

Short Blocklength Process Monitoring and Scheduling: Resolution and Data Freshness

Stefan Roth¹, Graduate Student Member, IEEE, Ahmed Arafa², Member, IEEE, Aydin Sezgin¹, Senior Member, IEEE, and H. Vincent Poor³, Life Fellow, IEEE

Abstract—In cyber-physical systems (CPSs) and internet-of-things applications, various sensor-actuator pairs are deployed for control purposes which require timely online communication. The sensors are measuring information about the CPS, e.g., process systems, whereas the actuators are using the information to take control actions. These sensor-actuator pairs usually communicate via the same wireless medium and thus their transmissions need to be scheduled in time. When transmitting the process data, a *short blocklength* source-channel coding approach is employed to reduce data errors. We investigate the influence of the decision policy consisting of communication parameters and scheduling design on data freshness and accuracy of process monitoring systems. An age-of-information (AoI) metric is used to assess data timeliness, while the mean square error (MSE) is used to assess the precision of the predicted process values. We characterize the AoI and MSE with closed-form expressions for the blocklengths and accuracy levels, for special types of scheduling strategies, namely, round-robin and maximum-age scheduling. We optimize the coding strategies by showing an *achievability region* of AoI and MSE. Other priority-based scheduling policies are also investigated. It is shown that the maximum-age policy provides excellent results in terms of AoI, while priority-based scheduling performs better in terms of MSE.

Index Terms—Short blocklength, age-of-information (AoI), process monitoring, scheduling.

I. INTRODUCTION

TRADITIONALLY, communication between devices has been considered for asymptotically large blocklengths. However, in delay-sensitive applications conducted in 5G and beyond, such as in cyber-physical systems, internet-of-things,

Manuscript received April 19, 2021; revised August 17, 2021; accepted November 17, 2021. Date of publication December 10, 2021; date of current version July 12, 2022. This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through Germany's Excellence Strategy—EXC 2092 CASA under Grant 390781972 and in part by the U.S. National Science Foundation under Grant CCF-1908308. An earlier version of this paper was presented in part at the 2020 IEEE International Conference on Communications (ICC) [DOI: 10.1109/ICC40277.2020.9149010]. The associate editor coordinating the review of this article and approving it for publication was X. Chen. (Corresponding author: Stefan Roth.)

Stefan Roth and Aydin Sezgin are with the Institute of Digital Communication Systems, Ruhr University Bochum, 44801 Bochum, Germany (e-mail: stefan.roth-k21@rub.de; aydin.sezgin@rub.de).

Ahmed Arafa is with the Department of Electrical and Computer Engineering, The University of North Carolina at Charlotte, Charlotte, NC 28223 USA (e-mail: aarafa@uncc.edu).

H. Vincent Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2021.3132462>.

Digital Object Identifier 10.1109/TWC.2021.3132462

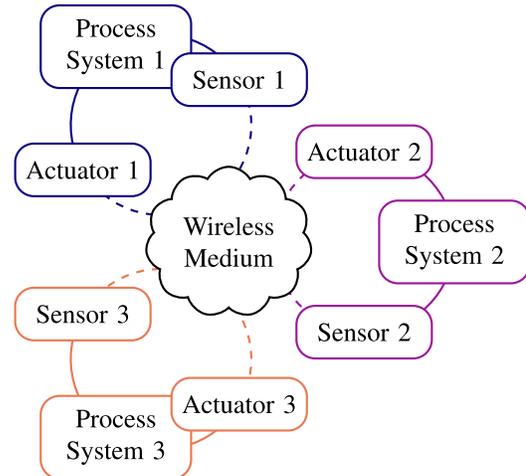


Fig. 1. Different sensor-actuator pairs are controlling multiple physical systems and communicating through the same wireless medium.

and networked control, it is crucial to transmit data in short packets. Thereby, the asymptotic capacity cannot be achieved due to the impacts of channel and source dispersions [2], [3]. The data throughput under the condition of short blocklengths has been investigated in [2]. When both source and channel coding are required in short blocklength scenarios, the relations between the throughput and distortion effects have been described in [3]. Our previous work [1] applies short blocklength source-channel coding results in the context of remote estimation of single-variate physical systems. In this work, we generalize these results to *multiple* and *multi-variate* systems that are remotely monitored for control purposes over a shared wireless medium.

In our scenario, there are multiple process systems, in which each system state value is measured by a sensor and can be controlled by an actuator. Hence, multiple sensor-actuator pairs are communicating via the same wireless medium, see Fig. 1. A central scheduler is employed to assign time slots to the sensor-actuator pairs for transmission. Hence, the central scheduler can decide about the transmission of sensor nodes with the corresponding waiting times, transmit time allocations and channel coding blocklengths. Thus, in general, the decision policy involves a source scheduling strategy, a sampling strategy and a coding strategy. However, practical implementations often rely on only a few scheduling policies, such as round-robin and maximum-age. Therefore, in addition

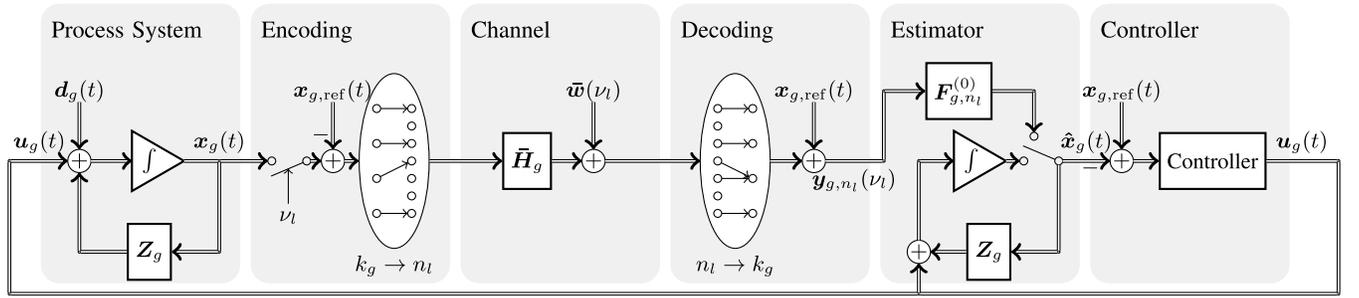


Fig. 2. Schematic description of the different parts of the system of sensor-actuator pair g when considering the transmission of packet l .

to the general formulation we will explicitly focus on these policies.

To evaluate the decision performance holistically, we jointly consider the *age-of-information (AoI)* metric [4] to assess timeliness of the estimates, and the *mean square error (MSE)* to assess the accuracy of the estimates. In more details, AoI is defined as the time elapsed since the generation time of the latest data measurement that has been received. Maintaining low AoI at the actuators is useful to diagnose errors and detect anomalies within the data quickly, such that immediate action can be taken. Compared to the MSE, the AoI is relevant in situations for which the system dynamics are hard to keep track of. In general, the two metrics are closely intertwined [5], [6].

With those metrics at hand, we focus on a Gauss-Markov process structure to describe the process systems. The transmission is done in packets encoded via a *short blocklength source-channel coding* framework and the received data are used to estimate and control the system state values until new measurements arrive. In general, with an increasing AoI, the estimate becomes less observant of the process, and the MSE increases as well. However, the MSE also depends on the blocklength used to describe each packet: (1) when the blocklength is large, very precise data can be transmitted (small MSE), but this also comes with longer transmission delays (large AoI); (2) when the blocklength is small, transmission delays become short (small AoI), whereas the data transmitted might be imprecise (large MSE). Therefore, there exists an inherent tradeoff between AoI and MSE, and blocklengths must be carefully designed to optimize them. On top of all this, the wireless medium contention among the different systems must be taken into consideration while computing the amounts of delay.

Contributions: We provide closed-form expressions for time-average AoI and MSE, in terms of the blocklengths and accuracy levels, for two specific scheduling policies, namely, round-robin and maximum-age scheduling. In addition, we present achievability regions describing a trade-off between AoI and MSE for different scheduling policies, which provides further insights on the intrinsic relationship between both parameters when remotely monitoring multiple processes in real-time. We use these results to numerically investigate if it is beneficial from a MSE-perspective to assign a higher priority to a sensor than to others. Our investigations suggest that excellent results in terms of AoI can be obtained by the

maximum-age policy, whereas assigning different priorities can be beneficial from an MSE perspective.

Related Works: Several works in the literature consider coding for AoI improvement with a single user, e.g., [7]–[15], of which [9] considers a study of short channel blocklengths in an AoI/delay minimization framework. Other works focus on estimation frameworks with AoI considerations for one user, e.g., [5], [6], [16]–[18]. The multiuser case has been investigated in [19]–[21]. The optimization of the AoI for robust control has been investigated in [22], whereas the impact of realistic communication assumptions on control has been investigated in [23]. The notion of updates with distortion has been studied in [24]. Our work differs from those works in the sense that we consider a joint source-channel short blocklength coding framework together with different scheduling policies to describe the relationship between AoI and MSE.

Notation: Matrices (vectors) are denoted in uppercase \mathbf{A} (lowercase \mathbf{a}). $(\cdot)^*$ and $(\cdot)^H$ refer to the complex-conjugate and conjugate-transpose, respectively. \mathbf{I}_k denotes the identity matrix of size $k \times k$; $e^{\mathbf{A}t}$ represents the matrix exponential; and $\text{eig}_i\{\mathbf{Q}\}$ indicates the i -th eigenvalue of a matrix \mathbf{Q} . $\|\cdot\|_2$ denotes the Euclidean norm. $\mathbb{E}[\cdot]$ is the expectation operator. $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy$ is the Q -function, which has an inverse of $Q^{-1}(x)$, whereas $\delta(x)$ refers to the Dirac-impulse.

II. SYSTEM MODEL

We consider the case of G sensor-actuator pairs communicating via the same wireless medium. Each sensor-actuator pair g consists of the following components (see Fig. 2):

- A *process system* that describes a vector-valued process to be controlled; the system state $\mathbf{x}_g(t)$ is controlled to become equal to a reference value $\mathbf{x}_{g,\text{ref}}(t)$, while being impacted by the system disturbance $\mathbf{d}_g(t)$.
- At the sensor, the system state vector is *encoded* such that it can be recovered after being transmitted through the channel. Here, we use a finite blocklength *joint source-channel coding* scheme.
- The transmission over a *channel* is typically imperfect. In this paper, we consider a noisy MIMO channel that leads to a distorted and delayed reception.
- At the receiver, *joint source-channel decoding* is used to decode the values received. Due to channel noise, the

decoded values remain imperfect and the probability of successful decoding is limited.

- The decoded variables are fed to an *estimator* to predict the system state of the following time instances until a new system state variable is decoded.
- A *controller* uses the estimated system state vector to generate an input signal $\mathbf{u}_g(t)$, such that $\mathbf{x}_g(t)$ approaches $\mathbf{x}_{g,\text{ref}}(t)$.

