Proposition 1.

**Illustrative example:** Assume $N_s = N_u = 3$, $c = 1$, $\lambda^i = 5$ and $\lambda^j = 2$. Fig. 2 shows the sampling and transmission patterns of each sub-system, wherein, each circle (square) shows that a new measurement sample from a stable (unstable) sub-system is sent to the buffer. The red-bordered ones are the scheduled data packets and the numbers inside circles and squares denote the age of that corresponding sub-system at that time-step. According to (21), every unstable system $j$ (denoted by US1, US2, US3 in Fig. 2) sends a fresh sample to the buffer at any time-step $k$ at which $\Delta^j_{k-1} > 2$. Hence, no data packet is injected to the buffer before time-step $k = 4$, at which all three unstable sub-systems will send a measurement sample to the buffer (see Fig. 2). Note that, at time-step $k = 4$, the samplers decide based on $\Delta^j_3 = 3 > 2$. The same occurs for the stable sub-systems (denoted by SS1, SS2, SS3 in Fig. 2), hence, they all send their first measurement samples to the buffer at time-step $k = 7$, knowing that $\Delta^i_6 = 6 > 5$. Since at time-step $k = 4$, there are three data packets all with identical highest ages, the AoI scheduler selects one of the three randomly, i.e., $\eta_4 = 3$ (see the second argument of(23)). This randomization is repeated again at the next time-step $k = 5$ now with only two data packets with similar ages belonging to US2 and US3 (US1 remains silent for the next two time-steps). At $k = 6$, there is only one data packet in the buffer and it is certainly scheduled as there is no competition for the single transmission resource. At time-step $k = 7$, there are 4 data packets belonging to SS1, SS2, SS3, US1. The data packet belonging to US1 will not be scheduled for transmission because it has a lower age compared to the other three. For the remaining ones with the same ages, one will be scheduled for transmission randomly (e.g., SS2 as in Fig. 2). At $k = 8$, random selection is done between only SS1 and SS3 since the existing data packets of US1 and US2 entail lower ages. Finally, at $k = 9$, SS1 is certainly scheduled for transmission as it has the highest age among all the data packets in the buffer. From this time-step forward, the same pattern of transmissions is repeated without any randomization.

As it is also illustrated by the above example, we state the following proposition for which we omit the lengthy but straightforward proof:

**Proposition 1.** *For the sketched heterogeneous NCS scenario, if* $c = 1$, $\lambda^i, \lambda^j < N$, $\forall i \in \mathrm{N}_s$ *and* $\forall j \in \mathrm{N}_u$, *then the following statements hold, asymptotically:*

1. *each sub-system is scheduled a transmission once every* $N$ *time-steps.*
2. *stable and unstable sub-systems send, respectively,* $N - \lambda^i$ *and* $N - \lambda^j$ *fresh samples to the buffer during every* $N$ *time-steps.*
3. *if* $\lambda^i = \lambda^j = N - 1$, *then the AoI sampling is equivalent with the time-triggered sampling.*

Now assume that $\lambda^i, \lambda^j \geq N$. We can express similar statements as in Proposition 1 and conclude that both stable and unstable sub-systems successfully transmit in asymptotic regime, respectively, every $\lambda^i + 1$ and $\lambda^j + 1$ time-steps, and they send only one sample to the buffer per each successful transmission. This is then clear that this scenario is also equivalent with the periodic transmission with periods of $\lambda^i + 1$ and $\lambda^j + 1$ for stable and unstable sub-systems, respectively.

If $\lambda^i \geq N$ and $\lambda^j < N$, the transmission pattern for each sub-system $i \in \mathrm{N}_s$ is similarly periodic with time period of $\lambda^i + 1$, and only one measurement sample is sent to the buffer per each transmission. For sub-systems $j \in \mathrm{N}_u$, however, the transmission pattern is not periodic with a unique period, i.e., the inter-transmission times vary between every two consecutive successful transmissions, if $N \leq \lambda^i < 2N$. In fact it changes between $\lambda^j + 1$ and $N$ for each $j \in \mathrm{N}_u$. When the inter-transmission time is $\lambda^j + 1$, no data sample is discarded in between, while, when it is $N$, each sub-system $j$ sends $N - \lambda^j$ number of samples per each transmission. In addition, if $\lambda^i \to \infty$, then every $j \in \mathrm{N}_u$ successfully transmits every $\lambda^j + 1$ time-steps during which each sub-system $j$ sends $\max(1, N_u - \lambda^j)$ number of samples to the buffer. The same can be said for $\lambda^j \to \infty$. These statements can be numerically tested by, for example, setting $\lambda^i = 7$ in the depicted illustrative example in Fig. 2.

From the above discussions, we can make two crucial conclusions. First, AoI/AoI co-design policy governed by the AoI threshold-based sampling law (21) and AoI-based highest-age-first scheduling law (23) is not equivalent to the unique fixed periodic transmission policy, although, for some specific parameters, e.g., $\lambda^i = \lambda^j = N - 1$, they coincide. Second, all the possible transmission patterns are determined by the capacity constraint (4), the AoI thresholds $\lambda^i$ and $\lambda^j$ and the number of network sharing sub-systems, that are all constants. Hence, the resulting transmission patterns are insensitive w.r.t. the dynamics of stable or unstable sub-systems. We may design the AoI thresholds $\lambda^i$ and $\lambda^j$ differently for stable and unstable sets of sub-systems, however, they are assumed to be constant parameters and not adjusted by changing the dynamics[1].

### 3.2. ET Sampling and AoI Scheduling Co-design

In this section, we study the ET/AoI co-design architecture, where the sampling is locally performed according to the ET law (20) and the scheduling is centrally governed by the highest-age-first law in (23). For the clarity of analysis and illustrative purposes, we first assume that the ET thresholds in (20) are deterministic and constant, i.e., $r^i_k = r^i \in \mathbb{R}_{\geq 0}$. We discuss in the next section how to extend the performance results to the ET sampling with stochastic thresholds.

Since the estimation error $e^i_k$ is a random variable (see (18) and (19)), there is generally no fixed pattern for transmission of each sub-system when sampling is controlled by the ET law in (20). Hence,

---

[1] The discussions of the Section 3.1 can be extended to cover the scenarios that the AoI thresholds are not identical within the set of stable or unstable sub-systems, i.e., if $i, l \in \mathrm{N}_u$, then $\lambda^i \neq \lambda^l$. Although this leads to more complex transmission patterns, it does not contradict the crucial conclusions of this section, as summarized in the last part of the Section 3.1.

we study the asymptotic transmission rate for which we try to find mathematical expressions or bounds. To do that, we first compute the asymptotic sampling rate for an arbitrary sub-system $i \in \mathrm{N}$, as follows:

$$\lim_{k\to\infty} \mathsf{E}[\delta_k^i] = \lim_{k\to\infty} \mathsf{P}(\delta_k^i = 1) = \lim_{k\to\infty} \mathsf{P}(\|e_k^i\|_2^2 > r^i) \leq \frac{\lim_{k\to\infty} \mathsf{E}[\|e_k^i\|_2^2]}{r^i}, \tag{25}$$

where the inequality in (25) is obtained using Markov's inequality knowing $\|e_k^i\|_2^2$ is a non-negative random variable and $r^i$ is a non-negative constant. To provide more meaningful bound, we first state the following Lemma which essentially states that the dynamics of a stable sub-system's estimation error variance becomes insensitive to closing the feedback loop, asymptotically.

**Lemma 1.** *For any stable LTI stochastic control system modeled by (1), the estimation error variance is asymptotically bounded regardless of how often the feedback loop is closed.*

**Proof.** From (19), we can express the estimation error at the sampler's side, assuming that no transmission has taken place from the initial time until the current time $k$, i.e., $\phi_1^i = \phi_2^i = \ldots = \phi_{k-1}^i = 0$, which ensures $\Delta_{k-1}^i = k-1$, as follows:

$$e_k^i = A_i^k\left(\tilde{e}_0^i - v_0^i\right) + v_k^i + \sum_{r=1}^{k} A_i^{r-1} w_{k-r}^i = A_i^k\left(x_0^i - \mu_{x_0^i}\right) + v_k^i + \sum_{r=1}^{k} A_i^{r-1} w_{k-r}^i,$$

where the second equality holds since $\tilde{e}_0^i = y_0^i - \mathsf{E}[x_0^i] = x_0^i + v_0^i - \mu_{x_0^i}$. According to the last expression, $e_k^i$ is zero-mean, hence, we can compute the asymptotic estimation error variance as

$$\begin{aligned}
\lim_{k\to\infty} \mathsf{E}\left[e_k^{i\top} e_k^i\right] &= \lim_{k\to\infty} \mathsf{E}\left[\|e_k^i\|_2^2\right] = \lim_{k\to\infty} \mathsf{E}\left[\|A_i^k\left(x_0^i - \mu_{x_0^i}\right) + v_k^i + \sum_{r=1}^{k} A_i^{r-1} w_{k-r}^i\|_2^2\right] \\
&= \lim_{k\to\infty} \mathsf{E}\left[\|A_i^k\left(x_0^i - \mu_{x_0^i}\right)\|_2^2\right] + \lim_{k\to\infty} \sum_{r=1}^{k} \mathsf{E}\left[\|A_i^{r-1} w_{k-r}^i\|_2^2\right] + \Sigma_{v^i} \qquad (26) \\
&\leq \lim_{k\to\infty} \sum_{r=1}^{k} \|A_i\|_2^{2(r-1)} \Sigma_{w^i} + \Sigma_{v^i} \\
&= \frac{\Sigma_{w^i}}{1 - \|A_i\|_2^2} + \Sigma_{v^i}, \qquad (27)
\end{aligned}$$

where, to obtain (26), we used the mutual statistical independence of $x_0^i$, $v_k^i$ and $\{w_0^i, w_1^i, \ldots, w_k^i\}$, and the third expression is derived using the sub-multiplicative property of vector norms and also knowing that $\lim_{k\to\infty} A_i^k = 0$, since $A_i$ is Hurwitz. The final bound (27) is obtained knowing that the infinite series $\lim_{k\to\infty} \sum_{r=1}^{k} \|A_i\|_2^{2(r-1)} \Sigma_{w^i}$ is convergent since $\|A_i\|_2^2 < 1$, which completes the proof. □

Having Lemma 1, we can then re-express (25) for all stable sub-systems $i \in \mathrm{N}_s$, as follows:

$$\lim_{k\to\infty} \mathsf{P}(\delta_k^i = 1 | i \in \mathrm{N}_s) \leq \frac{1}{r^i}\left(\frac{\Sigma_{w^i}}{1 - \|A_i\|_2^2} + \Sigma_{v^i}\right). \tag{28}$$

It should be noted that the expression (25) holds for both stable and unstable sub-systems, while (28) is valid only for the former ones. Moreover, the bound (28) becomes trivial if the right hand side of the inequality is bigger than one, so this might also be seen as a rule to design the threshold $r^i$. It is clear that the higher $r^i$ is, the lower the transmission rate of stable sub-systems becomes, which is expected.

