Predictive Control and Communication Co-Design via Two-Way Gaussian Process Regression and AoI-Aware Scheduling

Abanoub M. Girgis , *Student Member, IEEE*, Jihong Park , *Member, IEEE*,
Mehdi Bennis , *Fellow, IEEE*, and Mérouane Debbah, *Fellow, IEEE*

Abstract—This article studies the joint problem of uplink-downlink scheduling and power allocation for controlling a large number of control systems that upload their states to remote controllers and download control actions over wireless links. To overcome the lack of wireless resources, we propose a machine learning-based solution, where only one control system is controlled, while the rest of the control systems are actuated by locally predicting the missing state and/or action information using the previous uplink and/or downlink receptions via a Gaussian process regression (GPR). This GPR prediction credibility is determined using the age-of-information (AoI) of the latest reception. Moreover, the successful reception is affected by the transmission power, mandating a co-design of the communication and control operations. To this end, we formulate a network-wide minimization problem of the average AoI and transmission power under communication reliability and control stability constraints. To solve the problem, we propose a dynamic control algorithm using the Lyapunov drift-plus-penalty optimization framework. Numerical results corroborate that the proposed algorithm can stably control 2x more number of actuators compared to an event-triggered scheduling baseline with Kalman filtering and frequency division multiple access, which is 18x larger than a round-robin scheduling baseline.

Index Terms—Predictive control, Internet of Things (IoT), age of information (AoI), Gaussian process regression (GPR), ultra-reliable and low-latency (URLLC), beyond 5G (B5G), 6G.

I. Introduction

ULTRA-RELIABLE and low-latency communication (URLLC) is a key enabler for ensuring the stability of wirelessly networked control systems in real-time [2], [3]. By physically decoupling sensors, actuators, and controllers, the control system can exploit the recent progress in the fifth generation (5G) connectivity [4], machine learning, and edge computing [5], thereby spearheading many emerging applications ranging from large-scale smart industrial internet of things (IIoT) [6] to autonomous platooning [7]. The success of these applications relies on addressing several fundamental challenges emanating from unstable and intermittent wireless connectivity, which incurs distorted and delayed control information receptions, degrading control stability. Wireless resource allocation and scheduling are thus instrumental in not only improving communication efficiency but also in guaranteeing control stability.

Indeed, in the literature of wireless networked control systems (WNCS), resource allocation and scheduling of control systems play an important role as we shall review next. Static round-robin scheduling in [8], [9] is presumably the simplest algorithm in which each sensor/controller periodically transmits the state/action to a controller/actuator with fixed transmission power in a predefined repeating order, ignoring the communication and control dynamics. Dynamic round-robin scheduling in [10], [11] improves this idea by additionally adjusting the transmission power according to channel state information (CSI) in a way that the power control minimizes the system energy while ensuring communication reliability. Nonetheless, it still relies on the round-robin scheduler that is neither communication efficient nor control dynamics aware. In fact, only faction of control systems that are currently unstable require immediate state/action updates, and prioritizing them can significantly improve the communication efficiency and control stability of the entire control systems under limited wireless resources. To efficiently utilize wireless resources, event-triggered scheduling in [12], [13] carries out dynamic scheduling decisions in which each control system determines whether to transmit its state/action or not based on its last state and action information or the latest control dynamics.

Although effective in supporting a small number of control systems, the aforementioned frameworks [8]–[13] do not jointly design the communication and control operations, which are thus hardly scalable as pointed out in [14]. Furthermore, the scheduling decisions [8], [9], [12], [13] are not adaptive to CSI, which compromises the control stability particularly in large-scale systems. To fill this

void, communication and control co-design has been studied in [15]–[24]. To mention a few, opportunistic dynamic scheduling in [19]–[21] adjusts the transmission power allocations and scheduling decisions according to the given CSI to minimize the total power consumption while ensuring control stability under limited wireless resources. However, they ignore transmission latency, resulting in compromising the overall system control stability or wasting wireless resources. To fix this problem, control-aware scheduling in [22] carries out the scheduling decisions, based on communication reliability reflecting both control and channel states, thereby minimizing the overall transmission latency of the scheduled control systems.

Motivated by these prior works and recent advances in machine learning, in this paper we aspire to further improve the scalability and communication efficiency of WNCS to support a larger number of control systems under limited wireless bandwidth and transmit power. To this end, departing from the existing methods [15]–[24] whose control operations and scheduling decisions are *reactive* to the current control stability and CSI, we develop a *predictive* WNCS where the controller and each actuator locally predict their future states and actions, respectively, based on previously received data by individually running the Gaussian process regression (GPR) mechanisms at the controller and actuator, a Bayesian machine learning framework [25]. The prediction uncertainty of GPR depends significantly on the last received data's freshness, i.e., age of information (AoI) [26], in the uplink (UL) from each sensor to the controller and in the downlink (DL) from the controller to each actuator. To independently optimize the uncertainties of the controller and each actuator, we consider a UL-DL decoupled scheduler. In the UL, the scheduler serves only the sensor of a control system whose state prediction uncertainty and control instability are intolerable. Likewise, in the DL, the scheduler serves only the actuator of a control system whose action prediction uncertainty and control instability are unacceptable. Without scheduling the rest of the sensors and actuators, their corresponding control systems are operational by using locally predicted states and actions, thereby significantly reducing wireless resources without compromising control stability. Here, control stability depends on the GPR prediction uncertainty [27] that is determined by the AoI of the latest received signal. Meanwhile, the reliability of the received signal is dictated by the signal transmission power, highlighting the importance of co-designing communication, control, and machine learning operations.