In the next subsections, we elaborate on these components and the *decision policy* that describes the connection between all G sensor-actuator pairs.

A. The Process System

The process system is based on a Gauss-Markov process, which contains a state value that is changing over time. We consider each process system to be a linear system that can be described by a state-space equation. In case a system is non-linear, it is often possible to create a linear approximation nearby an operating point, such that the same methods are applicable [25]. The model of each system g evolves similarly to a form in [26]:

$$\dot{\mathbf{x}}_g(t) = \mathbf{Z}_g \mathbf{x}_g(t) + \mathbf{u}_g(t) + \mathbf{d}_g(t), \quad (1a)$$

$$\mathbf{x}_g(0) = \mathbf{x}_{g,0}, \quad (1b)$$

where $\mathbf{u}_g(t)$ is the input sequence. $\mathbf{d}_g(t)$ represents the disturbance, which is assumed to be Gaussian distributed ($\mathbf{d}_g(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_g^D)$). \mathbf{Z}_g refers to the system matrix, whereas the state of the system is described by the column vector $\mathbf{x}_g(t)$ of dimension k_g . The initial state value at time $t = 0$ is given by $\mathbf{x}_{g,0}$ and is assumed bounded.

At the receiver, we build an estimator that tracks the state of the system. The estimated state $\hat{\mathbf{x}}_g(t)$ follows the same dynamics as $\mathbf{x}_g(t)$ (apart from the unknown disturbance $\mathbf{d}_g(t)$) and can be used for control purposes. A typical controller is modeled as a linear system consisting of proportional, integrative and differential components [27], and designed such that the system state $\mathbf{x}_g(t)$ should reach the reference value $\mathbf{x}_{g,\text{ref}}(t)$ after a certain time. The objective of the integrative component is to reach the reference value $\mathbf{x}_{g,\text{ref}}(t)$ even in the case of a constant disturbance $\mathbf{d}_g(t)$ with a good precision, whereas the objective of the proportional and differential components is to reach $\mathbf{x}_{g,\text{ref}}(t)$ faster. For a given control signal sequence $\mathbf{u}_g(\mu)$ for all times μ , the system state at time t can be described explicitly by the equation of motion given by the following solution of (1):

$$\mathbf{x}_g(t) = e^{\mathbf{Z}_g t} \mathbf{x}_{g,0} + \int_0^t e^{\mathbf{Z}_g(t-\mu)} (\mathbf{u}_g(\mu) + \mathbf{d}_g(\mu)) d\mu, \quad t \geq 0. \quad (2)$$

Each sensor is regularly transmitting the corresponding system state value to its corresponding actuator.

B. Transmission Times and Decision Policy

We assume that a central scheduler decides which sensor is transmitting the next packet after the previous transmission has finished, when such sensor will be sampled, and how many codeword symbols are used to encode the packet. Packet l is

sampled at time ν_l by sensor g_l , encoded into a codeword of length n_l , and (if the transmission is successful) delivered at time D_l . We focus on signal-independent sampling, in which the sampling times $\{\nu_l\}$ are independent of the processes being monitored. For a given channel blocklength n_l , we model the time b_l required for transmission of the received packet l through the channel by the following linear model:

$$b_l = \alpha n_l + \beta, \quad (3)$$

where α represents the symbol duration and β refers to an extra channel-induced delay, which also includes a protocol overhead. If the reception is successful, we have $D_l = \nu_l + b_l$, and the actuator immediately feeds back an acknowledgement to the central scheduler. Otherwise, a negative acknowledgement is fed back. We assume the acknowledgements to be immediately available at the central scheduler, which uses the feedback to schedule upcoming transmissions.

The central scheduler might also decide to add a waiting time before sampling the scheduled sensor. Such waiting time is described by W_{l+1} , i.e., $\nu_{l+1} = D_l + W_{l+1}$. This means that the source *scheduling strategy* is described by $\pi \triangleq (g_1, g_2, \dots)$, whereas the *sampling strategy* is given by $f \triangleq (\nu_1, \nu_2, \dots)$, and the *coding strategy* is given by $c \triangleq (n_1, n_2, \dots)$. When combining these, the *decision policy* is given as $p = (\pi, f, c)$. Throughout this paper, we are focusing on *stationary* decision policies [28]. A decision policy is stationary if the distribution does not change over time, which consists of the AoI and blocklength of the most recently received packet at each actuator and the AoI, blocklength, and sensor-actuator pair of an potentially ongoing transmission.

Denote the set of all successfully received packets by \mathcal{L}_{SUC} . Then, the set of indices of successfully decoded measurements at actuator g is described by $\mathcal{L}_g = \{l : g_l = g\} \cap \mathcal{L}_{\text{SUC}}$. The latest received packet at time t therefore has the index

$$l_g(t) = \max \{l \in \mathcal{L}_g, D_l \leq t\}.$$

The time interval, in which a packet $l_g(t)$ is used for system state estimation depends on the delivery time of packet $l_g(t)$ and of the next packet that will be successfully received at the same receiver g . These are

$$\begin{aligned} \underline{D}_{l_g(t)} &= D_{l_g(t)}, \\ \overline{D}_{l_g(t)} &= \min \{D_{l'} : l' \in \mathcal{L}_g, D_{l'} > D_{l_g(t)}\}. \end{aligned}$$

The values of the variables introduced here mainly depend on the parameters used within the data transmission scheme.

C. Transmissions With a Short Blocklength Source-Channel Coding Scheme

At time ν_l , the sample value $\mathbf{x}_g(\nu_l)$ is encoded into n_l transmit signals $\bar{\mathbf{x}}_i(\nu_l)$, $1 \leq i \leq n_l$ for transmission, which takes place through a multiple-input multiple-output (MIMO) channel. The i th received signal is given by

$$\bar{\mathbf{y}}_i(\nu_l) = \bar{\mathbf{H}}_g \bar{\mathbf{x}}_i(\nu_l) + \bar{\mathbf{w}}_i(\nu_l), \quad (4)$$

where $\bar{\mathbf{H}}_g$ denotes the $O_R \times O_T$ channel matrix of the sensor-actuator pair g , and $\bar{\mathbf{w}}_i(\nu_l)$ represents zero-mean Gaussian noise with $\mathbb{E}[\bar{\mathbf{w}}_i(\nu_l) \bar{\mathbf{w}}_i^H(\nu_l)] = \mathbf{I}_{O_R}$. Here,

a joint source-channel-coding pre-processing scheme is used. Thereby, each data packet contains encoded information about the vector $\mathbf{x}_g(\nu_l)$. This introduces decoding errors and distortion with a non-zero probability. The final decoded version is referred to by $\mathbf{y}_{g,n_l}(\nu_l)$ and is given by

$$\mathbf{y}_{g,n_l}(\nu_l) = \mathbf{x}_g(\nu_l) + \mathbf{w}_{g,n_l}(\nu_l), \quad (5)$$

in which $\mathbf{w}_{g,n_l}(\nu_l)$ represents zero-mean Gaussian noise with covariance matrix $\mathbf{Q}_{g,n_l}^W = q_{g,n_l}^W \mathbf{I}_{k_g}$. Based on $\mathbf{y}_{g,n_l}(\nu_l)$, an estimate $\hat{\mathbf{x}}_g(t)$ of the state at times $t \geq \nu_l$ is formed, which we later specify.

Short blocklength coding has been developed for single-input single-output (SISO) systems with a focus on channel coding in [2], and on joint source-channel coding in [3]. This has been later extended in [29] for MIMO systems. Following this framework, we aim at designing a system in such a way that the distortion of one transmission remains below a certain tolerable value d with a pre-specified probability $1 - \varepsilon_g$. Decoding errors are declared by an additional error detection scheme otherwise, which is built on top of the short blocklength scheme. In particular, for such condition to be satisfied, k_g real source symbols are encoded into a channel code of blocklength n_l time instants, which should satisfy [3]

$$n_l C - k_g R(d) \approx \sqrt{n_l V_C + k_g V_S} Q^{-1}(\varepsilon_g), \quad (6)$$

where C is the channel capacity and $R(d)$ is the rate-distortion function [30], V_C is the channel dispersion [2], and V_S is the source dispersion [31]. Computing the above for our Gaussian source setting, we have the capacity given by [30]

$$C = \log_2 \left| \mathbf{I}_{O_R} + \bar{\mathbf{H}}_g \bar{\mathbf{Q}} \bar{\mathbf{H}}_g^H \right|, \quad (7)$$

in which $\bar{\mathbf{Q}} = \mathbb{E}[\bar{\mathbf{x}}_i(\nu_l) \bar{\mathbf{x}}_i^H(\nu_l)]$ is the covariance matrix maximizing the capacity under a power constraint. Therefore the optimal $\bar{\mathbf{Q}}$ has a rank between 1 and $\min(O_R, O_T)$ and can be obtained using water-filling as described in [32]. Moreover, we have [29], [31]

$$V_C = \left(O_R - \sum_{i=0}^{O_R-1} \frac{1}{(1 + \text{eig}_i \{ \bar{\mathbf{H}}_g \bar{\mathbf{Q}} \bar{\mathbf{H}}_g^H \})^2} \right) \log_2^2(e), \quad (8)$$

$$V_S = \frac{1}{2} \log_2^2(e). \quad (9)$$

Using (6), an expression of the rate-distortion $R(d)$ can be obtained, which also describes the amount of information transmitted per system state dimension within each packet. This also indicates the number of bits needed to represent the source symbol with distortion not surpassing d . Such $R(d)$ is given by [30], [33]

$$R(d) = \frac{1}{k_g} \sum_{i=1}^{k_g} \max \left\{ \frac{1}{2} \log_2 \left(\frac{\text{eig}_i \{ \mathbf{Q}_{g,n_l}^Y \}}{d} \right), 0 \right\}, \quad (10)$$

where $\mathbf{Q}_{g,n_l}^Y = \mathbf{Q}_g^X + \mathbf{Q}_{g,n_l}^W$ denotes the covariance matrix of $\mathbf{y}_{g,n_l}(\nu_l)$, and \mathbf{Q}_g^X denotes the covariance matrix of $\mathbf{x}_g(t)$, which we will elaborate on later. Once n_l and ε_g are known,

we can use this equation to determine the noise level present within the data after decoding. For the worst case $q_{g,n_l}^W = d$, we can obtain the noise level by, e.g., the Newton algorithm:

$$d_{j+1} = d_j + \bar{\gamma} \frac{\sum_{i=1}^{k_g} \log_2 \left(1 + \frac{\text{eig}_i \{ \mathbf{Q}_g^X \}}{d_j} \right) - 2k_g R(d_j)}{\sum_{i=1}^{k_g} \frac{\log_2(e) \text{eig}_i \{ \mathbf{Q}_g^X \}}{d_j (d_j + \text{eig}_i \{ \mathbf{Q}_g^X \})}}. \quad (11)$$

Thereby, $\bar{\gamma}$ is the gradient descent rate. For $k_g = 1$ and $k_g = 2$, (10) can be formulated as linear and quadratic equation, respectively, and solved analytically. In practice, polar codes and turbo codes are often used to implement short blocklength coding [34].