Lemma 1 does not apply to unstable sub-systems, hence, to derive similar upper-bound for sub-systems $j \in \mathrm{N}_u$, we compute $\lim_{k\to\infty} \mathsf{E}[\|e_k^j\|_2^2]$ according to the estimation error expression (18), as in the following:

$$\lim_{k\to\infty} \mathsf{E}[\|e_k^j\|_2^2|\Delta_{k-1}^j, j \in \mathrm{N}_u] =$$
$$\lim_{k\to\infty} \mathsf{E}\left[\left\|A_j^{\Delta_{k-1}^j+1}\left(\tilde{e}^j_{k-1-\Delta_{k-1}^j} - v^j_{k-1-\Delta_{k-1}^j}\right) + v_k^j + \sum_{r=1}^{\Delta_{k-1}^j+1} A_j^{r-1} w_{k-r}^j\right\|_2^2 \middle| \Delta_{k-1}^j, j \in \mathrm{N}_u\right] \leq$$
$$\Sigma_{v^j} + \lim_{k\to\infty} \sum_{r=1}^{1+\Delta_{k-1}^j} \|A_j^{r-1}\|_2^2 \Sigma_{w^j} + \lim_{k\to\infty} \|A_j\|_2^{2(\Delta_{k-1}^j+1)} \mathsf{E}\left[\left\|\tilde{e}^j_{k-1-\Delta_{k-1}^j} - v^j_{k-1-\Delta_{k-1}^j}\right\|_2^2 \middle| \Delta_{k-1}^j, j \in \mathrm{N}_u\right]. \quad (29)$$

Let us denote $\lim_{k\to\infty} \Delta_{k-1}^j = \Delta_\infty^j$. Having the highest-age-first scheduler (23), we can compute an upper-bound for (29) by evaluating the two disjoint cases $\lim_{k\to\infty} \|e_k^j\|_2^2 > r^j$ or $\lim_{k\to\infty} \|e_k^j\|_2^2 \leq r^j$, *almost surely*. If the first case holds, then according to (23) we know that a measurement sample belonging to the sub-system $j$ should exist in the buffer, and according to the Proposition 1 and the discussions afterwards, it yields that $\lim_{k\to\infty} \Delta_{k-1}^j \leq M^j \triangleq \max\{N, \lambda^j + 1\}$. Otherwise, if $\lim_{k\to\infty} \|e_k^j\|_2^2 \leq r^j$, *almost surely*, no measurement sample would be sent to the buffer asymptotically according to the ET sampling law (20), and then $\lim_{k\to\infty} \Delta_{k-1}^j > \max\{N, \lambda^j + 1\}$. Finally, knowing that $\tilde{e}^j_{k-1-\Delta_{k-1}^j}$ is the MMSE error computed by the Kalman filter having access to the measurement $y^j_{k-1-\Delta_{k-1}^j}$, we can rewrite (29) as

$$\lim_{k\to\infty} \mathsf{E}[\|e_k^j\|_2^2| \lim_{k\to\infty} \Delta_{k-1}^j \leq M^j, j \in \mathrm{N}_u] \leq \Sigma_{v^j} + \sum_{r=1}^{M^j+1} \|A_j^{r-1}\|_2^2 \Sigma_{w^j} + \lim_{k\to\infty} \|A_j\|_2^{2(M^j+1)} P^j_{k-1-\Delta_{k-1}^j}$$
$$= \Sigma_{v^j} + \sum_{r=1}^{M^j+1} \|A_j^{r-1}\|_2^2 \Sigma_{w^j} + \|A_j\|_2^{2(M^j+1)} P_\infty^j, \quad (30)$$

where, $P_\infty^j = \lim_{k\to\infty} P^j_{k-1-\Delta_{k-1}^j}$ is the asymptotic estimation error covanrince and will be obtained from the following algebraic Riccati equation

$$P_\infty^j = A_j \left(P_\infty^j - P_\infty^j (P_\infty^j + \Sigma_{v^j})^{-1} P_\infty^j\right) A_j^\top + \Sigma_{w^j}.$$

Finally, having (30), we can express the upper-bound in (25) for the unstable sub-systems $j \in \mathrm{N}_u$ as

$$\lim_{k\to\infty} \mathsf{P}(\delta_k^j = 1 | j \in \mathrm{N}_u) \leq \frac{\Sigma_{v^j} + \sum_{r=1}^{M^j+1} \|A_j^{r-1}\|_2^2 \Sigma_{w^j} + \|A_j\|_2^{2(M^j+1)} P_\infty^j}{r^j}. \quad (31)$$

Note that, if $\lim_{k\to\infty} \Delta_{k-1}^j \leq M^j$ holds for finite $M^j$, $P_\infty^j$ will also be finite as the Kalman filter receives state measurements asymptotically to compute the MMSE error. Even though the received measurements might not be fresh, the delay is finite and the Kalman filter algorithm converges.

### 3.3. Performance Comparisons

To conduct asymptotic performance analysis, we consider a heterogeneous NCS comprised of multiple stable and unstable sub-systems. Note that, for a NCS including all stable sub-systems asymptotic performance becomes independent of the sampling, scheduling and control policies due to the natural convergence of states. In fact, for a stochastic system of the form (1), the system states are expected to converge asymptotically to a bounded set around the origin where the set boundary is characterized by the moments of the primitive random variables, as we will show later in this section. Hence, to study the properties of the co-design we focus on the heterogeneous NCS scenario.

The overall asymptotic performance of the NCS is measured by the average cost functions $J$ introduced in (15). It can be seen from (A3) that the local LQG cost function $J_i$ can be minimized by the control law $\gamma^i$ in an inner optimization problem and then the residual cost becomes a function of the sampling law $\xi_i$. From (A8), we know that the residual local cost is a function of $\mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]$ and $\psi_k^i$ where the latter is shown in (A15) to be a function of the estimation error $\tilde{e}_t^i$ and its variance $P_t^i$, $t \geq k$. This should be remembered, however, that the resource constraint (4) does not allow the transmissions to be solely determined by the sampling law $\xi_i$, and therefore, the local cost $J_i$, from the overall perspective of control and network layers, becomes dependent also on the scheduling law $\pi$. The dependency appears in $\tilde{\mathcal{I}}_k^i$ which eventually affects both the estimate $\mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]$ and $\psi_k^i$. Hence, the optimal asymptotic cost function (15) should be minimized by both sampling and scheduling policies, which can be written as

$$J = \lim_{T\to\infty} \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\,\mathsf{E}\left[\min_{\delta^i_{[0,T-1]},\,\phi^i_{[0,T-1]}} \mathsf{E}\left[\mathsf{E}[x_0^i]^\top P_0^i\,\mathsf{E}[x_0^i] + \psi_0^i\right]\right] \tag{32}$$

wherein, $\mathsf{E}[x_0^i]$ is known *a priori*, and $\psi_0^i$ is a function of the estimation errors $\tilde{e}_k^i$ and their variances $P_k^i$, $k \in \{0,1,2,\ldots\}$, according to (A15).

Solving the optimization problem (32) is very challenging due to the coupling of the decision variables with $\psi_0^i$ through the end-to-end delay $\Delta_t^i$ and also the non-linear nature of the ET and AoI functions. The aim of this article is, therefore, to identify the appropriate class of policies for the sampling and scheduling that jointly result in an improved overall performance. Since the overall performance is a convex function of the estimation error according to (32), it is easier to study the asymptotic behavior of the estimation error of all sub-systems. In fact, if a certain co-design of sampling and scheduling policies results in a lower asymptotic average sum of estimation errors of all sub-systems, compared to another co-design, then it certainly results in a lower asymptotic overall cost (32) as well. Hence, for the performance analysis and make comparisons between different sampling/scheduling co-designs, we consider the following performance metric:

$$J_\infty^e = \lim_{k\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathsf{E}\left[\tilde{e}_k^{i\top}\tilde{e}_k^i\right] = \lim_{k\to\infty}\frac{1}{N}\left[\sum_{i=1}^{N_s}\mathsf{E}\left[\tilde{e}_k^{i\top}\tilde{e}_k^i\right] + \sum_{j=1}^{N_u}\mathsf{E}\left[\tilde{e}_k^{j\top}\tilde{e}_k^j\right]\right]. \tag{33}$$

Theorem 2 summarizes the second main result of this article on the appropriate sampling/scheduling co-design architecture. First, we define "non-trivial threshold-based functions", as follows:

**Definition 2.** *A threshold-based function with stochastic thresholds of the form (20) is said to be non-trivial if $\mathsf{P}[r_k^i \neq 0] > 0, \forall k$, and almost surely, $\mathsf{P}[r_k^i = \infty] = 0, \forall k$. For deterministic threshold-based functions of the form (21), we call the threshold-based function non-trivial if $\lambda^i \neq \{0,\infty\}$.*