Given the aforementioned predictive WNCS architecture, we focus on the problem of jointly optimizing the UL-DL scheduling and power allocation to minimize the network-wide average AoI and transmission power, while guaranteeing communication reliability and control stability. To solve the formulated non-convex stochastic optimization problem, we develop a dynamic control algorithm using Lyapunov optimization. Considering an inverted pendulum, numerical results demonstrate that the proposed scheduling method can stably support 2x more control systems compared to the event-triggered scheduling with Kalman filtering and frequency division multiple access (FDMA), which is 18x larger than

a time-triggered scheduling baseline. Furthermore, the results show that the proposed predictive control algorithm is more communication efficient while achieving faster control stability than the time-triggered and event-triggered control baselines, highlighting the effectiveness of the UL-DL decoupled scheduling and the use of two-way GPRs at both controller and actuator sides.

Note that this work has been extended from its conference version that has first proposed the GPR-based predictive WCNS framework [1]. Compared to that preliminary study where UL-DL scheduling is coupled, in this work we consider a UL-DL decoupled scheduler that substantially changes the analysis on communication reliability, control stability, and GPR prediction uncertainty, all of which are intertwined. Furthermore, the prior work only presents the analytic results while omitting their proofs, and demonstrates a couple of simulation examples. In contrast, in this work, we provide not only the details of all the derivations but also extensive simulation results and ablation studies including the comparison between analog and digital transmissions, which clarifies the theoretical contributions of this work while advocating the feasibility of the proposed WCNS framework and its scheduling solution. The remainder of this paper is organized as follows. In Section II, we specify the WNCS architecture including the system models of the control, communication, and GPR-based approach. In Section III, we formulate the communication, control, GPR-based co-design optimization problem and propose the stability-aware scheduling algorithm by leveraging the Lyapunov optimization framework to solve the co-design problem in Section IV. In Section V and Section VI, we present simulation results, and conclude the paper.

II. System Model

A. Wireless Networked Control System Architecture

As depicted in Fig. 1, the WNCS architecture under study consists of a set $\mathcal{M}$ of M independent linear *control systems* over a shared wireless channel. Each control system comprises a *plant*, a *sensor* that measures the plant's state, and an *actuator* that takes an action to control the plant's state. The action is computed by a remote *controller* based on the control and channel states. To this end, the plant's state is received by the controller in the *uplink*, and the controller's action is received by the actuator in the *downlink*. To avoid interference under limited communication bandwidth, each UL or DL channel is allocated to a single sensor-controller pair or controller-actuator pair per unit time, respectively. The rest of the sensor-controller and controller-actuator pairs without receptions locally predict their missing control states and actions, respectively, based on their previously received information, to be detailed in Sec. II-D. To be specific, the plant's state of control system $i \in \mathcal{M}$ at discrete control time $k \in \mathbb{Z}_+$ is denoted by $\mathbf{x}_{i,k}^u \in \mathbb{R}^D$. For a received action at an actuator $\mathbf{u}_{i,k}^a \in \mathbb{R}^P$ based on the computed action at controller $\mathbf{u}_{i,k}^d \in \mathbb{R}^P$, the state evolution of control system i at time k is described by the discrete-time linearized state-
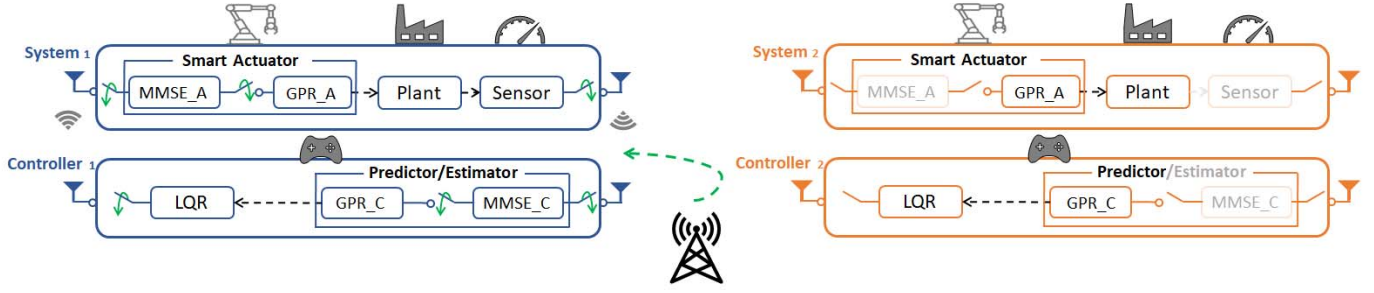
Fig. 1. An illustration of $M = 2$ WNCSs operated via both state/action measurement by remote sensors/controllers and state/actions prediction by GPR.