III. PROBLEM FORMULATION

Based on the system described above, we are interested in optimizing the time-average AoI and MSE at the actuators. In the following subsections, we will elaborate on the time-average AoI and MSE of each sensor. Then, we will formulate our problem as a multi-objective optimization problem between AoI and MSE.

A. Age of Information (AoI)

For the system performance, it is important to have fresh data available at the actuators. The AoI metric is well-suited to capture this notion. If the latest measurement available at actuator g at time t has been generated at time ν_l , then the instantaneous AoI is at this time is

$$a_g(t) = t - \nu_l, \quad l = l_g(t). \quad (12)$$

Note that this variable does not depend on the data resolution, such that the optimization of the AoI is not always equivalent to the optimization of the measurement accuracy. The time-average AoI can be expressed as [28]

$$A_g = \limsup_{l \rightarrow \infty, l \in \mathcal{L}_g} \frac{\mathbb{E} \left[\int_0^{\bar{D}_l} a_g(t) dt \right]}{\mathbb{E} [\bar{D}_l]}. \quad (13)$$

Thereby, the limit superior ensures existence. A_g is bounded if the sensor g is scheduled regularly and the decoding is successful with probability $1 - \varepsilon_g > 0$.

For a packet l received at actuator g , we define the AoI right after its reception as \underline{a}_l , an the AoI right before receiving the next packet as \bar{a}_l . Hence,

$$\begin{aligned} \underline{a}_l &= \underline{D}_l - \nu_l, \\ \bar{a}_l &= \bar{D}_l - \nu_l, \end{aligned}$$

respectively. An illustration of these variables is shown in Fig. 3. Note that \bar{a}_l depends on the ratio at which packet losses occur and the required transmission times between the successful receptions, a parameter which depends on the scheduling policy used.

We now divide both the numerator and the denominator by the number of packets evaluated. Under a stationary decision policy, (13) can be reduced to an evaluation over each successful transmission [28], i.e., the numerator becomes the

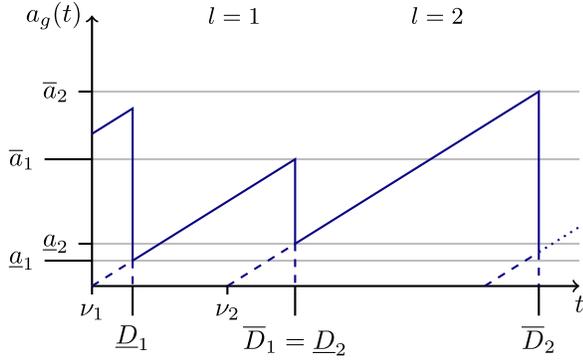


Fig. 3. Illustration of the AoI-related variables.

expectation of

$$\int_{\underline{a}_l}^{\bar{a}_l} a da = \frac{1}{2} (\bar{a}_l^2 - \underline{a}_l^2), \quad (14)$$

and the denominator becomes the expectation of $\bar{a}_l - \underline{a}_l$. Therefore, (13) can be expressed as

$$A_g = \frac{\mathbb{E}_{l \in \mathcal{L}_g} \left[\frac{1}{2} (\bar{a}_l^2 - \underline{a}_l^2) \right]}{\mathbb{E}_{l \in \mathcal{L}_g} [\bar{a}_l - \underline{a}_l]}. \quad (15)$$

When using this equation to analyze a decision policy, it has to be ensured that the numerator and denominator are implemented in a numerically stable way.

Next, we evaluate the time-average MSE.

B. Mean Square Error (MSE)

The MSE indicates the average squared difference between the system state and the estimate. The instantaneous MSE at actuator g is given by

$$m_g(t) = \mathbb{E} \left[\|\mathbf{x}_g(t) - \hat{\mathbf{x}}_g(t)\|_2^2 \right]. \quad (16)$$

Next, we use (2) to solve for $\mathbf{x}_g(t)$ in terms of the latest sampled value $\mathbf{x}_g(\nu_l)$, with $l = l_g(t)$. This is given by

$$\begin{aligned} \mathbf{x}_g(t) &= e^{\mathbf{Z}_g^a t} \mathbf{x}_g(\nu_l) \\ &+ \int_0^a e^{\mathbf{Z}_g^{(a-\mu)} t} (\mathbf{u}_g(\mu + \nu_l) + \mathbf{d}_g(\mu + \nu_l)) d\mu, \\ a &= a_g(t), \quad l = l_g(t), \quad t \geq \nu_l. \end{aligned} \quad (17)$$

We now consider the stationary covariance matrix \mathbf{Q}_g^X of the system state $\mathbf{x}_g(t)$. For diagonalizable \mathbf{Z}_g , this is presented in Lemma 1 below.

Lemma 1: For diagonalizable $\mathbf{Z}_g = \mathbf{U}_g \mathbf{\Lambda}_g \mathbf{U}_g^{-1}$, the covariance matrix of $\mathbf{x}_g(t)$ is given as

$$\mathbf{Q}_g^X = \mathbf{U}_g \tilde{\mathbf{Q}}_g^X \mathbf{U}_g^H, \quad (18)$$

in which

$$\left(\tilde{\mathbf{Q}}_g^X \right)_{i,j} = - \frac{\left(\mathbf{U}_g^{-1} \mathbf{Q}_g^D \mathbf{U}_g^{-H} \right)_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*} \quad (19)$$

and $\lambda_{g,i} = (\mathbf{\Lambda}_g)_{i,i}$.

Proof: For large t , we get a covariance matrix of the system state $\mathbf{x}_g(t)$ in (2) as

$$\begin{aligned} \mathbf{Q}_g^X &= \int_0^\infty e^{\mathbf{Z}_g \mu} \mathbf{Q}_g^D e^{\mathbf{Z}_g^H \mu} d\mu \\ &= \mathbf{U}_g \int_0^\infty e^{\mathbf{\Lambda}_g \mu} \mathbf{U}_g^{-1} \mathbf{Q}_g^D \mathbf{U}_g^{-H} e^{\mathbf{\Lambda}_g^H \mu} d\mu \mathbf{U}_g^H. \end{aligned} \quad (20)$$

When defining $\tilde{\mathbf{Q}}_g^D = \mathbf{U}_g^{-1} \mathbf{Q}_g^D \mathbf{U}_g^{-H}$, the element at location (i, j) of the integral is given as

$$\left(\tilde{\mathbf{Q}}_g^X \right)_{i,j} = \int_0^\infty e^{(\lambda_{g,i} + \lambda_{g,j}^*) \mu} \left(\tilde{\mathbf{Q}}_g^D \right)_{i,j} d\mu = - \frac{\left(\tilde{\mathbf{Q}}_g^D \right)_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*} \quad (21)$$

as long as $\lambda_{g,i} < 0$ for all i . Hence, the matrix \mathbf{Q}_g^X can be obtained as (18). \square

The input sequence $\mathbf{u}_g(t)$ equally influences the process state and the estimate.¹ Hence, by (16), the MSE can be described, for $a = a_g(t)$, $l = l_g(t)$, $t \geq \nu_l$, as

$$m_g(t) = \mathbb{E} \left[\left\| e^{\mathbf{Z}_g^a a} \mathbf{x}_g(\nu_l) + \int_0^a e^{\mathbf{Z}_g^{(a-\mu)} t} \mathbf{d}_g(\mu + \nu_l) d\mu - \mathbf{F}_{g,n_l}^{(a)} \mathbf{y}_{g,n_l}(\nu_l) \right\|_2^2 \right], \quad (22)$$

where $\mathbf{F}_{g,n_l}^{(a)}$ is the MSE-optimal estimator. We will consider two idealized cases of the MSE (22). In the first case, we consider *distortion-free* transmission, i.e., $\mathbf{w}_{g,n_l}(\nu_l) = \mathbf{0}$, in which case

$$\mathbf{F}_{g,n_l}^{(a)} \mathbf{y}_{g,n_l}(\nu_l) = e^{\mathbf{Z}_g^a a} \mathbf{x}_g(\nu_l), \quad (23)$$

and refer to the MSE as $m_g^i(t)$ in this case. Second, we consider the *input-disturbance-free* case, i.e., $\mathbf{d}_g(t) = \mathbf{0}$ for $t \geq \nu_l$, which we refer to as $m_g^c(t)$. Whereas the former depends on the input disturbance $\mathbf{d}_g(t)$, the latter depends on the channel noise $\mathbf{w}_{g,n_l}(\nu_l)$. As these two variables are uncorrelated, the instantaneous MSE can be separated as

$$m_g(t) = m_g^i(t) + m_g^c(t). \quad (24)$$

Similar to AoI case, the time-average MSE is formulated as

$$M_g = \limsup_{l \rightarrow \infty, l \in \mathcal{L}_g} \frac{\mathbb{E} \left[\int_0^{\bar{D}_l} m_g(t) dt \right]}{\mathbb{E} [\bar{D}_l]}. \quad (25)$$

This expression can be reduced to the case of minimizing the expected integrated MSE of each successful transmission, since the decision policy is stationary. Therefore, the numerator and denominator have to be divided by the number of packets evaluated. For $a = a_g(t)$, $l = l_g(t)$, $t \geq \nu_l$, the numerator becomes the expectation of

$$L_{g,n_l}(\underline{a}_l, \bar{a}_l) = \int_{\underline{a}_l}^{\bar{a}_l} m_g(\nu_l + a) da, \quad (26)$$

¹Note that if the receiver acts as actuator and controls the system state with an input signal $\mathbf{u}_g(t)$ this does not impact the MSE if the actuation changing the system state is also used to update the state estimate due to the linearity of the system.

with the denominator given by $\mathbb{E}[\bar{a}_l - \underline{a}_l]$. Therefore,

$$M_g = \frac{\mathbb{E}_{l \in \mathcal{L}_g} [L_{g,n_l}(\underline{a}_l, \bar{a}_l)]}{\mathbb{E}_{l \in \mathcal{L}_g} [\bar{a}_l - \underline{a}_l]}. \quad (27)$$

Similar to (24), we can separate $L_{g,n_l}(\underline{a}_l, \bar{a}_l) = L_g^i(\underline{a}_l, \bar{a}_l) + L_{g,n_l}^c(\underline{a}_l, \bar{a}_l)$ and $M_g = M_g^i + M_g^c$. Next, we will obtain closed form solutions of (26) for the two parts of the MSE.