**Theorem 2.** *Consider a heterogeneous NCS comprised of $N$ LTI stochastic sub-systems modeled as (1)-(2) from which $N_s$ sub-systems are stable and $N_u$ sub-systems are unstable. Let the network scheduler select only one sub-system per time-step to transmit its freshest state measurement in the buffer to the controller, i.e., $c = 1$ in (4). Then, for any non-trivial AoI sampling policy given in (21), there exists a non-trivial constant threshold ET sampling policy in form of (20) that asymptotically outperforms, in terms of (33), the AoI sampling policy, given that network scheduling policy obeys the AoI-based highest-age-first law in (24).*

**Proof.** See Appendix B. □

**Remark 6.** *Reminding the evolution of* $e_k^j$ *in* (18) *and* (19)*, it is clear that the estimation error at the sampler has a zero-mean but not normal distribution. For general square matrix* $A_j \in \mathbb{R}^{n^j \times n^j}$*, the asymptotic CDF of* $\|e_k^j\|_2^2$ *might not have an analytical form, but can be efficiently computed numerically. Indeed, the distribution of* $\|e_k^j\|_2^2$ *for general* $A_j$ *is determined by the distribution of its elements which are statistically dependent via the off-diagonal elements of* $A_j$*. For specific forms of* $A_j$*, however, the CDF has indeed an analytical form. If* $A_j$ *is a diagonal matrix, then the distribution of* $\|e_k^j\|_2^2$ *follows the sum of* $n_j$ *independent Gamma distributions which has an analytical CDF. For scalar systems, the distribution of* $\|e_k^j\|_2^2 = e_k^{j^2}$ *follows a single Gamma distribution.*

**Remark 7.** *Theorem* 2 *can be extended to the case that the thresholds* $r_k^i$ *and* $r_k^j$ *are stochastic, as in* (20)*. This would results in the Markov's inequality in* (25)*, and the expression* $\lim_{k\to\infty} \mathsf{P}(\|e_k^j\|_2^2 > r^j) = 1 - F^j_{\|e^j\|}(r^j)$ *not to be valid anymore due to the random nature of the thresholds. For stochastic thresholds, instead of Markov's inequality which holds for non-negative random variables, we can employ Chernoff bound which is a generalization of the Markov's inequality for real-valued random variables. In fact, if thresholds are stochastic, we can construct the new real-valued random variable* $\|e_k^j\|_2^2 - r_k^j$ *and find the upper-bound for it by applying the Chernoff bound. Further, we can write* $\lim_{k\to\infty} \mathsf{P}(\|e_k^j\|_2^2 - r_k^j > 0) = 1 - F^j_{\|e^j\|-r^j}(0)$*, where* $F^j_{\|e^j\|-r^j}$ *is now the asymptotic CDF of the constructed random variable* $\|e_k^j\|_2^2 - r_k^j$*. Note that,* $F^j_{\|e^j\|}(0) = 0$ *since* $\|e_k^j\|_2^2$ *is a non-negative random variable, however,* $F^j_{\|e^j\|-r^j}(0) > 0$ *since* $\|e_k^j\|_2^2 - r_k^j$ *is real-valued. The CDF* $F^j_{\|e^j\|-r^j}$ *may not have an analytical form, depending on the distributions of the random variables* $\|e_k^j\|_2^2$ *and* $r_k^j$*.*

In the following, we discuss that the pure AoI scheduling policy (23) may outperform the pure random transmission policy (24), but not always. In fact, we discuss that if the scheduler's highest-age-first prioritizing feature is applied first to the set of unstable sub-systems, then the AoI scheduling policy (23) certainly outperforms the pure random transmission policy. We define the highest-age-first policy for the unstable sub-systems similar to (23), with the exception that the law is applied, asymptotically, first on the set of unstable sub-systems and the resource is assigned to the unstable sub-system with the highest age, even if there are stable ones with higher age than the unstable ones.

**Corollary 1.** *For a fixed sampling policy, the highest-age-first threshold-based scheduling law* (23) *does not necessarily outperform the pure random scheduling policy* (24)*, asymptotically, in an NCS of heterogeneous stable and unstable control sub-systems sharing a capacity limited communication network. The AoI-based highest-age-first policy for unstable sub-systems, however, asymptotically outperforms the pure random scheduling policy* (24)*.*

**Proof.** As discussed before, the average sum of the estimation error variance of the set of stable sub-systems do not asymptotically change. Reminding (13), we see that the higher the age $\Delta_k^i$ is for unstable sub-systems, the larger the estimation error becomes. This is also true for the variance of the estimation error. Therefore, if a scheduling policy results in a higher transmission probability for the unstable sub-systems with the highest age, then the average sum of the estimation error variance will also be more reduced. According to (24), all sub-systems that have a data packet in the buffer are assigned identical probabilities of transmission $\frac{1}{|\mathrm{N}_k^b|}$, irrespective of their age or stability properties. For the same set of sub-systems with a packet in the buffer, the probability that the sub-system with the highest age, stable or unstable, successfully transmits is, according to (23), $\frac{1}{\eta_k}$, if there are $\eta_k \leq |\mathrm{N}_k^b|$ number of sub-systems all with identical highest age, which leads to $\frac{1}{\eta_k} \geq \frac{1}{|\mathrm{N}_k^b|}$. The equality occurs only if all sub-systems in the buffer have the similar age which is also the highest age. It should, however, be noted that if all the sub-systems that have the highest age are stable, then the unstable sub-systems in the buffer that may have relatively large age but not the highest are assigned with probability zero for successful transmissions, while this probability is $\frac{1}{|\mathrm{N}_k^b|}$ for the pure random scheduling policy that leads to a lower average sum of estimation error variance. With the modified prioritized highest-age-first policy for unstable sub-systems, however, the described problem can be easily considered in scheduling, and therefore, this policy always outperforms the pure random scheduling, for any fixed sampling policy. □

## 4. Numerical Evaluations

We consider different NCS setups with different number of stable and unstable sub-systems to numerically test the co-design architectures and compare with the common approaches. Number of sub-systems $N$ is chosen from the set $\{2, 4, 6, 8, 10\}$ with equal number of stable and unstable sub-systems. For the ease of interpretation, we choose scalar LTI sub-systems. The system matrices for stable and unstable sub-systems are selected to be 0.5 and 1.05, respectively. The system disturbance is modeled as $w_k^i \sim \mathcal{N}(0, 1)$, for all $i \in \mathrm{N}$, and $k \in \mathbb{N}_0$, and for the ease of illustrations we assume that measurements are noiseless. Each data point in the plots is generated by running the simulative setup for $10^6$ iterations. In the following we define the sampling and network scheduling strategies and the parameter values chosen for each scenario.

Sampling strategies

1. *Event triggering:* The sampler (sensor) samples the plant in each time-step, and if the value of the estimation error is greater than a threshold, then the sample is sent to the queue. The threshold is generated from an exponential distribution, and the mean of the distribution is chosen from the set $\{0, 0.1, 0.5, 1, 2 : 2 : 30\}$, where $2 : 2 : 30$ are integer values in $[2, 30]$ that are divisible by 2. We use a default setting where each sampler uses the same mean threshold.
2. *Period-n sampling:* Each sampler samples the plant periodically with period $n$.
3. *AoI sampling:* Each sampler samples the plant whenever the AoI at the sampler exceeds $N - 1$. The AoI at the sampler is equal to the AoI at the respective estimator from the previous time-step plus one.

*Network scheduling*

1. *Max AoI:* Under this scheduling policy, the network chooses the plant which has a packet in the queue and maximum AoI at the estimator, which is the highest-age-first policy.
2. *Pure randomized:* The network scheduler chooses a packet uniformly randomly from the queue.

In Figure 3, we plot the average estimation error variance (across all sub-systems) by varying the mean threshold. We observe that the estimation error variance is minimized for certain mean threshold values which increase with the number of sub-systems $N$. To understand this, note that when the thresholds are small, all the samplers will place a packet in the queue in almost every time-step. In this case, Max AoI does close to round-robin scheduling for the plants. Thus, plants with high or low estimation errors are treated rather indifferently leading to relatively high estimation error variance. The threshold values that attain minimum estimation error variances are such that the sub-systems with low estimation errors (usually the stable ones) do not contend for the network frequently as they do not exceed the thresholds frequently. This results in more often transmission of packets from sub-systems with high estimation errors, thus lowering the overall estimation error variance. In the following figures, we present the statistics of the event triggered sampling at mean thresholds that minimize the estimation error variance, that are marked by black circles in Figure 3.

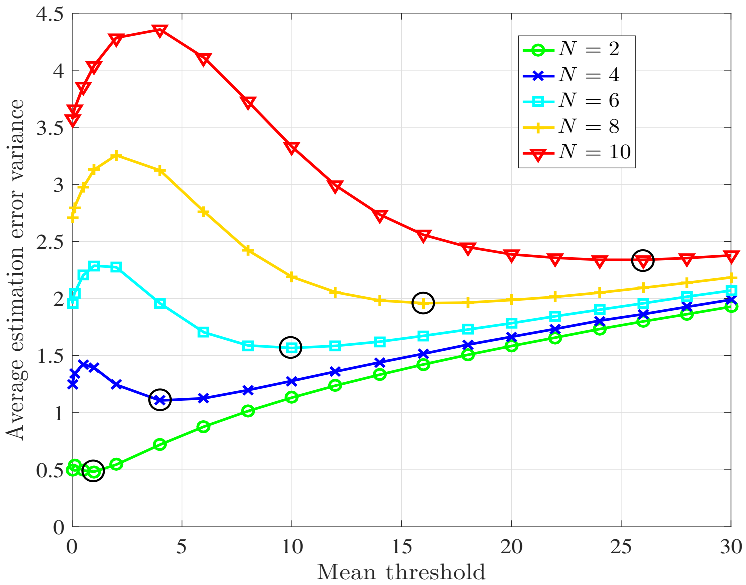


**Figure 3.** Average estimation error variance versus mean threshold under event triggering and Max AoI scheduling for different number of sub-systems $N$.