space representation as follows:

$$\mathbf{x}_{i,k+1}^{u} = \mathbf{A}_i\mathbf{x}_{i,k}^{u} + \mathbf{B}_i\mathbf{u}_{i,k}^{a} + \mathbf{w}_k, \tag{1}$$

where $\mathbf{A}_i \in \mathbb{R}^{D \times D}$ is a fixed state transition matrix of the i-th control system, $\mathbf{B}_i \in \mathbb{R}^{D \times P}$ is a fixed control action matrix of the i-th control system, and $\mathbf{w}_k \in \mathbb{R}^D$ is the plant noise at time k which is independent and identically distributed (IID) Gaussian noise with zero mean and covariance matrix $\mathbf{W}$. Note the plant noise also incorporates an error that comes from the linearization of the non-linear system dynamics. Here, to avoid a non-trivial problem, $\mathbf{A}_i$ is assumed to be unstable, i.e., $\mathbf{A}_i$'s spectral radius $\rho(\mathbf{A}_i) = \max\{|\lambda_1(\mathbf{A}_i)|, \cdots, |\lambda_D(\mathbf{A}_i)|\}$ is larger than unity, where $\lambda_D(\mathbf{A}_i)$ is the D-th eigenvalue of $\mathbf{A}_i$. This implies that the plant's state infinitely grows over time unless a proper control action $\mathbf{u}_{i,k}^{a}$ is provided. To stabilize such control system, each time k, the following four phase operations are considered.

1) *Sensing and Uplink Transmission (at a Sensor):* A centralized scheduler located at the base station (BS) shared among all control systems decides which sensor-controller pair is scheduled to transmit and close its sensing loop based on both the channel and control states. Then, the scheduled sensor transmits its state to its controller over a wireless UL fading channel using analog uncoded communication to be elaborated in Sec. II-B.

2) *State Reception or Prediction (at a Controller):* If the sensor-controller pair is scheduled, the controller obtains the current estimated state using the minimum mean square error (MMSE) estimator, and predicts the next state via GPR to be discussed in Sec. II-D. Otherwise, the controller directly predicts the current and the next states based on the state history using GPR. The current predicted state by GPR is fed to the linear quadratic regulator (LQR) to calculate the action unless the estimated state by the MMSE estimator is provided. The future predicted state is fed to the centralized scheduler to make the scheduling decisions.

3) *Action Computation and Downlink Transmission (at a Controller):* For a given control state, the controller computes the optimal action using LQR [28]. The controller transmits the computed action to the scheduled actuator over a wireless DL fading channel using analog uncoded communication to be elaborated in Sec. II-B.

4) *Action Reception or Prediction (at an Actuator):* If the controller-actuator pair is scheduled, the actuator obtains the currently estimated action using the MMSE estimator, and predicts the next action via GPR to be discussed in Sec. II-D. Otherwise, the actuator directly predicts the current and next actions based on the action history using GPR. For a given action, the actuator takes an action and subsequently, the plant's state is updated according to the control system dynamics in (1).

B. State and Action Communications

The UL state and DL action communications are elaborated, in terms of the received signal, signal-to-noise ratio (SNR), scheduling, and age-of-information (AoI) as follows.

Noisy State and Action Receptions: At the control time slot k, the received signal $\mathbf{y}_{i,k}^{l}$ at the l-th communication at the receiver of control system i is represented as

$$\mathbf{y}_{i,k}^{l} = \sqrt{P_{i,k}^{l}}\mathbf{C}_i\mathbf{H}_{i,k}^{l}\mathbf{q}_{i,k}^{l} + \mathbf{n}_k^{l}, \tag{2}$$

where $l \in \{u, d\}$ represents a communication indicator between the transmitter-receiver pair in which $l = u$ refers to the UL state communication between the sensor-controller pair while $l = d$ refers to the DL action communication between the controller-actuator pair. The transmitted signal, in the UL state communication (i.e., $\mathbf{q} = \mathbf{x}$ and $l = u$), $\mathbf{x}_{i,k}^{u} = [x_{i,k}^{u}(1) \cdots x_{i,k}^{u}(D)]$ is the plant's state transmitted by the sensor of a control system i at time k such that $\mathbb{E}\{|x_{i,k}^{u}(\iota)|^2\} = 1, \forall \iota \in \{1, \cdots, D\}$. The transmitted signal, in the DL action communication (i.e., $\mathbf{q} = \mathbf{u}$, $l = d$), $\mathbf{u}_{i,k}^{d} = [u_{i,k}^{d}(1) \cdots u_{i,k}^{d}(P)]$ is the action transmitted by the controller to an actuator of a control system i at time k such that $\mathbb{E}\{|u_{i,k}^{d}(p)|^2\} = 1, \forall p \in \{1, \cdots, P\}$. The matrix $\mathbf{H}_{i,k}^{l} \in \mathbb{R}^{\mathcal{F} \times \mathcal{F}}$ represents the wireless channel of the l-th communication between the transmitter-receiver pair of a control system i at time k, and $\mathcal{F} \in \{D, P\}$ represents the dimensions of the transmitted state or action, respectively. The channel is modeled as a Rayleigh block fading which is static and flat-fading within either UL or DL transmission time. The channel information is assumed to be known at the transmitters and receivers through the standard channel probing and estimation methods [29], while the channel estimation error analysis is deferred to future work, and $P_{i,k}^{l} \in [0, P_{max}^{l}]$ is the transmission power of control system i at time k with total transmission power P_{max}^{l}. Lastly, $\mathbf{n}_k^{l}$ is the additive white