1) *Distortion-Free Transmission*: We will now investigate the MSE defined in (16) for the case of no transmission errors, i.e., $\mathbf{y}_{g,n_l}(\nu_l) = \mathbf{x}_g(\nu_l)$. In this case, the optimal estimate is $\hat{\mathbf{x}}_g(t) = e^{\mathbf{Z}_g^a} \mathbf{y}_{g,n_l}(\nu_l)$, and therefore the MSE only describes the impacts of the later input disturbance. The instantaneous MSE is denoted for this case in lemma 2.

Lemma 2: Without transmission errors, the instantaneous MSE with AoI $a = a_g(t)$ is

$$m_g^i(t) = \text{trace} \left\{ \mathbf{Q}_g^X - e^{\mathbf{Z}_g^a} \mathbf{Q}_g^X e^{\mathbf{Z}_g^{H a}} \right\}. \quad (28)$$

Proof: When considering the distortion-free case in (23), (22) becomes

$$m_g^i(t) = \mathbb{E} \left[\left\| \int_0^a e^{\mathbf{Z}_g(a-\mu)} \mathbf{d}_g(\mu + \nu_l) d\mu \right\|_2^2 \right]. \quad (29)$$

As the input disturbance values $\mathbf{d}_g(t)$ and $\mathbf{d}_g(\mu)$ are uncorrelated for each $t \neq \mu$, this can be rephrased as

$$m_g^i(t) = \text{trace} \left\{ \int_0^a e^{\mathbf{Z}_g(a-\mu)} \mathbf{Q}_g^D e^{\mathbf{Z}_g^H(a-\mu)} d\mu \right\}. \quad (30)$$

From (30), we get

$$m_g^i(t) = \text{trace} \left\{ \mathbf{U}_g e^{\Lambda_g a(t)} \int_0^a e^{-\Lambda_g \mu} \tilde{\mathbf{Q}}_g^D e^{-\Lambda_g^* \mu} d\mu e^{\Lambda_g^* a} \mathbf{U}_g^{-1} \right\}, \quad (31)$$

in which $\tilde{\mathbf{Q}}_g^D = \mathbf{U}_g^{-1} \mathbf{Q}_g^D \mathbf{U}_g^{-H}$. Element i, j of the integral in the middle has a solution of

$$\begin{aligned} & \int_0^a e^{-(\lambda_{g,i} + \lambda_{g,j}^*) \mu} (\tilde{\mathbf{Q}}_g^D)_{i,j} d\mu \\ &= -\frac{(\tilde{\mathbf{Q}}_g^D)_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*} \left(e^{-(\lambda_{g,i} + \lambda_{g,j}^*) a} - 1 \right). \end{aligned} \quad (32)$$

Hence, the matrix representing this integral becomes

$$\int_0^a e^{-\Lambda_g \mu} \tilde{\mathbf{Q}}_g^D e^{-\Lambda_g^* \mu} d\mu = e^{-\Lambda_g a} \tilde{\mathbf{Q}}_g^X e^{-\Lambda_g^* a} - \tilde{\mathbf{Q}}_g^X. \quad (33)$$

When inserting this into (31), we get (28). \square

From $m_g^i(t)$, we can use (26) to determine the MSE integrated over the AoI interval $[\underline{a}_l, \bar{a}_l]$ of

$$\begin{aligned} L_g^i(\underline{a}_l, \bar{a}_l) &= \text{trace} \left\{ \int_{\underline{a}_l}^{\bar{a}_l} \mathbf{Q}_g^X - e^{\mathbf{Z}_g^a} \mathbf{Q}_g^X e^{\mathbf{Z}_g^H a} da \right\} \\ &= \text{trace} \left\{ (\bar{a}_l - \underline{a}_l) \mathbf{Q}_g^X + e^{\mathbf{Z}_g \bar{a}_l} \mathbf{S}_g^i e^{\mathbf{Z}_g^H \bar{a}_l} \right. \\ &\quad \left. - e^{\mathbf{Z}_g \underline{a}_l} \mathbf{S}_g^i e^{\mathbf{Z}_g^H \underline{a}_l} \right\}, \end{aligned} \quad (34)$$

in which

$$\mathbf{S}_g^i = \mathbf{U}_g \tilde{\mathbf{S}}_g^i \mathbf{U}_g^H \quad \text{and} \quad (\tilde{\mathbf{S}}_g^i)_{i,j} = -\frac{(\mathbf{U}_g^{-1} \mathbf{Q}_g^X \mathbf{U}_g^{-H})_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*}. \quad (35)$$

Thereby, the integrals are solved very similar to the ones in the proof of Lemma 2. Next, we consider the case of no input disturbance.

2) *No Input Disturbance*: After a successful decoding, the decoded system state contains noise as described in (5), which we ignored in the previous analysis. When employing this, the system state can be estimated using an MSE-optimal estimator $\mathbf{F}_{g,n_l}^{(a)}$ that depends on the blocklength n_l and sampling time ν_l of the most-recent available measurement in the form

$$\hat{\mathbf{x}}_g(\nu_l + a) = \mathbf{F}_{g,n_l}^{(a)} \mathbf{y}_{g,n_l}(\nu_l). \quad (36)$$

Lemma 3: When no input disturbance $\mathbf{u}_g(t)$ is present, the instantaneous MSE at receiver g for the estimator (36) can be obtained as

$$\begin{aligned} m_g^c(t) &= \text{trace} \left\{ \left(e^{\mathbf{Z}_g^a} - \mathbf{F}_{g,n_l}^{(a)} \right) \mathbf{Q}_g^X \left(e^{\mathbf{Z}_g^H a} - \left(\mathbf{F}_{g,n_l}^{(a)} \right)^H \right) \right\} \\ &\quad + \text{trace} \left\{ \mathbf{F}_{g,n_l}^{(a)} \mathbf{Q}_{g,n_l}^W \left(\mathbf{F}_{g,n_l}^{(a)} \right)^H \right\}, \end{aligned} \quad (37)$$

in which $a = a_g(t)$ is the current AoI.

Proof: We are assuming the case of having no system-input $\mathbf{u}_g(t)$ after a packet has been transmitted. The MSE in this case describes only the impacts of the channel noise and distortion

$$m_g^c(t) = \mathbb{E} \left[\left\| e^{\mathbf{Z}_g^a} \mathbf{x}_g(\nu_l) - \mathbf{F}_{g,n_l}^{(a)} (\mathbf{x}_g(\nu_l) + \mathbf{w}_{g,n_l}(\nu_l)) \right\|_2^2 \right]. \quad (38)$$

Using the independence of $\mathbf{w}_{g,n_l}(\nu_l)$ and $\mathbf{x}(\nu_l)$, as well as (17), (5), and (36), the instantaneous MSE becomes (37). \square

Taking derivative of (37) and setting it to 0, the optimal estimator can be obtained as

$$\mathbf{F}_{g,n_l}^{(a)} = e^{\mathbf{Z}_g^a} \mathbf{Q}_g^X \left(\mathbf{Q}_g^X + \mathbf{Q}_{g,n_l}^W \right)^{-1}. \quad (39)$$

Similar to (34), the MSE integrated over the AoI interval $[\underline{a}_l, \bar{a}_l]$ can be obtained by integrating (37). We get

$$\begin{aligned} L_{g,n_l}^c(\underline{a}_l, \bar{a}_l) &= \text{trace} \left\{ -e^{\mathbf{Z}_g \bar{a}_l} \mathbf{S}_{g,n_l}^{cX} e^{\mathbf{Z}_g^H \bar{a}_l} + e^{\mathbf{Z}_g \underline{a}_l} \mathbf{S}_{g,n_l}^{cX} e^{\mathbf{Z}_g^H \underline{a}_l} \right\} \\ &\quad + \text{trace} \left\{ -e^{\mathbf{Z}_g \bar{a}_l} \mathbf{S}_{g,n_l}^{cW} e^{\mathbf{Z}_g^H \bar{a}_l} + e^{\mathbf{Z}_g \underline{a}_l} \mathbf{S}_{g,n_l}^{cW} e^{\mathbf{Z}_g^H \underline{a}_l} \right\}, \end{aligned} \quad (40)$$

where

$$\mathbf{S}_{g,n_l}^{cX} = \mathbf{U}_g \tilde{\mathbf{S}}_{g,n_l}^{cX} \mathbf{U}_g^H \quad (41)$$

$$(\tilde{\mathbf{S}}_{g,n_l}^{cX})_{i,j} = -\frac{\left(\mathbf{U}_g^{-1} \bar{\mathbf{F}}_{g,n_l}^{(0)} \mathbf{Q}_g^X \left(\bar{\mathbf{F}}_{g,n_l}^{(0)} \right)^H \mathbf{U}_g^{-H} \right)_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*},$$

$$\mathbf{S}_{g,n_l}^{cW} = \mathbf{U}_g \tilde{\mathbf{S}}_{g,n_l}^{cW} \mathbf{U}_g^H \quad (42)$$

$$(\tilde{\mathbf{S}}_{g,n_l}^{cW})_{i,j} = -\frac{\left(\mathbf{U}_g^{-1} \mathbf{F}_{g,n_l}^{(0)} \mathbf{Q}_{g,n_l}^W \left(\mathbf{F}_{g,n_l}^{(0)} \right)^H \mathbf{U}_g^{-H} \right)_{i,j}}{\lambda_{g,i} + \lambda_{g,j}^*},$$

and $\bar{\mathbf{F}}_{g,n_l}^{(0)} = \mathbf{I} - \mathbf{F}_{g,n_l}^{(0)}$. When defining $\mathbf{S}_{g,n_l}^c = \mathbf{S}_{g,n_l}^{cX} + \mathbf{S}_{g,n_l}^{cW}$, we can write (40) as a function of the joint covariance matrix instead of the individual matrices.

3) *Channel Noise and Input Disturbance*: When combining both parts of the packet-integrated MSE, i.e., (34) and (40), we have

$$L_{g,n_l}(\underline{a}_l, \bar{a}_l) = \text{trace} \left\{ (\bar{a}_l - \underline{a}_l) \mathbf{Q}_g^X + e^{\mathbf{Z}_g \bar{a}_l} \mathbf{S}_{g,n_l} e^{\mathbf{Z}_g^H \bar{a}_l} - e^{\mathbf{Z}_g \underline{a}_l} \mathbf{S}_{g,n_l} e^{\mathbf{Z}_g^H \underline{a}_l} \right\}, \quad (43)$$

in which $\mathbf{S}_{g,n_l} = \mathbf{S}_g^i - \mathbf{S}_{g,n_l}^c$. This expression can be inserted in (27) to obtain the time-average MSE. Going back to the results of Lemma 2 and Lemma 3, we note that the instantaneous MSE mainly depends on the product of system matrix and instantaneous AoI.

Remark 1: Both parts of the instantaneous MSE $m_g(t)$ depend on $e^{\mathbf{Z}_g a}$. This means that the MSE is small when either the AoI a or the absolute values of the eigenvalues of \mathbf{Z}_g are small. Hence, it can be beneficial for the maximum MSE to schedule the sensors with larger-amplitude \mathbf{Z}_g more often.