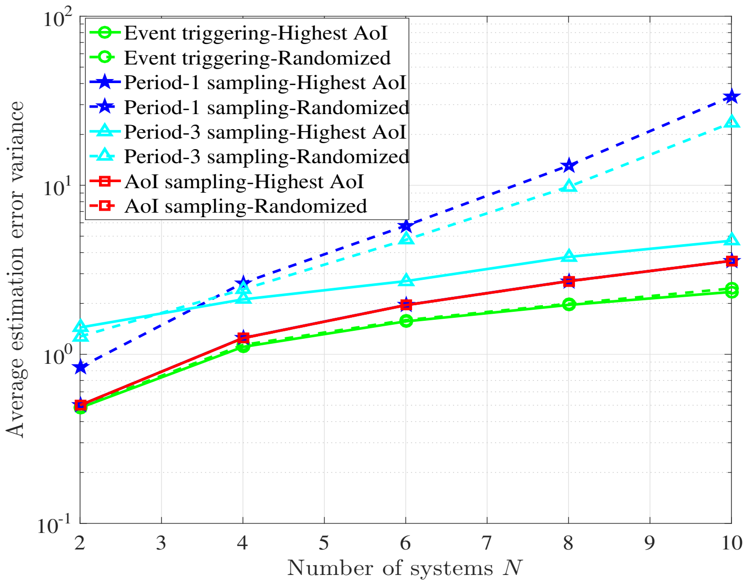


**Figure 4.** Comparison of estimation error variance for various sampling/scheduling architectures.

In Figure 4, we compare the estimation error variance achieved under different schemes. For event triggered sampling, we plot the minimum estimation error variance achieved over different thresholds for each $N$. We observe that event triggered sampling strategy obtains the lower estimation error variance, $30 - 40\%$ lower than that of AoI sampling when $N = 10$. While AoI sampling and period-1 sampling results in same variance, the sampling frequency of AoI sampling is much lower and equals $\frac{1}{N-1}$. Also, it can be observed that, in general, using Max AoI scheduling results in lower estimation error variance especially as the number of sub-systems grows.

In Figure 5, we compare the normalized total number of network transmissions that occur under different schemes. While period-1 sampling and AoI sampling result in a transmission in each time-step, event triggered results in transmissions 80% of the time for varying number of sub-systems. This is because, the queue remains empty 20% (on average) under event triggered sampling since only sub-systems with estimation errors greater than the threshold are allowed to place a packet in the queue. Therefore, event triggering not only provides lower estimation error variance, but also reduces the number of network transmissions.

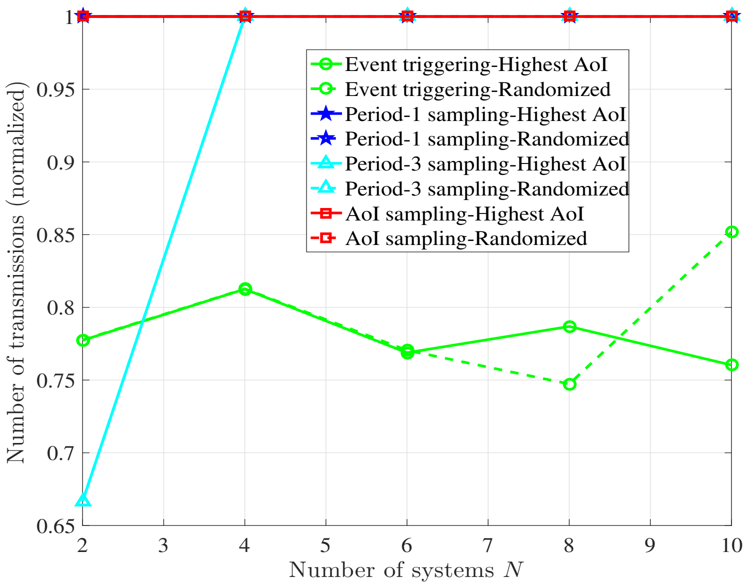


**Figure 5.** Normalized total number of network transmissions under different schemes.

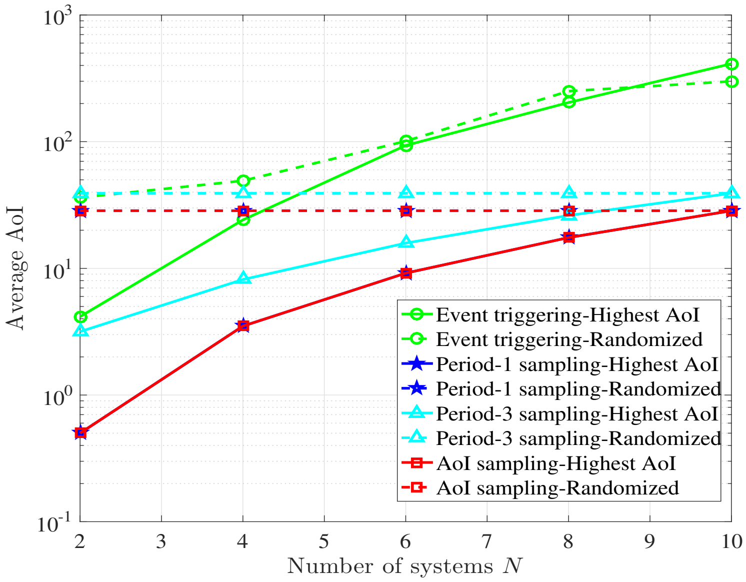


**Figure 6.** Average AoI achieved under different schemes.

In Figure 6, we compare the average AoI (averaged over all the estimators), achieved under different schemes. Since AoI sampling samples a plant based on its AoI at the estimator, this strategy results in the lowest average AoI. On the other hand, event triggering results in higher average AoI, as it samples based on estimation error, which increases non-linearly with AoI. Also, since AoI Max scheduling picks the plant with highest AoI and transmits its packet, this strategy results in lower average AoI across different sampling strategies. The main conclusion is, although ET sampling policy does not result in the lowest average AoI across the NCS, it results in the lowest achieved estimation error variance.

## 5. Conclusions

In this article, the major goal is to propose a co-design networked control architecture of sampling, scheduling and control for NCSs comprised of multiple heterogeneous LTI stochastic control systems that close their sensor-to-controller loops over a shared capacity-limited communication network. We first show that under mild assumptions on the information structure of each policy maker, the optimal control law is of certainty equivalence form. We then investigate various combinations of decentralized sampling and centralized scheduling architectures employing the applicable concepts of event-triggered and AoI utility functions. We analytically show that the event-triggered sampling is capable of asymptotically outperforming AoI sampling policy when the communication resources are limited, while we demonstrate AoI-based prioritizing scheduling may outperform the pure random scheduling policy under appropriate prioritization metric. To discuss the effectiveness of each co-design, we measure the overall NCS performance by the average sum of local LQG cost functions. Our theoretical analyses are successfully validated for the proposed co-designs and comparisons are made with conventional periodic and pure random access approaches through simulations on different NCS scenarios.

**Author Contributions:** "Conceptualization, M.H.M., J.P.C., J.G. and K.H.J.; methodology, M.H.M., J.P.C., J.G. and K.H.J.; software, M.H.M., J.P.C.; validation, M.H.M., J.P.C., J.G. and K.H.J.; formal analysis, M.H.M., J.P.C., J.G. and K.H.J.; investigation, M.H.M., J.P.C., J.G. and K.H.J.; resources, M.H.M., J.P.C., J.G. and K.H.J.; data curation, M.H.M. and J.P.C.; writing–original draft preparation, M.H.M. and J.P.C.; writing–review and editing, M.H.M., J.P.C., J.G. and K.H.J.; visualization, M.H.M. and J.P.C.; supervision, J.G. and K.H.J.; project administration, J.G. and K.H.J.; funding acquisition, J.G. and K.H.J. All authors have read and agreed to the published version of the manuscript."

**Funding:** This research was funded by Knut and Alice Wallenberg Foundation, the Swedish Strategic Research Foundation, and the Swedish Research Council.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| NCS | Networked Control System |
| AoI | Age-of-Information |
| ET | Event-Triggered |
| QoS | Quality-of-Service |
| QoC | Quality-of-Control |
| LQG | Linear-quadratic Gaussian |
| CE | Certainty Equivalence |

## Appendix A. Proof of Theorem 1

From the perspective of each local sub-system $i$, the expected local cost (14) changes depending on the $\mathcal{I}_k^i$-measurable sampling policy $\xi_i(\mathcal{I}_k^i)$ and the $\tilde{\mathcal{I}}_k^i$-measurable control policy $\gamma_i(\tilde{\mathcal{I}}_k^i)$. Using the law of total expectation[2], we can re-write (14) over the horizon $[0, T]$, as follows:

$$J_i(\xi_i, \gamma_i) = \frac{1}{T}\,\mathsf{E}\left[\mathsf{E}\left[x_T^{i\top} Q_i^2 x_T^i + \sum_{k=0}^{T-1} x_k^{i\top} Q_i^1 x_k^i + u_k^{i\top} R_i u_k^i \middle| \mathcal{I}_k^i\right]\right]. \tag{A1}$$

[2] Let a random variables $X$ be measurable w.r.t. to some $\sigma$-algebra $\mathcal{H}$, then we have $\mathsf{E}[\mathsf{E}[X|\mathcal{H}]] = \mathsf{E}[X]$.