Gaussian noise at the receiver with zero-mean and covariance matrix $\mathbb{E}\{\mathbf{n}_k^{l^T}\mathbf{n}_k^l\} = N_0\mathbf{I}_{\mathcal{F}}$, where N_0 is the measurement noise power spectral density, and $\mathbf{I}_{\mathcal{F}}$ is the $\mathcal{F} \times \mathcal{F}$ identity matrix. The matrix $\mathbf{C}_i \in \mathbb{R}^{\acute{\mathcal{F}} \times \acute{\mathcal{F}}}$ is the observation matrix of the control system i that equals $\mathcal{F} \times \mathcal{F}$ identity matrix in the DL action communication while equals $\acute{\mathcal{F}} \times \mathcal{F}$ rectangular matrix in the UL state communication to characterizes the minimum states that fully describes the control system. The measurement noise $\mathbf{n}_k^l$ and the plant noise $\mathbf{w}_k$ in (1) are uncorrelated zero-mean Gaussian noise with unit variance. Hence, the SNR at the receiver of the l-th communication of a control system i at time k is

$$\text{SNR}_{i,k}^l = \frac{P_{i,k}^l \|\mathbf{H}_{i,k}^l\|^2}{N_0\,\omega}, \tag{3}$$

where ω is the bandwidth and the SNR in (3) is equivalent to the signal-to-distortion ratio (SDR) as a result of the channel theoretical limit in which the rate-distortion function of the source equals the channel capacity and the number of source samples is matched to the number of channels under analog uncoded communications [30]. In this work, we consider analog uncoded communications, in which the discrete-time continuous amplitude source samples are amplified and transmitted to the receiver over wireless channels. Compared to digital coded communications, uncoded communications are favorable for achieving low latency thanks to skipping channel encoding, at the cost of increasing the transmission power to ensure communication reliability [30], [31]. Furthermore, analog uncoded communication is robust to channel conditions and performs well at different SNRs compared to digital communication that is sensitive to any degradation in channel conditions. To ensure reliable communication for control stability, the successful detection of the transmitted signal is described by the indicator function $\mathbb{I}_{\{\text{SNR}_{i,k}^l \geq \text{SNR}_{\text{th}}^l\}}$ for a target SNR threshold SNR_{th}^l.

Scheduling and AoI: At the time k, the centralized scheduler located at the BS and shared among all control systems schedules at most one sensor-controller pair of control system i in the UL state communication, and most one controller-actuator pair of the control system i in the DL action communication, while the unscheduled pairs are controlled by utilizing the GPR at the controller/actuator sides. Let $\alpha_{i,k}^l \in \{0,1\}$ be the scheduling variable of the l-th communication of the control system i at time k, where $\alpha_{i,k}^l = 1$ when the transmitter-receiver pair of the l-th communication of control system i is scheduled at time k and $\alpha_{i,k}^l = 0$ otherwise.

The freshness of the received information is measured using AoI, i.e., the elapsed time since the generation of the latest received information [26]. AoI is composed of the inter-arrival time that is defined as the time elapsed between two consecutive update generations and the service time defined as the transmission time of update information. In analog uncoded communications, AoI depends only on the inter-arrival time since the service time is deterministic based on channel bandwidth. Hence, the AoI of the l-th communication of the control system i at the receiver linearly increases with time if it is not scheduled or its SNR is below a threshold. Formally, the AoI of the l-th communication of control system

i at the receiver is given as:

$$\beta_{i,k}^l = \begin{cases} 1 + \beta_{i,k-1}^l, & \text{if } \xi_{i,k}^l = 0, \\ 1, & \text{o.w.,} \end{cases} \tag{4}$$

where $\beta_{i,k}^l \in \mathbb{Z}_{++}$ is the AoI of the l-th communication of the control system i at time k at the receiver, and $\xi_{i,k}^l = \alpha_{i,k}^l \mathbb{I}_{\{\text{SNR}_{i,k}^l \geq \text{SNR}_{th}^l\}}$ is the transmission indicator variable of l-th communication of control system i at time k that depends on both scheduling variable and SNR indicator function.

C. State and Action Estimation Over Noisy Communications

The UL received states and DL received actions are distorted by Rayleigh fading channels. The transmitted signals are estimated using the MMSE estimator as detailed next. When one transmitter-receiver pair of the l-th communication of control system i is scheduled, i.e., $\alpha_{i,k}^l = 1$, the receiver applies the MMSE estimator to restore the transmitted signal from the noisy received signal in (2). The resultant estimated signal $\tilde{\mathbf{q}}_{i,k}^l$ is given as

$$\tilde{\mathbf{q}}_{i,k}^l = \mathbb{E}\{\mathbf{q}_{i,k}^l | \mathbf{y}_{i,k}^l\} = \mathbf{G}_{i,k}^l \mathbf{y}_{i,k}^l = \mathbf{q}_{i,k}^l + \mathbf{v}_{i,k}^l, \tag{5}$$

where $\mathbf{G}_{i,k}^l \in \mathbb{R}^{\mathcal{F} \times \mathcal{F}}$ is the linear MMSE matrix at the receiver of the l-th communication of the control system i at time k that minimizes the mean-squared error (MSE) between the original and estimated signals as [32]

$$\mathbf{G}_{i,k}^l = \sqrt{P_{i,k}^l} \mathcal{S}_q \mathbf{H}_{i,k}^{l^T} \left(P_{i,k}^l \mathbf{H}_{i,k}^l \mathcal{S}_q \mathbf{H}_{i,k}^{l^T} + N_0\mathbf{I}_{\mathcal{F}}\right)^{-1} \tag{6}$$