With the expressions of AoI and MSE stated, we are now able to consider the multi-objective optimization of both.

C. Multi-Objective Optimization

To formulate a joint optimization problem covering all sensor-actuator pairs, we are normalizing the AoI by the symbol duration α . The MSE instead is normalized by the asymptotic MSE, which is reached if there are no successful transmissions over a long time, such that it is bounded to one,

$$\hat{A}_g = \frac{A_g}{\alpha}, \quad (44)$$

$$\hat{M}_g = \frac{M_g}{\text{trace} \left\{ \mathbf{Q}_g^X \right\}}. \quad (45)$$

This normalization also ensures comparability in the case of system states with very different amplitudes. The normalized AoI and MSE are now optimized jointly as

$$\underset{\substack{\pi=(g_1, g_2, \dots) \\ f=(\nu_1, \nu_2, \dots) \\ c=(n_1, n_2, \dots)}}{\text{minimize}} \left(\hat{A}_1, \dots, \hat{A}_G, \hat{M}_1, \dots, \hat{M}_G \right). \quad (46)$$

The optimal decision policy can be found by traversing through the set of possible scheduling, sampling, and coding strategies. For each set of strategies, the AoI and MSE can be analyzed algorithmically by simulating the transmission of a high number of packets. Therefore, the different packet-integrated MSE-values in (43) can be used to obtain the AoI and MSE in (15) and (27).

However, the computational resources required for finding an optimal decision policy are rather expensive. Hence, we instead focus in the following on two widely-used scheduling policies, i.e., round-robin and maximum-age. The goal is to optimize the coding strategy $c = (n_1, n_2, \dots)$ such that the combination of AoI and MSE is minimized, while the scheduling strategy $\pi = (g_1, g_2, \dots)$ and the sampling strategy $f = (\nu_1, \nu_2, \dots)$ are fixed. To characterize the Pareto-boundary, the AoI and MSE have to be optimized jointly as

$$\underset{c=(n_1, n_2, \dots)}{\text{minimize}} \left(\hat{A}_1, \dots, \hat{A}_G, \hat{M}_1, \dots, \hat{M}_G \right). \quad (47)$$

In order to obtain an achievability region of such a multi-objective problem [35], we will traverse through the set of different coding strategies. Next, we will be investigating the impact of the scheduling policies the AoI and MSE. Later on, we will elaborate on how to choose the sampling and scheduling strategies.

IV. AOI AND MSE FOR SPECIAL POLICIES

We now focus on special policies in order to compute their Pareto-optimal boundary for problem (47). First, we consider the coding strategy in which each sensor g uses a constant blocklength n_g to communicate. In this case, the transmission time of all packets transmitted to actuator g is b_g . Hence, one can show that the optimal sampling strategy f is a *zero-wait policy*, in which a new sample is generated and transmitted right after the g th sensor's turn comes up according to the central scheduler [36, Theorem 2].

We now focus on the scheduling strategy π . We analyze two classes of scheduling strategies, namely *round-robin* and *maximum-age*, which we detail next.

A. Round-Robin Scheduling

In round-robin scheduling, the sensors are scheduled in the same order $\{1, 2, \dots, G\}$ regardless of failures. That is, once the transmission of sensor g is finished, sensor $g+1$ is scheduled regardless of whether sensor g 's transmission has been successfully-received at actuator g . This means that the time required for one round of transmissions is the sum of the transmission times of all sensor-actuator pairs. Now, let us define the length of one round as

$$\hat{b} = \sum_{g'=0}^G b_{g'}. \quad (48)$$

Therefore, the length of the AoI interval $[\underline{a}_l, \bar{a}_l)$ for sensor g contains

$$\underline{a}_l = b_g, \quad \bar{a}_l = b_g + \hat{b}_g, \quad \forall l \in \mathcal{L}_g, \quad (49)$$

where the term \hat{b}_g depends on the number of transmissions of sensor g in the interval $[\underline{a}_l, \bar{a}_l)$, and so is given by

$$\hat{b}_g \sim \sum_{v=0}^{\infty} \varepsilon_g^v (1 - \varepsilon_g) \delta(\hat{b}_g - (v+1)\hat{b}), \quad (50)$$

where v denotes the number of failed transmissions. An example is shown in Fig. 4. We now have the following lemma regarding the AoI:

Lemma 4: In case of a round-robin policy, the time-average AoI at sensor g is

$$A_g = b_g + \frac{1 + \varepsilon_g \hat{b}}{2(1 - \varepsilon_g)}. \quad (51)$$

Proof: When assuming $\underline{a}_l = b_g$ and $\bar{a}_l = b_g + \hat{b}'_{g,g}$, (15) becomes

$$A_g = b_g + \frac{\mathbb{E}_{\hat{b}_g} \left[\frac{1}{2} (\hat{b}_g)^2 \right]}{\mathbb{E}_{\hat{b}_g} [\hat{b}_g]}. \quad (52)$$

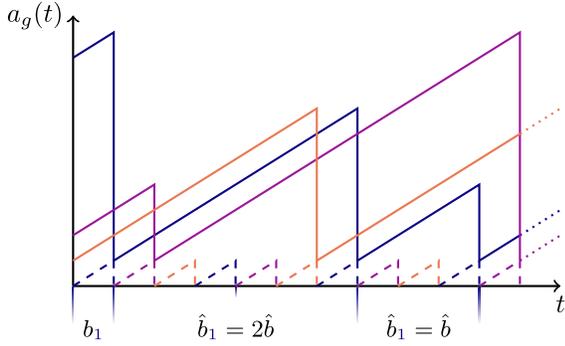


Fig. 4. Example of how the AoI changes over time for three sensor-actuator pairs with the round-robin policy.

The first and second moments of \hat{b}_g are now expressed as follows:

$$\mathbb{E}_{\hat{b}_g} [\hat{b}_g] = \sum_{v=0}^{\infty} \varepsilon_{g'}^v (1 - \varepsilon_{g'}) (v + 1) \hat{b} = \frac{1}{1 - \varepsilon_{g'}} \hat{b}. \quad (53)$$

$$\begin{aligned} \mathbb{E}_{\hat{b}_g} \left[\left(\hat{b}_g \right)^2 \right] &= \sum_{v=0}^{\infty} \varepsilon_{g'}^v (1 - \varepsilon_{g'}) \left((v + 1) \hat{b} \right)^2 \\ &= (1 - \varepsilon_{g'}) \hat{b}^2 \frac{1 + \varepsilon_{g'}}{(1 - \varepsilon_{g'})^3} = \frac{1 + \varepsilon_{g'}}{(1 - \varepsilon_{g'})^2} \hat{b}^2. \end{aligned} \quad (54)$$

When substituting (53) and (54) into (52), we get (51). \square

Next, we have the following lemma regarding the MSE:

Lemma 5: In case of a round-robin policy, the time-average MSE at sensor g is

$$M_g = \frac{\mathbb{E}_{\hat{b}_g} [L_{g,n_g}(\underline{a}_l, \bar{a}_l)]}{\frac{1}{1 - \varepsilon_g} \hat{b}}, \quad (55)$$

in which the numerator is

$$\begin{aligned} \mathbb{E}_{\hat{b}_g} [L_g(\underline{a}_l, \bar{a}_l)] &= \text{trace} \left\{ \left(\frac{1}{1 - \varepsilon_g} \hat{b} \right) \mathbf{Q}_g^X + e^{\mathbf{Z}_g^H (b_g + \hat{b})} \Psi(g, \hat{b}) \right. \\ &\quad \left. \times e^{\mathbf{Z}_g (b_g + \hat{b})} \mathbf{S}_{g,n_g} - e^{\mathbf{Z}_g^H b_g} e^{\mathbf{Z}_g b_g} \mathbf{S}_{g,n_g} \right\}. \end{aligned} \quad (56)$$

Thereby, we have $\hat{b} = \sum_{g=0}^G b_g$ and

$$\Psi(g, \hat{b}) = \mathbf{U}_g^{-H} \tilde{\Psi}(g, \hat{b}) \mathbf{U}_g^{-1}, \quad (57)$$

in which

$$\left(\tilde{\Psi}(g, \hat{b}) \right)_{i,j} = \frac{1 - \varepsilon_g}{1 - \varepsilon_g e^{(\lambda_{g,i}^* + \lambda_{g,j}) \hat{b}}} \left(\mathbf{U}_g^H \mathbf{U}_g \right)_{i,j}. \quad (58)$$

Proof: After matrix rotation and inserting (50), the expectation of parts of (43) can be written as

$$\Psi(g, \hat{b}) = \sum_{v=0}^{\infty} \varepsilon_g^v (1 - \varepsilon_g) e^{\mathbf{Z}_g^H \hat{b} v} e^{\mathbf{Z}_g \hat{b} v}, \quad (59)$$

in which \hat{b} is the time elapsed between two transmission trials. When diagonalizing $\mathbf{Z}_g = \mathbf{U}_g \mathbf{\Lambda}_g \mathbf{U}_g^{-1}$, we can obtain the

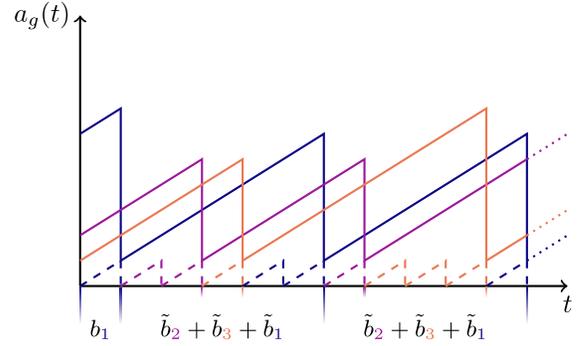


Fig. 5. A possible example of how the AoI develops over time for three sensor-actuator pairs and the maximum-age policy.

elements of $\tilde{\Psi}(g, \hat{b})$ as

$$\left(\tilde{\Psi}(g, \hat{b}) \right)_{i,j} = \sum_{v=0}^{\infty} \varepsilon_g^v (1 - \varepsilon_g) e^{\lambda_{g,i}^* \hat{b} v} \left(\mathbf{U}_g^H \mathbf{U}_g \right)_{i,j} e^{\lambda_{g,j} \hat{b} v}. \quad (60)$$

By applying the geometric series, we get (58). When defining (57), the expectation of (43) becomes (56). From (27), we get a closed-form MSE expression of (55). \square