From (8) and (10), we know that $\mathcal{I}_k^i \subset \tilde{\mathcal{I}}_k^i$, therefore, we can re-write (A1) by employing the general law of total expectation[3], as

$$J_i(\xi_i, \gamma_i) = \frac{1}{T}\,\mathsf{E}\left[\mathsf{E}\left[\mathsf{E}\left[x_T^{i\top} Q_i^2 x_T^i + \sum_{k=0}^{T-1} x_k^{i\top} Q_i^1 x_k^i + u_k^{i\top} R_i u_k^i \middle| \tilde{\mathcal{I}}_k^i\right] \middle| \mathcal{I}_k^i\right]\right]. \tag{A2}$$

Define the LQG cost-to-go at time-step $k$ as $V_k^i(\xi_i, \gamma_i) = x_T^{i\top} Q_i^2 x_T^i + \sum_{t=k}^{T-1} x_t^{i\top} Q_i^1 x_t^i + u_t^{i\top} R_i u_t^i$. We then have from (A2)

$$J_i^* = \frac{1}{T}\min_{\xi_i,\gamma_i} J_i(\xi_i, \gamma_i) = \mathsf{E}\left[\min_{\delta^i_{[0,T-1]}} \mathsf{E}\left[\min_{u^i_{[0,T-1]}} \mathsf{E}\left[V_0^i(\xi_i,\gamma_i)|\tilde{\mathcal{I}}_0^i\right] |\mathcal{I}_0^i\right]\right]. \tag{A3}$$

We, moreover, define the optimal stage cost $J_i^*(k)$ as follows:

$$J_i^*(k) = \min_{\delta^i_{[k,T-1]}} \mathsf{E}\left[\min_{u^i_{[k,T-1]}} \mathsf{E}\left[V_k^i(\xi_i,\gamma_i)|\tilde{\mathcal{I}}_k^i\right] |\mathcal{I}_k^i\right], \tag{A4}$$

which results in the compact form $J_i^* = \mathsf{E}[J_i^*(0)]$.

The LQG optimal cost-to-go at time-step $k+1$ has the following form:

$$V_{k+1}^{i,*} = \min_{u^i_{[k+1,T-1]}} \mathsf{E}\left[x_T^{i\top} Q_i^2 x_T^i + \sum_{t=k+1}^{T-1} x_t^{i\top} Q_i^1 x_t^i + u_t^{i\top} R_i u_t^i \middle| \tilde{\mathcal{I}}_{k+1}^i\right]. \tag{A5}$$

Knowing that $\tilde{\mathcal{I}}_k^i \subset \tilde{\mathcal{I}}_{k+1}^i$, we have from the law of total expectation that

$$\mathsf{E}\left[V_{k+1}^{i,*}|\tilde{\mathcal{I}}_k^i\right] = \min_{u^i_{[k+1,T-1]}} \mathsf{E}\left[x_T^{i\top} Q_i^2 x_T^i + \sum_{t=k+1}^{T-1} x_t^{i\top} Q_i^1 x_t^i + u_t^{i\top} R_i u_t^i \middle| \tilde{\mathcal{I}}_k^i\right]. \tag{A6}$$

Having (A6), we obtain

$$V_k^{i,*} = \min_{u^i_{[k,T-1]}} \mathsf{E}\left[x_k^{i\top} Q_i^1 x_k^i + u_k^{i\top} R_i u_k^i + V_{k+1}^{i,*}|\tilde{\mathcal{I}}_k^i\right]. \tag{A7}$$

Let us assume that $V_k^{i,*}$ can be expressed in the following form:

$$V_k^{i,*} \triangleq \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]^\top P_k^i\, \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i] + \psi_k^i, \tag{A8}$$

where $\psi_k^i$ is a control-input-independent expression. We will show later in this proof that (A8) is indeed an authentic assumption. According to (A8), we can re-express (A7) as follows:

$$V_k^{i,*} = \min_{u_k^i} \mathsf{E}\left[x_k^{i\top} Q_i^1 x_k^i + u_k^{i\top} R_i u_k^i + \mathsf{E}[x_{k+1}^i|\tilde{\mathcal{I}}_{k+1}^i]^\top P_{k+1}^i\, \mathsf{E}[x_{k+1}^i|\tilde{\mathcal{I}}_{k+1}^i] + \psi_{k+1}^i|\tilde{\mathcal{I}}_k^i\right]. \tag{A9}$$

---

[3] Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two sub-$\sigma$-algebras of a probability space with $\sigma$-algebra $\mathcal{H}$, and $X$ is defined on that probability space. If $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{H}$, then we have $\mathsf{E}[\mathsf{E}[X|\mathcal{H}_2]|\mathcal{H}_1] = \mathsf{E}[X|\mathcal{H}_1]$.

We have for the *a priori* state estimate that $\hat{x}^{i-}_{k+1} = \mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k]$. Since $\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k]$ is $\tilde{\mathcal{I}}^i_k$-measurable, we obtain

$$\begin{aligned}\mathsf{E}\left[\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_{k+1}]^\top P^i_{k+1}\,\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k]|\tilde{\mathcal{I}}^i_k\right] &= \mathsf{E}\left[\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_{k+1}]^\top|\tilde{\mathcal{I}}^i_k\right] P^i_{k+1}\,\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k] \\ &= \mathsf{E}\left[x^i_{k+1}|\tilde{\mathcal{I}}^i_k\right]^\top P^i_{k+1}\,\mathsf{E}\left[x^i_{k+1}|\tilde{\mathcal{I}}^i_k\right] = \hat{x}^{i-\top}_{k+1}P^i_{k+1}\hat{x}^{i-}_{k+1}, \end{aligned} \tag{A10}$$

where, for the first equality we use the conditional expectation property of $\mathsf{E}[XY|\mathcal{H}] = X\,\mathsf{E}[Y|\mathcal{H}]$ if $X$ is $\mathcal{H}$-measurable, and the second equality holds according to the law of total expectation knowing $\tilde{\mathcal{I}}^i_k \subset \tilde{\mathcal{I}}^i_{k+1}$. Similarly, it can be shown that $\mathsf{E}\left[\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k]^\top P^i_{k+1}\,\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_{k+1}]|\tilde{\mathcal{I}}^i_k\right] = \hat{x}^{i-\top}_{k+1}P^i_{k+1}\hat{x}^{i-}_{k+1}$. Define $\epsilon^i_{k+1} \triangleq \mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_{k+1}] - \mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_k] = \hat{x}^i_{k+1} - \hat{x}^{i-}_{k+1}$. It is straightforward to see $\epsilon^i_{k+1}$ is independent of $u^i_k$. From (A10), we conclude $\mathsf{E}\left[\epsilon^{i\top}_{k+1}P^i_{k+1}\hat{x}^{i-}_{k+1}\right] = \mathsf{E}\left[\hat{x}^{i-\top}_{k+1}P^i_{k+1}\epsilon^i_{k+1}\right] = 0$. Using the equivalence $\mathsf{E}[x^i_{k+1}|\tilde{\mathcal{I}}^i_{k+1}] = \epsilon^i_{k+1} + \hat{x}^{i-}_{k+1}$, together with knowing $\epsilon^i_{k+1}$ and $\psi^i_{k+1}$ are independent of control inputs, we can re-write (A9) as

$$\begin{aligned} V^{i,*}_k &= \min_{u^i_k}\mathsf{E}\left[x^{i\top}_k Q^1_i x^i_k + u^{i\top}_k R_i u^i_k|\tilde{\mathcal{I}}^i_k\right] + \min_{u^i_k}\mathsf{E}\left[(A_i\hat{x}^i_k + B_iu^i_k)^\top P^i_{k+1}(A_i\hat{x}^i_k + B_iu^i_k)|\tilde{\mathcal{I}}^i_k\right] \\ &+ \mathsf{E}\left[\epsilon^{i\top}_{k+1}P^i_{k+1}\epsilon^i_{k+1} + \psi^i_{k+1}|\tilde{\mathcal{I}}^i_k\right]. \end{aligned} \tag{A11}$$

Since the last term after the equality above is $u^i_k$-independent, finding the optimal control $u^{i*}_k$ is straightforward and can be obtained by setting the derivative of the first two terms in (A11) w.r.t. $u^i_k$ to zero, which results in

$$u^{i*}_k = -\left(R_i + B_i^\top P^i_{k+1}B_i\right)^{-1} B_i^\top P^i_{k+1}A_i\hat{x}^i_k.$$

Defining $L^i_k = -\left(R_i + B_i^\top P^i_{k+1}B_i\right)^{-1} B_i^\top P^i_{k+1}A_i$, (16) will be obtained. We still need to show that $\psi^i_{k+1}$ is indeed independent of control inputs. By plugging in (16) in the optimal cost-to-go (A11), and also using $x^i_k = \tilde{e}^i_k + \hat{x}^i_k - v^i_k$ (see (11)), we have

$$\begin{aligned} V^{i,*}_k &= \mathsf{E}\left[(\tilde{e}^i_k + \hat{x}^i_k - v^i_k)^\top Q^1_i(\tilde{e}^i_k + \hat{x}^i_k - v^i_k) + (L^i_k\hat{x}^i_k)^\top R_i(L^i_k\hat{x}^i_k)|\tilde{\mathcal{I}}^i_k\right] \\ &+ \mathsf{E}\left[(A_i\hat{x}^i_k + B_iL^i_k\hat{x}^i_k)^\top P^i_{k+1}(A_i\hat{x}^i_k + B_iL^i_k\hat{x}^i_k)|\tilde{\mathcal{I}}^i_k\right] + \mathsf{E}\left[\epsilon^{i\top}_{k+1}P^i_{k+1}\epsilon^i_{k+1} + \psi^i_{k+1}|\tilde{\mathcal{I}}^i_k\right] \\ &= \hat{x}^{i\top}_k\left(L^{i\top}_k R_iL^i_k + Q^1_i + (A_i + B_iL^i_k)^\top P^i_{k+1}(A_i + B_iL^i_k)\right)\hat{x}^i_k \\ &+ \mathsf{E}\left[\tilde{e}^{i\top}_k Q^1_i\tilde{e}^i_k|\tilde{\mathcal{I}}^i_k\right] - \mathsf{E}\left[v^{i\top}_k Q^1_i v^i_k\right] + \mathsf{E}\left[\epsilon^{i\top}_{k+1}P^i_{k+1}\epsilon^i_{k+1} + \psi^i_{k+1}|\tilde{\mathcal{I}}^i_k\right], \end{aligned} \tag{A12}$$