The term $\mathbf{v}_{i,k}^l$ in (5) is the MMSE estimation error following a zero-mean Gaussian random vector with the covariance matrix $\mathbf{V}_{i,k}^l \in \mathbb{R}^{\mathcal{F} \times \mathcal{F}}$. Following [33], we assume that $\mathbf{q}_{i,k}^l$ follows a zero-mean Gaussian distribution with the covariance matrix $\mathcal{S}_q \in \mathbb{R}^{\mathcal{F} \times \mathcal{F}}$, then we have

$$\begin{aligned} \mathbf{V}_{i,k}^l = \mathbb{E}\{\mathbf{v}_{i,k}^l \mathbf{v}_{i,k}^{l^T}\} &= \mathbb{E}\left\{\left(\tilde{\mathbf{q}}_{i,k}^l - \mathbf{q}_{i,k}^l\right)\left(\tilde{\mathbf{q}}_{i,k}^l - \mathbf{q}_{i,k}^l\right)^T\right\} \\ &= \mathcal{S}_q - \mathbf{G}_{i,k}^l \sqrt{P_{i,k}^l} \mathbf{H}_{i,k}^l \mathcal{S}_q \end{aligned} \tag{7}$$

D. State and Action Prediction Without Communication

When one transmitter-receiver pair of the l-th communication of the control system i at time k is not scheduled, i.e., $\alpha_{i,k}^l = 0$, a receiver applies parallel GPRs proportional to the missing signal dimensions to predict both the missing current signal and the next signal using the previously received signals. Each individual GPR learns the functional relationship $g \in \mathbb{R}$ between the control discrete-time $k' \in \mathbb{Z}_+$ and each output of the received signal. This means that each output of the MMSE estimated signal $\tilde{\mathbf{q}}_{i,k'}^l$ in (5) is the state observation $\tilde{\mathbf{x}}_{i,k'}^u \in \mathbb{R}^D$ in the UL and the action $\tilde{\mathbf{u}}_{i,k'}^d \in \mathbb{R}^P$ in the DL. This is accomplished by assuming the latent function g is a random variable and drawn from a Gaussian distribution, as well as, any finite subset of these random variables that are taken from a Gaussian distribution has a joint Gaussian distribution. In this way, we learn a latent function of the following regression model $\tilde{q}_{i,k'}^l(j) = g_j(k') + \epsilon, \quad j \in \{1, \cdots, \mathcal{F}\}, \quad \forall i, l, k'$, where

$\tilde{q}_{i,k'}^l(j) \in \mathbb{R}$ is the j-th output of the estimated signal of the l-th communication of the control system i at time k', g_j is a j-th output latent function, and $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ is an IID Gaussian noise distribution with zero mean and variance σ_n^2 that accounts for the measurements or modeling errors [34]. Specifically, to predict the missing received signal $\tilde{\mathbf{q}}_{i,k}^l$ of the l-th communication of the control system i at test time k, we exploit $\mathcal{F}$ individual GPRs, and feeding each individual GPR with a training set $\mathcal{D}_{i,n_l}^{l,j}$ of each output of the previously received signals associated with its observation time k', given as $\mathcal{D}_{i,n_l}^{l,j} = \{(k', \xi_{i,k'}^l \, \tilde{q}_{i,k'}^l(j)) \,|\, j = 1, \cdots, \mathcal{F}, \; k' = 1, \cdots, n_l, \; i = 1, \cdots, M, \; l \in \{u, d\}\}$. Here, $n_l = \sum_{k'} \xi_{i,k'}^l$ counts the number of received signals of the l-th communication until time k' in the training set of the control system i. Hence, the last time instant in which the transmitter of the l-th communication of the control system transmitted its observation to the receiver is given as $\tilde{n}_l = k' - \beta_{i,k'}^l + 1$. It is obvious that a large value of AoI decreases the number of observations at the receiver that affects the signal prediction credibility at a particular level.