B. Maximum-Age Scheduling

In maximum-age scheduling, the sensor with the maximum AoI gets scheduled. This way, transmitting data from a scheduled sensor continuous until successful reception, since the AoI order does not change with failures. Different from round-robin scheduling, the length of the time interval $[\underline{a}_l, \bar{a}_l]$ for sensor g is now given by

$$\underline{a}_l = b_g, \quad \bar{a}_l = b_g + \sum_{g'=1}^G \tilde{b}_{g'}, \quad \forall l \in \mathcal{L}_g, \quad (61)$$

where $\tilde{b}_{g'}$ denotes the time needed for each sensor to finish its *successful* transmission, and is distributed as

$$\tilde{b}_{g'} \sim \sum_{v=0}^{\infty} \varepsilon_{g'}^v (1 - \varepsilon_{g'}) \delta \left(\tilde{b}_{g'} - (v + 1) b_{g'} \right). \quad (62)$$

An example is shown in Fig. 5. We now have the following lemma regarding the AoI:

Lemma 6: In case of a maximum-age policy, the time-average AoI at sensor g is

$$A_g = b_g + \frac{1}{2} \sum_{g'=1}^G \frac{1}{1 - \varepsilon_{g'}} b_{g'} + \frac{1}{2} \frac{\sum_{g'=1}^G \frac{\varepsilon_{g'}}{(1 - \varepsilon_{g'})^2} b_{g'}^2}{\sum_{g'=1}^G \frac{1}{1 - \varepsilon_{g'}} b_{g'}}. \quad (63)$$

Proof: From (15), we get

$$A_g = b_g + \frac{\mathbb{E}_{\tilde{b}_{g'}} \left[\frac{1}{2} \left(\sum_{g'=1}^G \tilde{b}_{g'} \right)^2 \right]}{\mathbb{E}_{\tilde{b}_{g'}} \left[\sum_{g'=1}^G \tilde{b}_{g'} \right]}.$$

When utilizing the first-order and second-order moments

$$\begin{aligned}\mathbb{E}_{\tilde{b}_{g'}}[\tilde{b}_{g'}] &= \sum_{v=0}^{\infty} \varepsilon_{g'}^v (1 - \varepsilon_{g'}) (v + 1) b_{g'} \\ &= \frac{1}{1 - \varepsilon_{g'}} b_{g'}.\end{aligned}\quad (64)$$

$$\begin{aligned}\mathbb{E}_{\tilde{b}_{g'}}\left[\left(\tilde{b}_{g'}\right)^2\right] &= \sum_{v=0}^{\infty} \varepsilon_{g'}^v (1 - \varepsilon_{g'}) ((v + 1) b_{g'})^2 \\ &= (1 - \varepsilon_{g'}) b_{g'}^2 \frac{1 + \varepsilon_{g'}}{(1 - \varepsilon_{g'})^3} = \frac{1 + \varepsilon_{g'}}{(1 - \varepsilon_{g'})^2} b_{g'}^2,\end{aligned}\quad (65)$$

we obtain (63). \square

Next, we have the following lemma regarding the MSE:

Lemma 7: In case of a maximum-age policy, the time-average MSE at sensor g is

$$M_g = \frac{\mathbb{E}_{l \in \mathcal{L}_g} [L_{g,n_g}(\underline{a}_l, \bar{a}_l)]}{\sum_{g'=1}^G \frac{1}{1 - \varepsilon_{g'}} b_{g'}}.\quad (66)$$

Here, the expectation in the numerator is

$$\begin{aligned}\mathbb{E}_{l \in \mathcal{L}_g} [L_{g,n_g}(\underline{a}_l, \bar{a}_l)] &= \text{trace} \left\{ \left(\sum_{g'=1}^G \frac{1}{1 - \varepsilon_{g'}} b_{g'} \right) \mathbf{Q}_g^X + e^{\mathbf{Z}_g^H (b_g + \hat{b})} \mathbf{\Psi}_g e^{\mathbf{Z}_g (b_g + \hat{b})} \right. \\ &\quad \left. \times \mathbf{S}_{g,n_g} - e^{\mathbf{Z}_g^H b_g} e^{\mathbf{Z}_g b_g} \mathbf{S}_{g,n_g} \right\},\end{aligned}\quad (67)$$

in which $\mathbf{\Psi}_g = \mathbf{U}_g^{-H} \tilde{\mathbf{\Psi}}_g \mathbf{U}_g^{-1}$ and

$$\left(\tilde{\mathbf{\Psi}}_g\right)_{i,j} = \left(\prod_{g'=1}^G \frac{1 - \varepsilon_{g'}}{1 - \varepsilon_{g'} e^{(\lambda_{g',i}^* + \lambda_{g',j}) b_{g'}}} \right) \left(\mathbf{U}_g^H \mathbf{U}_g \right)_{i,j}.\quad (68)$$

Proof: Similar to the definition of $\mathbf{\Psi}(g, p)$ in the round-robin policy, we can define here

$$\begin{aligned}\mathbf{\Psi}_g &= \sum_{(v_1, \dots, v_G) = (0, \dots, 0)}^{(\infty, \dots, \infty)} \left(\prod_{g'=1}^G \varepsilon_{g'}^{v_{g'}} (1 - \varepsilon_{g'}) \right) \\ &\quad \times \exp \left(\mathbf{Z}_g^H \sum_{g'=1}^G b_{g'} v_{g'} \right) \exp \left(\mathbf{Z}_g \sum_{g'=1}^G b_{g'} v_{g'} \right).\end{aligned}\quad (69)$$

After diagonalizing the matrix \mathbf{Z}_g , we can apply the geometric series on each element of $\tilde{\mathbf{\Psi}}_g$ once for each sensor-actuator pair to obtain (68). With this definition, the expectation of (43) becomes (67). From (27), we get a closed-form MSE expression of (66). \square

V. GENERALIZATION OF THE RESULTS

A. Weighted MSE

The calculations of the MSE are well-suited, if the different entries of the system state should be weighted equally. But, the entries of each vector $\mathbf{x}_g(t)$ might cover very different operational meanings and also different units. Therefore, it might be beneficial to introduce weight matrices \mathbf{R}_g of full rank,

which are applied on each system state, i.e., considering the transformed system states $\mathbf{x}'_g(t) = \mathbf{R}_g \mathbf{x}_g(t)$. To use these variables in (1), also the transformations $\mathbf{Z}'_g = \mathbf{R}_g \mathbf{Z}_g \mathbf{R}_g^{-1}$, $\mathbf{u}'_g(t) = \mathbf{R}_g \mathbf{u}_g(t)$ and $\mathbf{Q}_g^{\text{D}'} = \mathbf{R}_g \mathbf{Q}_g^{\text{D}} \mathbf{R}_g^H$ have to be applied. With these transformations, the methods presented above are also suitable for this case.

B. Multi-Subcarrier

Above, we have formulated the system model for a single subcarrier. If a system enables support for multiple subcarriers, the central scheduler might also decide about the subcarrier used, such that multiple transmissions are active at the same time. This means that the decision policy has to be extended by a subcarrier policy. Moreover, the channel matrices $\tilde{\mathbf{H}}_g$ also have to depend on the subcarrier selected for each transmission. The generalized results in (46) also hold in this case, whereas the analytical results obtained for the round-robin and maximum-age policies have to be adapted. Therefore, one option is to first allocate the sensor-actuator pairs to the different subcarriers, before then applying round-robin or maximum-age scheduling to each subcarrier.

We now summarize our approach to characterize Pareto optimal boundaries for problem (47), which also describes an achievability region between AoI and MSE. For either round-robin or maximum-age scheduling, combined with zero-wait sampling, we have expressions for the time-average AoI and MSE. These expressions, however, depend on the coding strategy being used. Each point on the Pareto boundary, therefore, will correspond to a specific choice of blocklengths. We provide more detailed examples on this in the next section.

VI. NUMERICAL RESULTS

In this section, we present detailed numerical examples to further illustrate the results of this paper. We focus on the case of two sensor-actuator pairs, i.e., $G = 2$.

We will contrast the round-robin and maximum-age scheduling policies with another scheduling policy that we term *asynchronous*-(T_1, T_2). In such policy, sensor g gets T_g consecutive trials to transmit. When either the number of trials is reached or one transmission is successful, the other sensor g' gets to transmit for a maximum of $T_{g'}$ number of consecutive trials, and so on. This way, round-robin scheduling equivalent to asynchronous-(1, 1) scheduling, while maximum-age scheduling is equivalent to asynchronous-(∞, ∞) scheduling. This way, asynchronous-(T_1, T_2) generalizes round-robin and maximum-age scheduling, and will be shown useful in instance where the processes monitored have variant parameters.

We now describe the system parameters. We first assume that the first process system is characterized by

$$\begin{aligned}\mathbf{Z}_1 &= \begin{pmatrix} -0.08 & 0.03 & -0.05 \\ -0.01 & -0.07 & 0.05 \\ 0.2 & 0.35 & -0.55 \end{pmatrix}, \\ \mathbf{Q}_1^{\text{D}} &= \begin{pmatrix} 4 & 1 & 3 \\ 1 & 0.25 & 0.75 \\ 3 & 0.75 & 2.25 \end{pmatrix},\end{aligned}$$

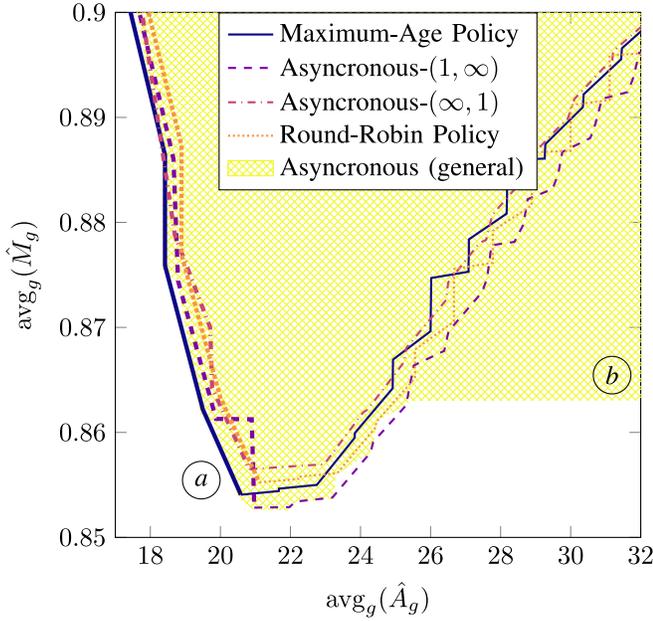


Fig. 6. Achievability region of time-average AoI and MSE for a system with smaller eigenvalues in dependency of the coding strategy for different scheduling policies. The Pareto-optimal parts (a) on the left are drawn slightly thicker. (b) shows the case of giving full priority to one sensor.

whereas the second process system is characterized by

$$\mathbf{Z}_2 = \begin{pmatrix} -0.04 & 0 \\ 0 & -0.03 \end{pmatrix}, \quad \mathbf{Q}_2^D = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.6 \end{pmatrix}.$$

The channels between the sensors and actuators are specified by

$$\bar{\mathbf{H}}_1 = \begin{pmatrix} 0.9 & 0 \\ 0 & 0.95 \end{pmatrix}, \quad \bar{\mathbf{H}}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0.8 \end{pmatrix},$$

with $\alpha = 1$ and $\beta = 3$. We assume that the coding is done in a way that decoding is unsuccessful at a fraction of $\varepsilon_1 = \varepsilon_2 = 0.1$.