where (A12) is obtained noting that $\mathsf{E}\left[v^{i\top}_k Q^1_i\hat{x}^i_k|\tilde{\mathcal{I}}^i_k\right] = \mathsf{E}\left[\hat{x}^{i\top}_k Q^1_i v^i_k|\tilde{\mathcal{I}}^i_k\right] = \mathsf{E}\left[\tilde{e}^{i\top}_k Q^1_i\hat{x}^i_k|\tilde{\mathcal{I}}^i_k\right] = \mathsf{E}\left[\hat{x}^{i\top}_k Q^1_i\tilde{e}^i_k|\tilde{\mathcal{I}}^i_k\right] = 0$, and $\mathsf{E}\left[\tilde{e}^{i\top}_k Q^1_i v^i_k|\tilde{\mathcal{I}}^i_k\right] = \mathsf{E}\left[v^{i\top}_k Q^1_i\tilde{e}^i_k|\tilde{\mathcal{I}}^i_k\right] = \mathsf{E}\left[v^{i\top}_k Q^1_i v^i_k\right]$. Now, comparing (A12) with (A8), we conclude the two following statements:

$$P^i_k = L^{i\top}_k R_iL^i_k + Q^1_i + (A_i + B_iL^i_k)^\top P^i_{k+1}(A_i + B_iL^i_k), \tag{A13}$$

$$\begin{aligned} \psi^i_k &= \mathsf{E}\left[\tilde{e}^{i\top}_k Q^1_i\tilde{e}^i_k|\tilde{\mathcal{I}}^i_k\right] - \mathsf{E}\left[v^{i\top}_k Q^1_i v^i_k\right] + \mathsf{E}\left[\epsilon^{i\top}_{k+1}P^i_{k+1}\epsilon^i_{k+1} + \psi^i_{k+1}|\tilde{\mathcal{I}}^i_k\right] \\ &= \mathsf{E}\left[\sum_{t=k}^{T-1}\tilde{e}^{i\top}_t Q^1_i\tilde{e}^i_t + \tilde{e}^{i\top}_T Q^2_i\tilde{e}^i_T|\tilde{\mathcal{I}}^i_k\right] - \mathsf{E}\left[\sum_{t=k}^{T-1}v^{i\top}_t Q^1_i v^i_t + v^{i\top}_T Q^2_i v^i_T\right] + \mathsf{E}\left[\sum_{t=k+1}^{T}\epsilon^{i\top}_t P^i_t\epsilon^i_t|\tilde{\mathcal{I}}^i_k\right]. \end{aligned} \tag{A14}$$

From the definitions of $\tilde{e}_k^i$, and $\epsilon_k^i$, and using $u_k^i = L_k^i \,\mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i]$, we obtain the following:

$$\begin{aligned}\epsilon_k^i + \tilde{e}_k^i &= \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i] - \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_{k-1}^i] + x_k^i + v_k^i - \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_k^i] \\ &= x_k^i - \mathsf{E}[x_k^i|\tilde{\mathcal{I}}_{k-1}^i] + v_k^i \\ &= A_i(x_{k-1}^i - \mathsf{E}[x_{k-1}^i|\tilde{\mathcal{I}}_{k-1}^i]) + v_k^i + w_{k-1}^i \\ &= A_i(\tilde{e}_{k-1}^i - v_{k-1}^i) + v_k^i + w_{k-1}^i\end{aligned}$$

Knowing that $\mathsf{E}\left[\epsilon_k^{i\top} P_k^i \,\tilde{e}_k^i|\tilde{\mathcal{I}}_k^i\right] = 0$ and $\mathsf{E}\left[\tilde{e}_{k-1}^{i\top} v_{k-1}^i|\tilde{\mathcal{I}}_k^i\right] = \mathsf{E}\left[v_{k-1}^{i\top} v_{k-1}^i\right]$, we can write

$$\begin{aligned}\mathsf{E}\left[\epsilon_k^{i\top} P_k^i \epsilon_k^i|\tilde{\mathcal{I}}_k^i\right] + \mathsf{E}\left[\tilde{e}_k^{i\top} P_k^i \tilde{e}_k^i|\tilde{\mathcal{I}}_k^i\right] &= \mathsf{E}\left[(\epsilon_k^i + \tilde{e}_k^i)^\top P_k^i(\epsilon_k^i + \tilde{e}_k^i)|\tilde{\mathcal{I}}_k^i\right] \\ &= \mathsf{E}\left[(A_i(\tilde{e}_{k-1}^i - v_{k-1}^i) + v_k^i + w_{k-1}^i)^\top P_k^i(A_i(\tilde{e}_{k-1}^i - v_{k-1}^i) + v_k^i + w_{k-1}^i)|\tilde{\mathcal{I}}_k^i\right] \\ &= \mathsf{E}\left[\tilde{e}_{k-1}^{i\top} A_i^\top P_k^i A_i \tilde{e}_{k-1}^i|\tilde{\mathcal{I}}_k^i\right] - \mathsf{E}\left[v_{k-1}^{i\top} A_i^\top P_k^i A_i v_{k-1}^i\right] + \mathsf{E}\left[v_k^{i\top} P_k^i v_k^i\right] + \mathsf{E}\left[w_{k-1}^{i\top} P_k^i w_{k-1}^i\right] \\ &= \mathsf{E}\left[\tilde{e}_{k-1}^{i\top} A_i^\top P_k^i A_i \tilde{e}_{k-1}^i|\tilde{\mathcal{I}}_k^i\right] + \mathrm{tr}\left((P_k^i - A_i^\top P_k^i A_i)\Sigma_{v^i}\right) + \mathrm{tr}\left(P_k^i \Sigma_{w^i}\right).\end{aligned}$$

From the above expression, therefore, we obtain

$$\begin{aligned}\mathsf{E}\left[\sum_{t=k+1}^{T} \epsilon_t^{i\top} P_t^i \epsilon_t^i|\tilde{\mathcal{I}}_k^i\right] = \mathsf{E}\left[\sum_{t=k+1}^{T} \tilde{e}_{t-1}^{i\top} A_i^\top P_t^i A_i \tilde{e}_{t-1}^i|\tilde{\mathcal{I}}_k^i\right] + \sum_{t=k+1}^{T} \mathrm{tr}\left((P_t^i - A_i^\top P_t^i A_i)\Sigma_{v^i}\right) \\ + \sum_{t=k+1}^{T} \mathrm{tr}\left(P_t^i \Sigma_{w^i}\right) - \mathsf{E}\left[\sum_{t=k+1}^{T} \tilde{e}_t^{i\top} P_t^i \tilde{e}_t^i|\tilde{\mathcal{I}}_k^i\right].\end{aligned}$$

Now, defining $\tilde{P}_t^i = A_i^\top P_{t+1}^i A_i - P_t^i + Q_i^1$, we can rewrite (A14) and derive $\psi_k^i$ as follows:

$$\begin{aligned}\psi_k^i &= \mathsf{E}\left[\sum_{t=k}^{T-1} \tilde{e}_t^{i\top}(Q_i^1 + A_i^\top P_{t+1}^i A_i)\tilde{e}_t^i + \tilde{e}_T^{i\top} Q_i^2 \tilde{e}_T^i|\tilde{\mathcal{I}}_k^i\right] - \mathsf{E}\left[\sum_{t=k+1}^{T} \tilde{e}_t^{i\top} P_t^i \tilde{e}_t^i|\tilde{\mathcal{I}}_k^i\right] \\ &+ \sum_{t=k+1}^{T} \mathrm{tr}\left((P_t^i - A_i^\top P_t^i A_i - Q_i^1)\Sigma_{v^i}\right) - \mathrm{tr}\left(Q_i^2\Sigma_{v^i}\right) + \sum_{t=k+1}^{T} \mathrm{tr}\left(P_t^i\Sigma_{w^i}\right) \\ &= \mathsf{E}\left[\tilde{e}_k^{i\top} P_k^i \tilde{e}_k^i + \sum_{t=k}^{T-1} \tilde{e}_t^{i\top} \tilde{P}_t^i \tilde{e}_t^i + \tilde{e}_T^{i\top} Q_i^2 \tilde{e}_T^i|\tilde{\mathcal{I}}_k^i\right] + \sum_{t=k+1}^{T} \mathrm{tr}\left((P_t^i - A_i^\top P_t^i A_i - Q_i^1)\Sigma_{v^i}\right) \\ &- \mathrm{tr}\left(Q_i^2\Sigma_{v^i}\right) + \sum_{t=k+1}^{T} \mathrm{tr}\left(P_t^i\Sigma_{w^i}\right).\end{aligned} \tag{A15}$$

According to (12) and (13), $\tilde{e}_k^i$ is independent of the control inputs $u_t^i$, $t \le k$ and $k \in \mathbb{N}_0$. Therefore, $\psi_k^i$ expressed in (A15) is shown to be control-independent, and the proof in then complete.