In each individual GPR, according to the Gaussian process (GP) characteristics where any finite subset of random variables taken from a realization of a GP follows a joint Gaussian distribution, each j-th output latent function g_j of the vector-valued latent function $\mathbf{g}(k) = [g_1(k) \cdots g_{\mathcal{F}}(k)]$ is assumed to follow a GP as $g_j(k) \sim \mathcal{GP}(m_j(k), \mathcal{R}_j(k, k'))$, where $m_j(k)$ is the mean function of the j-th output of the missing received signal which is usually taken as zero without loss of generality [25], and $\mathcal{R}_j(k, k')$ is the covariance function of the j-th output of the missing received signal between the outputs at time k and k' that defines the correlation between the outputs according to the input times. It is noted that the stationary covariance function between the outputs is based on the difference between their corresponding input times $|k - k'|$ in which the two outputs are strongly correlated if their corresponding input times are sufficiently close to each other. Since we focus on time-series data, we utilize information from previously received signals to describe the current data depending on the past observations. Hence, we use a squared exponential kernel function coupled with a periodic kernel function, to model the correlation between the outputs according to their temporal behaviors, as defined in [25]

$$
\mathcal{R}(k, k') = h_q^2 \exp\left[\frac{-(k - k')^2}{2h_k^2}\right] \\
+ \exp\left\{-2\sin^2\left[\nu\pi(k - k')\right]\right\}, \quad (8)
$$

where the first term represents the stationary covariance function that depends on when the signal $|k - k'|$ was received with h_k and h_q being hyperparameters representing the time-scaling and output-scaling of a squared exponential function, respectively, and the second term gives the periodicity with hyperparameter ν representing frequency. For a set of j-th output observations $\tilde{\mathbf{q}}_i^l(j) = \{\tilde{q}_{i,1}^l(j), \cdots, \tilde{q}_{i,n_l}^l(j)\}^T$ and the associated observation times $\mathbf{k}' = \{1, \cdots, n_l\}^T$, the joint distribution of the j-th output past observations $\tilde{\mathbf{q}}_{i,k'}^l(j)$ together

with the j-th output $g_j(k)$ at test time k is given as

$$
\begin{bmatrix} \tilde{\mathbf{q}}_i^l(j) \\ g_j(k) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{R}_j(\mathbf{k}', \mathbf{k}') & \mathbf{r}_j(\mathbf{k}', k) \\ \mathbf{r}_j(k, \mathbf{k}') & \mathcal{R}_j(k, k) \end{bmatrix} \right), \quad (9)
$$

where $\mathcal{R}_j(k, k) \in \mathbb{R}$ is the prior covariance function of j-th output observation at a test time k, and $\mathbf{R}_j(\mathbf{k}', \mathbf{k}') \in \mathbb{R}^{n_l \times n_l}$ is the symmetric and positive semi-definite covariance matrix of j-th output past observations with the elements $\mathcal{R}_j(\mathbf{k}'(a), \mathbf{k}'(b))$ for $a, b = 1, \cdots, n_l$. Following [34], we treat the prediction mean as the j-th output predicted signal $\hat{q}_{i,k}^l(j)$, the posterior distribution of $g_j(k)$ at test time k based on the training set $\mathcal{D}_{i,n_l}^{l,j}$ can be analytically derived as

$$
\Pr\left(g_j(k) \,|\, \mathcal{D}_{i,n_l}^{l,j}, k, \mathbf{\Theta}_j\right) \sim \mathcal{N}\left(\hat{q}_{i,k}^l(j), \sigma_{i,k}^2(j)\right). \quad (10)
$$

Following [34], the j-th output prediction mean $\hat{q}_{i,k}^l(j)$, and the j-th output prediction variance $\sigma_{i,k}^2(j)$ are respectively given as

$$
\hat{q}_{i,k}^l(j) = \mathbf{r}_j(k, \mathbf{k}')^T \mathbf{R}_j(\mathbf{k}', \mathbf{k}')^{-1} \tilde{\mathbf{q}}_i^l(j) = q_{i,k}^l(j) + e_{i,k}^l(j), \quad (11)
$$

$$
\begin{aligned}
\sigma_{i,k}^2(j) &= \mathbb{E}\{e_{i,k}^l(j) e_{i,k}^{l\,T}(j)\} \\
&= \mathbb{E}\left\{ \left(\hat{q}_{i,k}^l(j) - q_{i,k}^l(j)\right) \left(\hat{q}_{i,k}^l(j) - q_{i,k}^l(j)\right)^T \right\} \\
&= \mathcal{R}_j(k, k) - \mathbf{r}_j(k, \mathbf{k}') \mathbf{R}_j(\mathbf{k}', \mathbf{k}')^{-1} \mathbf{r}_j(k, \mathbf{k}')^T, \quad (12)
\end{aligned}
$$

where $\mathbf{r}_j(\mathbf{k}', k) \in \mathbb{R}^{n_l \times 1}$ is j-th output observation covariance between the outputs at the n_l observation times and a test time k, and the term $e_{i,k}^l(j)$ is j-th output prediction error defined as the difference between true and predicted outputs. Moreover, $\mathbf{\Theta}_j$ in (10) is the j-th output hyperparameters of the covariance function $\mathcal{R}$. Finally, the predicted signal at the receiver of the l-th communication of control system i at time k and its prediction error covariance matrix are

$$
\hat{\mathbf{q}}_{i,k}^l = \{\hat{q}_{i,k}^l(1) \cdots \hat{q}_{i,k}^l(\mathcal{F})\}^T = \mathbf{q}_{i,k}^l + \mathbf{e}_{i,k}^l, \quad (13)
$$

$$
\mathcal{J}_{i,k}^l = \begin{bmatrix} \sigma_{i,k}^2(1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{i,k}^2(\mathcal{F}) \end{bmatrix}. \quad (14)
$$

E. Action Computation and Actuation

By feeding the estimated or predicted state, the controller computes the action using LQR. Then, the actuator applies the estimated or predicted action to stabilize state as detailed next.