In Fig. 6, we plot the *average* (over the number of sensors) MSE versus the *average* AoI using the different scheduling policies mentioned above. Differently, in Fig. 7, we plot the *maximum* MSE versus the *maximum* AoI. In both figures, each point on the curves is achieved by a certain blocklength pair (n_1, n_2) . All points are then connected to form the whole boundary/region. The points on the left hand side of the boundaries represent Pareto-optimal points, indicated by the letter *a* in the figure. In general, the boundaries show that there exists an intrinsic relationship between AoI and MSE. Basically, precise information requires more transmission time whereas fast transmission incurs higher distortion. One can make either the AoI or the MSE sufficiently small, but not both simultaneously. The boundaries also show that neither small or large values of the AoI are MSE-optimal. Instead, there exists an optimal coding strategy beyond which the MSE is not enhanced since larger delays are incurred, and before which the MSE is also not enhanced since larger distortions are incurred. Finally, we note that the boundaries represent *achievability regions* between MSE and AoI, in which each

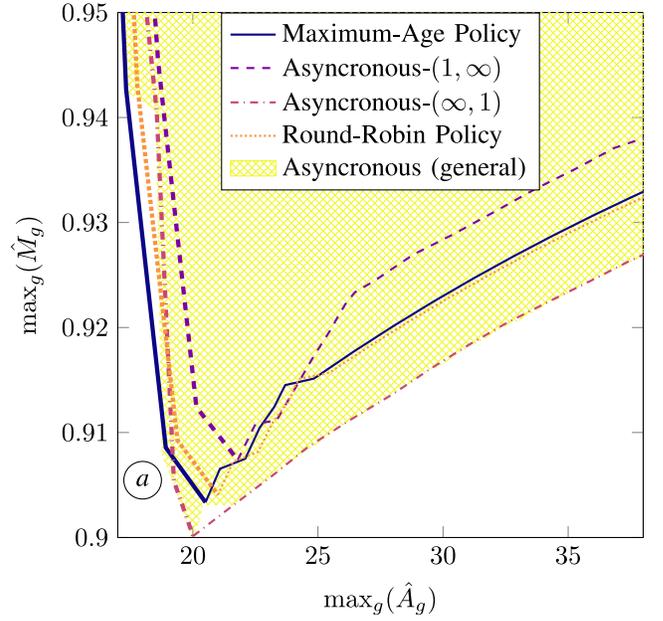


Fig. 7. Achievability region of maximum AoI and MSE for a system with smaller eigenvalues in dependency of the coding strategy for different scheduling policies. On the left (a), the Pareto-optimal parts are shown slightly thicker.

point *above* the boundary lines is achievable using a specific coding strategy.

In the figures, we also plot the performance achieved by a general asynchronous- (T_1, T_2) policy by traversing through all possible (T_1, T_2) pairs; this gives the colored background region in the figures. If one sets $T_1 = 0$, this means we never schedule sensor 1. In this case, the average and maximum AoI grow infinitely large. The MSE of sensor 1 is bounded by its process variance (normalized to 1), and so the average MSE will always stay below 1 (as shown in the right hand side of Fig. 6 indicated by the letter *b*), yet the maximum MSE grows to 1 (as shown in the right hand side Fig. 7).

In both figures, the maximum-age policy is providing the lowest (average and maximum) AoI values. However, an asynchronous transmission policy that schedules the sensors in accordance to the eigenvalues of \mathbf{Z}_g differently can be beneficial from an MSE perspective. In our scenario, the lowest eigenvalue of \mathbf{Z}_1 is given by -0.5628 , which is much smaller than the lowest eigenvalue of \mathbf{Z}_2 , given by -0.04 . This means that the state of process system 1 is changing much more dynamically than the state of process system 2, leading to an asynchronous scheduling scheme performing best.

To additionally evaluate the behavior in case of other process systems, we now change the system matrices to

$$\mathbf{Z}_1 = \begin{pmatrix} -0.04 & 0.03 & -0.05 \\ -0.01 & -0.06 & 0.05 \\ 0.2 & 0.15 & -0.4 \end{pmatrix},$$

$$\mathbf{Z}_2 = \begin{pmatrix} -0.02 & 0 \\ 0 & -0.03 \end{pmatrix}.$$

This means that the lowest eigenvalues of \mathbf{Z}_1 and \mathbf{Z}_2 become -0.3910 and -0.03 , respectively, which is larger than in the previous example. The according average MSE and average

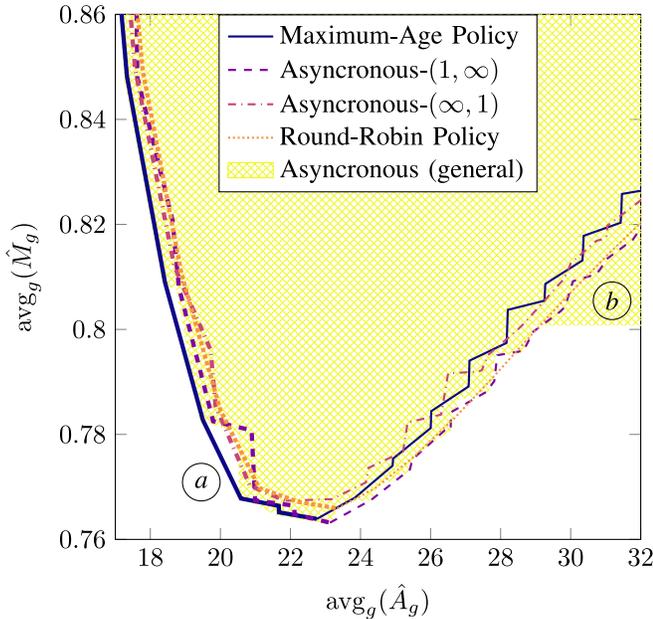


Fig. 8. Achievability region of time-average AoI and MSE for a system with larger eigenvalues in dependency of the coding strategy for different scheduling policies. The Pareto-optimal parts (a) on the left are drawn slightly thicker. (b) shows the case of giving full priority to one sensor.

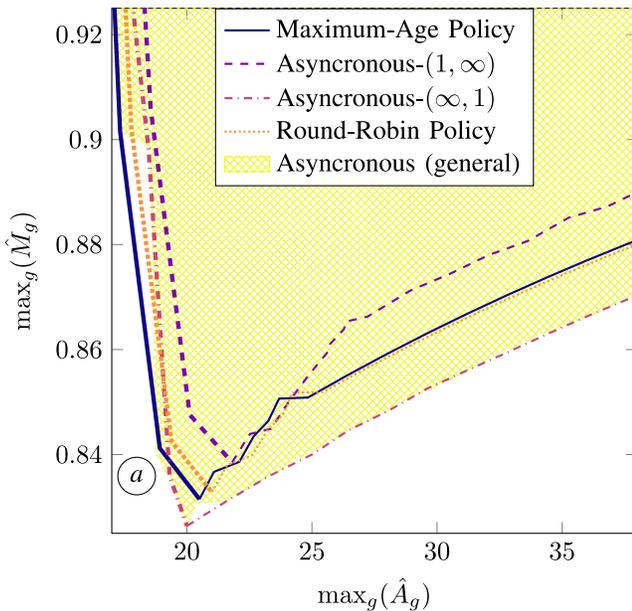


Fig. 9. Achievability region of maximum AoI and MSE for a system with larger eigenvalues in dependency of the coding strategy for different scheduling policies. On the left (a), the Pareto-optimal parts are shown slightly thicker.

AoI are shown in Fig. 8, whereas the maximum MSE and maximum AoI are shown in Fig. 9. These plots indicate that there is a very similar general behavior in both pairs of process systems. However, the amplitudes of the MSE values are very different. Also, the ratio between the MSE for the case of giving full priority to one sensor (indicated with the letter b) and the MSE of the other four policies differs.

To further evaluate which sensor-actuator pair should be given priority in an asynchronous scheduling policy, the

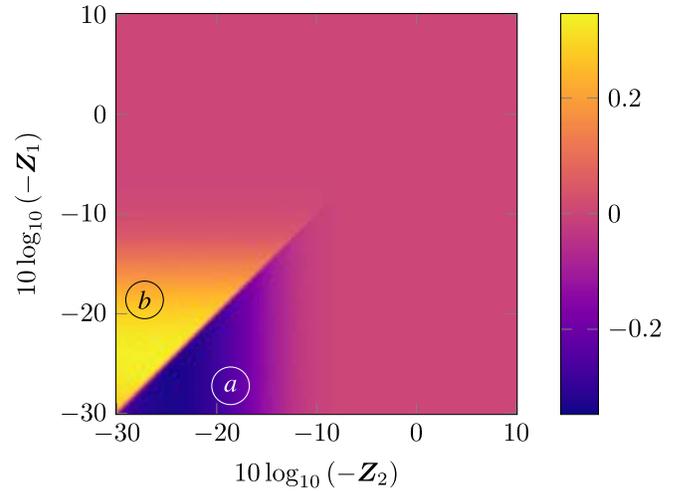


Fig. 10. Logarithmic fraction of maximum MSE achieved with the two asynchronous policies with parameters (1, ∞) and (∞ , 1) in (70) over the value of the process system matrices \mathbf{Z}_i . Blue (a) and yellow (b) indicate that sensors 1 and 2 should be given priority, respectively.