## Appendix B. Proof of Theorem 2

Incorporating the scheduling decision, no matter which type of scheduling policy has generated it, we can rewrite (33) as follows:

$$\begin{aligned}\lim_{k\to\infty} \frac{1}{N}\sum_{i=1}^{N} \mathsf{E}\left[\tilde{e}_k^{i\top}\tilde{e}_k^i\right] = \lim_{k\to\infty}\frac{1}{N}\Bigg[\sum_{i=1}^{N_s}\Big(\mathsf{P}(\phi_k^i = 1)\,\mathsf{E}\Big[\|\tilde{e}_k^i\|_2^2|\phi_k^i = 1\Big] + \mathsf{P}(\phi_k^i = 0)\,\mathsf{E}\Big[\|\tilde{e}_k^i\|_2^2|\phi_k^i = 0\Big]\Big) \\ + \sum_{j=1}^{N_u}\Big(\mathsf{P}(\phi_k^j = 1)\,\mathsf{E}\Big[\|\tilde{e}_k^j\|_2^2|\phi_k^j = 1\Big] + \mathsf{P}(\phi_k^j = 0)\,\mathsf{E}\Big[\|\tilde{e}_k^j\|_2^2|\phi_k^j = 0\Big]\Big)\Bigg].\end{aligned} \tag{A16}$$

According to Lemma 1, for stable sub-systems, we know that the expression (26) converges to a constant value that depends only on $A_i, \Sigma_{v^i}, \Sigma_{w^i}$. The exact constant expression equality for non-scalar systems, however, is non-trivial to derive. For scalar systems, i.e., if $0 < A_i < 1$, then the inequality in (27) becomes equality. For non-scalar case, though, we can use Cauchy-Schwarz inequality to find a constant upper-bound. What we need for the proof of Theorem 2 is not the exact expression for $\lim_{k\to\infty} \mathsf{E}\big[\|\tilde{e}^i_k\|_2^2\big]$, but only knowing that $\lim_{k\to\infty} \mathsf{E}\big[\|\tilde{e}^i_k\|_2^2 | \phi^i_k = 1\big] = \lim_{k\to\infty} \mathsf{E}\big[\|\tilde{e}^i_k\|_2^2 | \phi^i_k = 0\big]$. This equality is clear from (26) since $\|A_i\|_2^2 < 1$ and this diminishes the role of time $k$ in the expression for the estimation error variance. This essentially concludes that, for the set of stable sub-systems, transmissions in asymptotic regime do not influence the estimation error variance, and hence, (A16) can be rewritten as

$$\lim_{k\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathsf{E}\left[\tilde{e}^{i\top}_k \tilde{e}^i_k\right] = \lim_{k\to\infty} \frac{1}{N} \Bigg[ \sum_{i=1}^{N_s} \mathsf{E}\left[\|\tilde{e}^i_k\|_2^2\right] + \sum_{j=1}^{N_u} \left( \mathsf{P}(\phi^j_k = 1)\, \mathsf{E}\left[\|\tilde{e}^j_k\|_2^2 | \phi^j_k = 1\right] + \mathsf{P}(\phi^j_k = 0)\, \mathsf{E}\left[\|\tilde{e}^j_k\|_2^2 | \phi^j_k = 0\right] \right) \Bigg]. \tag{A17}$$

The problematic term in the above expression that may lead to increase the asymptotic average estimation error variance is $\mathsf{E}\left[\|\tilde{e}^j_k\|_2^2 | \phi^j_k = 0\right]$ while $\|\tilde{e}^j_k\|_2^2 > r^i$. Hence, the aim of the co-design policy is to increase $\lim_{k\to\infty} \mathsf{P}(\phi^j_k = 1)$ which consequently leads to a decrease in $\lim_{k\to\infty} \mathsf{P}(\phi^i_k = 1)$. Simply, we would like to assign the transmission opportunities more often to the unstable sub-systems, asymptotically. In addition, we are interested in not only a successful transmission, but a successful transmission of a low-age state measurement. This means more frequent sampling and more frequent scheduling of unstable sub-systems, in probabilistic sense. To achieve this, we should first notice from the statements of the Proposition 1 and the discussions afterwards in Section 3.1 that, in the non-trivial AoI/AoI co-design architecture, the minimum sampling rate of stable or unstable sub-systems with the AoI threshold $\lambda^i$ is $\lim_{k\to\infty} \mathsf{P}(\delta^i_k = 1) = \frac{1}{\max\{N, \lambda^i + 1\}} = \frac{1}{M^i}$. To have a higher sampling rate for the ET sampling law compared to the AoI sampling law, we need to show

$$\lim_{k\to\infty} \mathsf{P}(\delta^i_k = 1 | i \in \mathrm{N}_s, ET/AoI) < \lim_{k\to\infty} \mathsf{P}(\delta^i_k = 1 | i \in \mathrm{N}_s, AoI/AoI). \tag{A18}$$

Hence, using (28), the inequality (A18) is satisfied if

$$\frac{1}{r^i} \left( \frac{\Sigma_{w^i}}{1 - \|A_i\|_2^2} + \Sigma_{v^i} \right) < \frac{1}{M^i}, \tag{A19}$$

which results in the following lower-bound for the ET thresholds for stable sub-systems:

$$r^i > M^i \left( \frac{\Sigma_{w^i}}{1 - \|A_i\|_2^2} + \Sigma_{v^i} \right). \tag{A20}$$

For unstable sub-systems, we need to show

$$\lim_{k\to\infty} \mathsf{P}(\delta^j_k = 1 | j \in \mathrm{N}_u, ET/AoI) > \lim_{k\to\infty} \mathsf{P}(\delta^j_k = 1 | j \in \mathrm{N}_u, AoI/AoI). \tag{A21}$$

We know $\lim_{k\to\infty} \mathsf{P}(\|e^j_k\|_2^2 > r^j) = 1 - F^j_{\|e^j\|}(r^j)$, where $F^j_{\|e^j\|}$ is the asymptotic cumulative distribution function (CDF) of the random process $\|e^j_k\|_2^2$ and $F^j_{\|e^j\|}(r^j)$ is the value of the asymptotic CDF at $r^j$. Hence, (A21) is satisfied if

$$1 - \frac{1}{M^j} > F^j_{\|e^j\|}(r^j). \tag{A22}$$

The CDF $F^j_{\|e^j\|}$ is a monotonically non-decreasing function w.r.t. $r^j$, hence, the lower $M^j$ is (i.e., either lower $N$ or lower $\lambda^j$), the ET thresholds for unstable sub-systems should also be decreased to asymptotically out-sample the AoI/AoI architecture, and vice-versa, which is an expected conclusion. Having (A19) and (A22) satisfied, it is ensured that, first, the asymptotic sampling rate of stable sub-systems is lower in the ET/AoI co-design compared to the AoI/AoI, and second, the sampling rate of unstable sub-systems is higher for the former approach. Hence, not only the probability that the unstable sub-systems transmit is higher for the ET/AoI compared to the AoI/AoI policy, but also the scheduled transmissions that are determined by the AoI-based highest-age-first policy in (23) have lower average age for the ET/AoI co-design. This means, $\lim_{k\to\infty} \mathsf{P}(\phi^j_k = 1 | j \in \mathrm{N}_u, ET/AoI) > \lim_{k\to\infty} \mathsf{P}(\phi^j_k = 1 | j \in \mathrm{N}_u, AoI/AoI)$ and therefore, lower asymptotic average estimation error variance in (A17). Finally, knowing that the asymptotic behavior of the stable set of sub-systems are independent of the sampling and scheduling policies, the asymptotic average estimation error variance in (A17) can be upper-bounded as follows:

$$\lim_{k\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathsf{E}\left[\tilde{e}^{i\top}_k \tilde{e}^i_k\right] \leq \frac{1}{N} \sum_{i=1}^{N_s} \left[\frac{\Sigma_{w^i}}{1-\|A_i\|_2^2} + \Sigma_{v^i}\right] + \frac{1}{N} \sum_{j=1}^{N_u} \left[\mathsf{P}(\phi^j_k = 1)\left(\Sigma_{v^j} + \sum_{r=1}^{N} \|A_j^{r-1}\|_2^2 \Sigma_{w^j} + \|A_j\|_2^{2N} P^j_\infty\right) + \mathsf{P}(\phi^j_k = 0) r^j\right]. \tag{A23}$$

The bound is trivial if $\|A_i\|_2^2 = 1$ or $P^j_\infty \to \infty$. The first one is avoided due to assuming $A_i$ is Hurwitz, and $P^j_\infty$ is bounded due to the fact that if $\lim_{k\to\infty} \|e^j_k\|_2^2 > r^j$ asymptotically, then $\Delta^j_k \leq M^j$ which means, in the worst case, there is one state information with bounded delay to construct the Kalman estimate. This completes the proof.

## References


1. Zhang, X.; Han, Q.; Yu, X. Survey on recent advances in networked control systems. *IEEE Transactions on Industrial Informatics* **2016**, *12(5)*, 1740–1752.
2. Park, P.; Coleri Ergen, S.; Fischione, C.; Lu, C.; Johansson, K. H. Wireless network design for control systems: A survey. *IEEE Communications Surveys & Tutorials* **2018**, *20(2)*, 978–1013.
3. Baillieul, J.; Antsaklis, P. J. Control and communication challenges in networked real-time systems. *Proceedings of the IEEE* **2007**, *95(1)*, 9–28.
4. Pajic, M.; Sundaram, S.; Pappas, G. J.; Mangharam, R. The wireless control network: A new approach for control over networks. *IEEE Transactions on Automatic Control* **2011**, *56(10)*, 2305–2318.
5. Liu, X.; Goldsmith, A. Cross-layer Design of Control over Wireless Networks. In *Advances in Control, Communication Networks, and Transportation Systems*; Editor1, F., Editor2, A., Eds.; Birkhäuser Boston: Boston, USA, 2005; pp. 111–136.
6. Park, P.; Marco, P. D.; Johansson, K. H. Cross-layer optimization for industrial control applications using wireless sensor and actuator mesh networks. *IEEE Transactions on Industrial Electronics* **2017**, *64(4)*, 3250–3259.
7. Molin, A.; Esen, H.; Johansson, K. H. Scheduling networked state estimators based on Value-of-Information. *Automatica* **2019**, *110*.
8. Tanaka, T.; Esfahani, P. M.; Mitter, S. K. LQG control with minimum directed information: Semidefinite programming approach. *IEEE Transactions on Automatic Control* **2018**, *63(1)*, 37–52.
9. Kosta, A.; Pappas, N.; Angelakis, V. Age-of-Information: A new concept, metric, and tool. *Foundations and Trends in Networking* **2017**, *12(3)*, 162–259.
10. Yates, R. D.; Kaul, S. K. The Age-of-Information: Real-time status updating by multiple sources. *IEEE Transactions on Information Theory* **2019**, *65(3)*, 1807–1827.