Action Computation After State Estimation/Prediction: For a given estimated state (i.e., MMSE output) in (5) or predicted state (i.e., GPR output) in (13), at time k, the state $\mathbf{x}_{i,k}^c$ available at the controller based on the UL transmission indicator variable is given as

$$
\mathbf{x}_{i,k}^c = \xi_{i,k}^u \tilde{\mathbf{x}}_{i,k}^u + \left(1 - \xi_{i,k}^u\right) \hat{\mathbf{x}}_{i,k}^u. \quad (15)
$$

The received state $\mathbf{x}_{i,k}^c$ is used in the LQR located at the controller, and the optimal action of a control system i at time k is given by the following linear feedback control law as

$$
\mathbf{u}_{i,k}^d = -\mathbf{\Phi}_i \mathbf{x}_{i,k}^c, \quad (16)
$$

where $\mathbf{u}_{i,k}^d \in \mathbb{R}^P$ is the computed action at the controller, $\mathbf{\Phi}_i = \left(\mathbf{Z}^u + \mathbf{B}_i^T \mathbf{P} \mathbf{B}_i\right)^{-1} \mathbf{B}_i^T \mathbf{P} \mathbf{A}_i$ is the feedback gain matrix

of the control system i, $\mathbf{Z}^s \in \mathbb{S}_+^{D \times D}$ is a positive semi-definite weight matrix of the state deviation cost, and $\mathbf{Z}^u \in \mathbb{S}_{++}^{P \times P}$ is a positive definite weight matrix of the action cost. The term $\mathbf{P} = \mathbf{A}_i^T \mathbf{P} \mathbf{A}_i - \mathbf{A}_i^T \mathbf{P} \mathbf{B}_i \left(\mathbf{B}_i^T \mathbf{P} \mathbf{B}_i + \mathbf{Z}^u \right)^{-1} \mathbf{B}_i^T \mathbf{P} \mathbf{A}_i + \mathbf{Z}^s$ is the unique positive definite matrix which satisfies the discrete-time algebraic Riccati equation (DARE). Then, the controller transmits the computed action $\mathbf{u}_{i,k}^d$ in (16) to an actuator of control system i at time k in the DL, if $\xi_{i,k}^d = 1$, as discussed in Sec. II-B.

1) Actuation After Action Estimation/Prediction: For a given estimated control action (i.e., MMSE output) in (5) or predicted action (i.e., GPR output) in (13), at time k, the action $\mathbf{u}_{i,k}^a$ available at the actuator based on the DL transmission indicator variable is given as

$$\mathbf{u}_{i,k}^a = \xi_{i,k}^d \, \tilde{\mathbf{u}}_{i,k}^d + \left(1 - \xi_{i,k}^d\right) \hat{\mathbf{u}}_{i,k}^d. \quad (17)$$

Note that the UL and DL transmission indicator variables are periodically generated by the centralized scheduler in which, within each unit control time duration, the UL state communication can be firstly activated for sensing the plant's state based on the UL transmission indicator variable. Then, the DL action communication can be activated for actuation based on the DL transmission indicator variable. Consequently, for a given pair of UL and DL transmission indicator variables with (15) and (17), the actuator takes a control action that changes the plant's state of the control system i at time k in (1) into four cases of state evolution as follows:

$$\mathbf{x}_{i,k+1}^o = \mathbf{A}_i \mathbf{x}_{i,k}^u - \mathbf{B}_i (\mathbf{\Phi}_i \hat{\mathbf{x}}_{i,k}^u + \mathbf{e}_{i,k}^d) + \mathbf{w}_k, \quad (18)$$

$$\mathbf{x}_{i,k+1}^s = \mathbf{A}_i \mathbf{x}_{i,k}^u - \mathbf{B}_i (\mathbf{\Phi}_i \tilde{\mathbf{x}}_{i,k}^u + \mathbf{e}_{i,k}^d) + \mathbf{w}_k, \quad (19)$$

$$\mathbf{x}_{i,k+1}^a = \mathbf{A}_i \mathbf{x}_{i,k}^u - \mathbf{B}_i (\mathbf{\Phi}_i \hat{\mathbf{x}}_{i,k}^u + \mathbf{v}_{i,k}^d) + \mathbf{w}_k, \quad (20)$$

$$\mathbf{x}_{i,k+1}^c = \mathbf{A}_i \mathbf{x}_{i,k}^u - \mathbf{B}_i (\mathbf{\Phi}_i \tilde{\mathbf{x}}_{i,k}^u + \mathbf{v}_{i,k}^d) + \mathbf{w}_k. \quad (21)$$

where the open-loop in (18) holds if $\xi_{i,k}^u = 1$ and $\xi_{i,k}^d = 0$, the sensing-loop in (19) holds if $\xi_{i,k}^u = 1$ and $\xi_{i,k}^d = 0$, the actuating-loop in (20) holds if $\xi_{i,k}^u = 0$ and $\xi_{i,k}^d = 1$, and the closed-loop in (21) holds if $\xi_{i,k}^u = 1$ and $\xi_{i,k}^d = 1$.