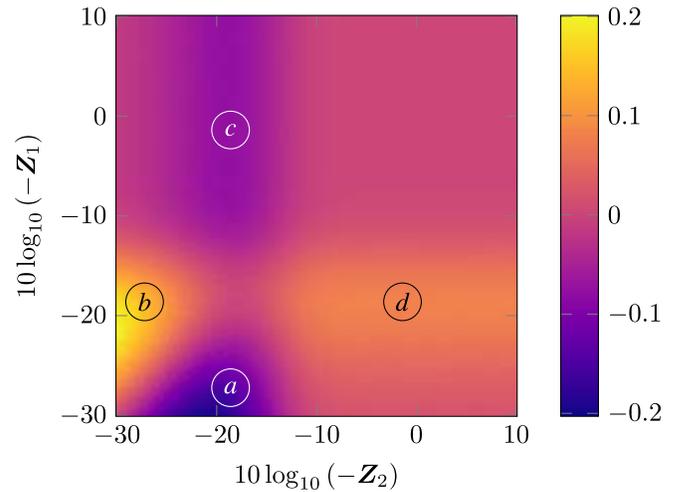


Fig. 11. Logarithmic fraction of average MSE achieved with the two asynchronous policies with parameters (1, ∞) and (∞ , 1) in (71) over the value of the process system matrices \mathbf{Z}_i . Areas shaded in blue (a, c) and yellow (b, d) indicate that sensor 1 and 2 should be given priority, respectively.

dependency of the eigenvalues on the MSE is investigated. Towards that end, we consider a simplified setting with two process systems, where \mathbf{Z}_g 's are 1×1 matrices and $\mathbf{Q}_g^D = 1$. The communication parameters are specified by $\bar{\mathbf{H}}_g = 1$, $n_g = 7$, $\alpha = 1$ and $\beta = 3$. Similar to above, the decoding fails at a fraction of $\varepsilon_1 = \varepsilon_2 = 0.1$. We define the maximum and average MSE with the asynchronous policies as $\hat{M}_{\max}^{(T_1, T_2)}$ and $\hat{M}_{\text{avg}}^{(T_1, T_2)}$, respectively. In Fig. 10 and Fig. 11, the logarithmic fraction of the maximum and average MSE of the two policies, i.e.,

$$10 \log_{10} \left(\frac{\hat{M}_{\max}^{(1, \infty)}}{\hat{M}_{\max}^{(\infty, 1)}} \right), \quad (70)$$

$$10 \log_{10} \left(\frac{\hat{M}_{\text{avg}}^{(1, \infty)}}{\hat{M}_{\text{avg}}^{(\infty, 1)}} \right), \quad (71)$$

are plotted versus the values of Z_1 and Z_2 . This means that in regions where this value has a low-amplitude (indicated with the letter *a*), sensor 1 should be scheduled more often, whereas sensor 2 should be prioritized in areas with a high-amplitude (indicated with the letter *b*). These figures show that it is beneficial to schedule the sensor with the eigenvalue of the higher amplitude more often (especially for eigenvalues near zero). On and near the diagonal, both systems are almost identical. Hence, both asynchronous policies behave equally well and the maximum-age policy is therefore optimal. In the case of optimizing the average MSE, there are also exceptions to prioritizing the higher-amplitude eigenvalues more often (as indicated with the letters *c* and *d*). Transmission of the state-value of the less-dynamical system can be more beneficial to the average MSE under certain conditions. However, considering the maximum MSE promotes fairness, and hence priority is given to the more dynamical system.

VII. CONCLUSION

The quality of the monitoring of process systems has a significant impact on the control performance. Two metrics widely used for measuring the monitoring quality are the AoI and the MSE. In this paper, we have investigated the trade-off between the maximum and average of these two metrics. Whereas the AoI mainly depends on the system delays and the scheduling policy, the MSE also depends on the transmission noise. As the impact of the transmission noise is especially relevant when the delays within the system are small, both objectives show a similar behavior when the blocklengths are large. However, when the communication parameters are chosen to have low blocklengths, this leads to small average AoI values together with high MSE values, i.e., these two metrics show the opposite behavior. A minimal AoI could be reached by choosing the shortest possible blocklength, whereas an intermediate blocklength is MSE-optimal.

We have derived closed-form expressions for the AoI and MSE for two widely used scheduling policies, i.e., the round-robin policy and the maximum-age policy. The maximum-age policy provides excellent results in terms of AoI when the probability of a successful transmission of the sensors is equal. However, numerical results suggest that asynchronous priority-based scheduling policies that schedule sensors whose processes are more dynamic more often can provide a lower maximum MSE.

REFERENCES

- [1] S. Roth, A. Arafa, H. V. Poor, and A. Sezgin, "Remote short blocklength process monitoring: Trade-off between resolution and data freshness," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [2] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [3] V. Kostina and S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2545–2575, May 2013.
- [4] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2731–2735.
- [5] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the Wiener process over a channel with random delay," in *Proc. IEEE Int. Symp. Inform. Theory*, Jun. 2017, pp. 321–325.
- [6] T. Z. Ornee and Y. Sun, "Sampling for remote estimation through queues: Age of information and beyond," in *Proc. Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOPT)*, Jun. 2019, pp. 1–8.
- [7] E. Najm, R. Yates, and E. Soljanin, "Status updates through M/G/1/1 queues with HARQ," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 131–135.
- [8] H. Sac, T. Bacinoglu, E. Uysal-Biyikoglu, and G. Durisi, "Age-optimal channel coding blocklength for an M/G/1 queue with HARQ," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2018, pp. 1–5.
- [9] R. Devassy, G. Durisi, G. C. Ferrante, O. Simeone, and E. Uysal-Biyikoglu, "Delay and peak-age violation probability in short-packet transmissions," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 2471–2475.
- [10] E. Najm, E. Telatar, and R. Nasser, "Optimal age over erasure channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 335–339.
- [11] P. Parag, A. Taghavi, and J.-F. Chamberland, "On real-time status updates over symbol erasure channels," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.
- [12] E. T. Ceran, D. Gündüz, and A. Gyöngy, "Average age of information with hybrid ARQ under a resource constraint," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.
- [13] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, "On timely channel coding with hybrid ARQ," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [14] S. Feng and J. Yang, "Age-optimal transmission of rateless codes in an erasure channel," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
- [15] R. Wang, Y. Gu, H. Chen, Y. Li, and B. Vucetic, "On the age of information of short-packet communications with packet management," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [16] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid ARQ in wireless networked control," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3490–3504, May 2020.
- [17] J. Yun, C. Joo, and A. Eryilmaz, "Optimal real-time monitoring of an information source under communication costs," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 4767–4772.
- [18] J. Chakravorty and A. Mahajan, "Distortion-transmission trade-off in real-time transmission of Gauss–Markov sources," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1387–1391.
- [19] A. Mitra, J. A. Richards, S. Bagchi, and S. Sundaram, "Finite-time distributed state estimation over time-varying graphs: Exploiting the age-of-information," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2019, pp. 4006–4011.
- [20] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, Mar. 2019.
- [21] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4019–4034, Jun. 2021.
- [22] B. Han, S. Yuan, Z. Jiang, Y. Zhu, and H. D. Schotten, "Robustness analysis of networked control systems with aging status," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Jul. 2020, pp. 1360–1361.
- [23] K. Gatsis, H. Hassani, and G. J. Pappas, "Latency-reliability tradeoffs for state estimation," *IEEE Trans. Autom. Control*, vol. 66, no. 3, pp. 1009–1023, Mar. 2021.
- [24] M. Bastopcu and S. Ulukus, "Age of information for updates with distortion," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Aug. 2019, pp. 1–5.
- [25] L. Rambault, E. Etien, S. Cautet, G. Champenois, and O. Bachelier, "Linearization by using two methods: Exact input-output linearization and sliding mode control," in *Proc. Eur. Control Conf. (ECC)*, Sep. 2001, pp. 1040–1043.
- [26] K. J. Astrom and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton, NJ, USA: Princeton Univ. Press, Apr. 2008.
- [27] K. H. Ang, G. Chong, and Y. Li, "PID control system analysis, design, and technology," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 4, pp. 559–576, Jul. 2005.
- [28] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksals, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

- [29] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jun. 2014.
- [30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, Sep. 2006.
- [31] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3309–3338, Jun. 2012.
- [32] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
- [33] J. Gutiérrez-Gutiérrez, M. Zárrega-Rodríguez, F. Villar-Rosety, and X. Insausti, "Rate-distortion function upper bounds for Gaussian vectors and their applications in coding AR sources," *Entropy*, vol. 20, no. 6, p. 399, May 2018.
- [34] M. Shirvanimoghaddam *et al.*, "Short block-length codes for ultra-reliable low latency communications," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 130–137, Feb. 2019.
- [35] E. Björnson, E. A. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 14–23, Nov. 2014.
- [36] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, "Timely estimation using coded quantized samples," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 1812–1817.



Stefan Roth (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degree in electrical engineering and information technology from Ruhr University Bochum, Germany, in 2016 and 2019, respectively, where he is currently pursuing the Ph.D. degree with the Institute of Digital Communication Systems.

From September 2016 to September 2017, he studied at the University of Birmingham, U.K. From June to August 2019, he was a Visiting Student at Princeton University, NJ, USA. His research interest

includes physical layer security and communication in the internet of things (IoT).



Ahmed Arafa (Member, IEEE) received the B.Sc. degree (Hons.) in electrical engineering from Alexandria University, Egypt, in 2010, the M.Sc. degree in wireless technologies from the Wireless Intelligent Networks Center (WINC), Nile University, Egypt, in 2012, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, USA, in 2016 and 2017, respectively.

He has been with the Department of Electrical Engineering, Princeton University, as a Post-Doctoral Research Associate, from 2017 to 2019. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, The University of North Carolina at Charlotte. His research interests include communication theory, information theory, machine learning, and signal processing, with recent focus on federated learning, timely information processing and transfer (age-of-information), information-theoretic security and privacy, and energy harvesting communications. He was a recipient of the Distinguished Dissertation Award from the Department of Electrical and Computer Engineering, University of Maryland, in 2017, for his Ph.D. thesis work on optimal energy management policies in energy harvesting communication networks with system costs.



Aydin Sezgin (Senior Member, IEEE) received the Dipl. Ing. (M.S.) degree in communications engineering from Technische Fachhochschule Berlin (TFH), Berlin, in 2000, and the Dr. Ing. (Ph.D.) degree in electrical engineering from TU Berlin, in 2005.

From 2001 to 2006, he was with the Heinrich-Hertz-Institut, Berlin. From 2006 to 2008, he held a post-doctoral position and was also a Lecturer with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA, USA. From 2008 to 2009, he held

a post-doctoral position at the Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA, USA. From 2009 to 2011, he was the Head of the Emmy-Noether-Research Group on Wireless Networks, Ulm University. In 2011, he joined TU Darmstadt, Germany, as a Professor. He is currently a Professor of information systems and sciences with the Department of Electrical Engineering and Information Technology, Ruhr-Universität Bochum, Germany. He is interested in signal processing, communication, and information theory, with a focus on wireless networks. He has published several book chapters more than 50 journals and 170 conference papers in these topics. He has coauthored a book on multi-way communications. He is a winner of the ITG-Sponsorship Award in 2006. He was a first recipient of the prestigious Emmy-Noether Grant by the German Research Foundation in communication engineering in 2009. He has coauthored papers that received the Best Poster Award at the IEEE Communication Theory Workshop in 2011, the Best Paper Award at ICCSPA in 2015, and the Best Paper Award at ICC in 2019. He has served as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2009 to 2014.



H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University in 1977.

From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at

Berkeley and Cambridge. His research interests are in the areas of information theory, machine learning and network science, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the forthcoming book *Machine Learning and Wireless Communications* (Cambridge University Press). Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences and is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the IEEE Alexander Graham Bell Medal in 2017.