11. Wang, X..; Lemmon, M. D. Event-triggering in distributed networked control systems. *IEEE Transactions on Automatic Control* **2011**, *56(3)*, 586–601.
12. Lunze, J.; Lehmann, D. A state-feedback approach to event-based control. *Automatica* **2010**, *46(1)*, 211–215.
13. Wang, H.; Liao, C.; Tian Z. Providing quality of service over time delay networks by efficient queue management. *IEEE Conference on Local Computer Networks* **2011**, 275–278.
14. Zhang, Q.; Zhang, Y. Cross-layer design for QoS support in multihop wireless networks. *IEEE Proceedings* **2008**, *96(1)*, 64–76.
15. Rajandekar, A.; Sikdar, B. A survey of MAC layer issues and protocols for machine-to-machine communications. *IEEE Internet of Things Journal* **2015**, *2(2)*, 175–186.
16. Elia, N.; Eisenbeis, J. N. Limitations of linear control over packet drop networks. *IEEE Transactions on Automatic Control* **2011**, *56(4)*, 826–841.
17. Nayyar, N.; Kalathil, D.; Jain, R. Optimal decentralized control in unidirectional one-step delayed sharing pattern. *Annual Conference on Communication, Control, and Computing (Allerton)* **2013**, 374–380.
18. Åstrom, K. J.; Bernhardsson, B. Comparison of periodic and event based sampling for first-order stochastic systems. *Proceedings of IFAC World Congress* **1999**, 301–306.
19. Rabi, M.; Moustakides, G. V.; Baras, J. S. Adaptive sampling for linear state estimation. *SIAM Journal on Control and Optimization* **2012**, *50(2)*, 672–702.
20. Tabuada, P. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control* **2007**, *52(9)*, 1680–1685.
21. Heemels, W. P. M. H.; Johansson, K. H.; Tabuada, P. An introduction to event-triggered and self-triggered control. *IEEE Conference on Decision and Control* **2012**, 3270–3285.
22. Molin, A.; Hirche, S. On the optimality of certainty equivalence for event-triggered control systems. *IEEE Transactions on Automatic Control* **2013**, *58(2)*, 470–474.
23. Weerakkody, S.; Mo, Y.; Sinopoli, B.; Han, D.; Shi, L. Multi-sensor scheduling for state estimation with event-based stochastic triggers. *IEEE Transactions on Automatic Control* **2016**, *61(9)*, 2695–2701.
24. Blind, R.; Allgöwer, F. Analysis of networked event-based control with a shared communication medium: part II - slotted ALOHA. *IFAC Proceedings Volumes* **2011**, *44(1)*, 8830–8835.
25. Molin, A.; Hirche, S. A bi-level approach for the design of event-triggered control systems over a shared network. *Discrete Event Dynamic Systems* **2014**, *24*, 153–171.
26. Al-Areqi, S.; Görges, D.; Liu, S. Event-based networked control and scheduling co-design with guaranteed performance. *Automatica* **2015**, *57*, 128–134.
27. Brunner, F. D.; Antunes, D.; Allgöwer, F. Stochastic thresholds in event-triggered control: A consistent policy for quadratic control. *Automatica* **2018**, *89*, 376–381.
28. Mamduhi, M. H.; Molin, A.; Tolic, D.; Hirche, S. Error-dependent data scheduling in resource-aware multi-loop networked control systems. *Automatica* **2017**, *81*, 209–216.
29. Peng, C.; Zhang, J. Event-triggered output-feedback $\mathcal{H}_\infty$ control for networked control systems with time-varying sampling. *IET Control Theory & Applications* **2015**, *9(9)*, 1384–1391.
30. Miskowicz, M. Event-based sampling strategies in networked control systems. *IEEE Workshop on Factory Communication Systems* **2014**, 1–10.
31. Yu, H.; Antsaklis, P. J. Event-triggered output feedback control for networked control systems using passivity: Achieving $\mathcal{L}_2$ stability in the presence of communication delays and signal quantization. *Automatica* **2013**, *49(1)*, 30–38.
32. Hu, S.; Yue, D.; Xie, X.; Du, Z. Event-triggered $\mathcal{H}_\infty$ stabilization for networked stochastic systems with multiplicative noise and network-induced delays. *Information Sciences* **2015**, *229*, 178–197.
33. Mamduhi, M. H.; Molin, A.; Hirche, S. Stability analysis of stochastic prioritized dynamic scheduling for resource-aware heterogeneous multi-loop control systems. *IEEE Conference on Decision and Control* **2013**, 7390–7396.
34. Quevedo, D. E.; Gupta, V.; Ma, W.; Yüksel, S. Stochastic stability of event-triggered anytime control. *IEEE Transactions on Automatic Control* **2014**, *59(12)*, 3373–3379.
35. Cloosterman, M. B. G.; van de Wouw, N.; Heemels, W. P. M. H.; Nijmeijer, H. Stability of networked control systems with uncertain time-varying delays. *IEEE Transactions on Automatic Control* **2009**, *54(7)*, 1575–1580.
36. Kartakis, S.; Fu, A.; Mazo, M.; McCann, J. A. Communication schemes for centralized and decentralized event-triggered control systems. *IEEE Transactions on Control Systems Technology* **2018**, *26(6)*, 2035–2048.

37. Shao, C.;Leng, S.; Zhang, Y.; Fu, H. A multi-priority supported medium access control in vehicular Ad Hoc networks. *Computer Communications* **2014**, *39*, 11–21.
38. Vilgelm, M.; Mamduhi, M. H.; Kellerer, W.; Hirche, S. Adaptive decentralized MAC for event-triggered networked control systems. *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control* **2016**, 165–174.
39. Mamduhi, M. H.; Kneissl, M.; Hirche, S. Decentralized event-triggered medium access control for networked control systems. *IEEE Conference on Decision and Control* **2016**, 513–519.
40. Peng, C.; Yang, T. C. Event-triggered communication and $\mathcal{H}_\infty$ control co-design for networked control systems. *Automatica* **2013**, *49(5)*, 1326–1332.
41. Garone, E.; Sinopoli, B.; Goldsmith, A.; Casavola, A. LQG control for MIMO systems over multiple erasure channels with perfect acknowledgment. *IEEE Transactions on Automatic Control* **2012**, *57(2)*, 450–456.
42. Maity, D.; Mamduhi, M. H.; Hirche, S.; Johansson, K. H.; Baras, J. S. Optimal LQG control under delay-dependent costly information. *IEEE Control Systems Letters* **2019**, *3(1)*, 102–107.
43. Molin, A.; Hirche, S. On LQG joint optimal scheduling and control under communication constraints. *IEEE Conference on Decision and Control* **2009**, 5832–5838.
44. Mamduhi, M. H.; Baras, J. S.; Johansson, K. H.; Hirche, S. State-dependent data queuing in shared-resource networked control systems. *IEEE Conference on Decision and Control* **2018**, 1731–1737.
45. Kaul, S.; Gruteser, M.; Rai, V.; and Kenney, J. Minimizing age of information in vehicular networks. *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks* **2011**, 350–358.
46. Kaul, S.; Yates, R.; Gruteser, M. Real-time status: How often should one update? *Proceedings IEEE INFOCOM* **2012**, 2731–2735.
47. Yates, R. D.; Kaul, S. Real-time status updating: Multiple sources. *IEEE International Symposium on Information Theory Proceedings* **2012**, 2666–2670.
48. Yates, R. D. Lazy is timely: Status updates by an energy harvesting source. *IEEE International Symposium on Information Theory* **2015**, 3008–3012.
49. Champati, J. P.; Al-Zubaidy, H.; Gross, J. Statistical guarantee optimization for age of information for the D/G/1 queue. *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* **2018**, 130–135.
50. Sun, Y.; Uysal-Biyikoglu, E.; Yates, R. D.; Koksal, C. E.; Shroff, N. B. Update or wait: How to keep your data fresh. *IEEE Transactions on Information Theory* **2017**, *63(11)*, 7492–7508.
51. Sun, Y.; Polyanskiy, Y.; and Uysal-Biyikoglu, E. Remote estimation of the Wiener process over a channel with random delay. *IEEE International Symposium on Information Theory* **2017**, 321–325.
52. Champati, J. P.; Mamduhi, M. H.; Johansson, K. H.; Gross, J. Performance characterization using AoI in a single-loop networked control system. *IEEE Conference on Computer Communications Workshops* **2019**, 197–203.
53. Ayan, O.; Vilgelm, M.; Klügel, M.; Hirche, S.; Kellerer, W. Age-of-Information vs. Value-of-Information scheduling for cellular networked control systems. *Proceedings of the ACM/IEEE International Conference on Cyber-Physical Systems* **2019**, 109–117.
54. Sinha, D.; Roy, R. Scheduling status update for optimizing Age of Information in the context of industrial cyber-physical system. *IEEE Access* **2019**, 95677–95695.
55. Klügel, M.; Mamduhi, M. H.; Hirche, S.; Kellerer, W. AoI-penalty minimization for networked control systems with packet loss. *IEEE Conference on Computer Communications Workshops* **2019**, 189–196.
56. Ayan, O.; Vilgelm, M.; Kellerer, W. Optimal scheduling for discounted age penalty minimization in multi-loop networked control. *IEEE Consumer Communications & Networking Conference* **2020**, 1–7.
57. Rabi, M.; Ramesh, C.; Johansson, K. H. Separated design of encoder and controller for networked linear quadratic optimal control. *SIAM Journal on Control and Optimization* **2016**, *54(2)*, 662–689.