Timing Diagram: Based on the UL and DL transmission indicator variables within each unit control time duration, the timing diagram of a control system is illustrated in Fig. 2. The centralized scheduler shared among all control systems, within each unit control time duration, primarily transmits the UL and DL transmission indicator variables to the sensor, controller, and actuator sides of all control systems. Then, the state is only transmitted by the sensor if the control system has a reliable UL communication and has valuable information affecting the control stability (i.e, $\xi_{i,k}^u = 1$) which results in saving wireless communication resources. After that, LQR located at the controller computes the action based on the state available at the controller in (15) in which the predicted state, if $\xi_{i,k}^u = 0$, is applied to the LQR. Lastly, the action is transmitted by the controller if it has reliable DL communication and valuable information affecting the control stability (i.e, $\xi_{i,k}^d = 1$). This is a result of assuming that the UL and DL transmission indicator variables as being periodically transmitted by the centralized scheduler every control time duration, the controller periodically calculates the action

depending on the available state, and the actuator periodically applies the action depending on the available action. Moreover, the discrete-time control time k equals the continuous-time control time duration unit Δk comprising the UL and DL transmission times while ignoring the computational delay.

III. COMMUNICATION CONTROL CO-DESIGN

A. Control-Constrained Problem Formulation

Our objective is to minimize the total communication cost per control system subject to ensuring communication reliability and control stability. The total communication cost incorporates the AoI and transmission power since the AoI indirectly affects the control stability through the GPR prediction stability and wireless resources consumption, while transmit power affects communication reliability and energy consumption. Formally speaking, we have:

$$\mathcal{C}\left(\{\bar{\beta}_i^l\}, \{\bar{\hat{P}}_i^l\}\right) = \omega_{\beta_l} \sum_{i=1}^M \mathcal{G}_\beta(\bar{\beta}_i^l) + \omega_{P_l} \sum_{i=1}^M \mathcal{G}_P(\bar{\hat{P}}_i^l), \quad (22)$$

$\forall l \in \{u, d\}$, where the non-decreasing concave functions $\mathcal{G}_\beta(\beta) = \log(1 + \beta)$ and $\mathcal{G}_P(\hat{P}) = \log(1 + \hat{P})$ are proportionally fair cost functions of the AoI and the transmission power function for each control system, respectively [35]. The transmission power function that depends on the scheduling variable is given as $\hat{P}_{i,k}^l = \alpha_{i,k}^l P_{i,k}^l$, and the given positive weights ω_{β_l} and ω_{P_l} adjust the relative importance of the corresponding cost functions. Throughout this work, the following notation for the long-term time-averaged of any quantity z is defined as $\bar{z} \triangleq \limsup_{K \to \infty} \frac{1}{K} \sum_{k=1}^K z$. In particular, $\bar{\beta}_i^l$ and $\bar{\hat{P}}_i^l$ are the long-term time-averaged of β_i^l and $\hat{P}_i^l$, respectively.

To evaluate control stability, we consider the quadratic Lyapunov function that measures the performance of each control system as a function of the state expressed as

$$\mathcal{L}(\mathbf{x}_{i,k}^u) = \mathbf{x}_{i,k}^{u^T} \mathcal{Z} \mathbf{x}_{i,k}^u, \quad \forall \mathcal{Z} \in \mathbb{S}_{++}^D, \quad (23)$$

where $\mathcal{Z} \in \mathbb{S}_{++}^D$ is a unique positive definite solution to the discrete Lyapunov equation $\mathbf{A}_i^{c^T} \mathcal{Z} + \mathcal{Z} \mathbf{A}_i^c = -\mathbb{I}_D$, and $\mathbf{A}_i^c$ is a closed-loop state transition matrix defined as $\mathbf{A}_i^c = \mathbf{A}_i - \mathbf{B}_i \mathbf{\Phi}_i$. Because the centralized scheduler has only access to the predicted state, the expected current value of $\mathcal{L}(\mathbf{x}_{i,k}^u)$ is calculated in the following lemma.

Lemma 1: Given the predicted state $\hat{\mathbf{x}}_{i,k}^u$ and the state prediction error covariance matrix $\mathcal{J}_{i,k}^u$ at the controller, the expected current value of $\mathcal{L}(\mathbf{x}_{i,k}^u)$ is given as

$$\mathbb{E}\left[\mathcal{L}(\mathbf{x}_{i,k}^u)|\hat{\mathbf{x}}_{i,k}^u\right] = \|\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + Tr\left[\mathcal{Z} \, \mathcal{J}_{i,k}^u\right], \; \forall \mathcal{Z} \in \mathbb{S}_{++}^D. \quad (24)$$

Proof: Please refer to Appendix.A ∎

Note that the expected current value of $\mathcal{L}(\mathbf{x}_{i,k}^u)$ of the control system i at time k naturally grows as the predicted state, and the prediction error get larger as a result of increasing AoI and/or the insufficiency of received observations number in the training set. Control stability requires that the expected future value of $\mathcal{L}(\mathbf{x}_{i,k+1}^u)$ should decrease at a given rate $\zeta_i \in (0, 1]$ of its expected current value of $\mathcal{L}(\mathbf{x}_{i,k}^u)$, which means the
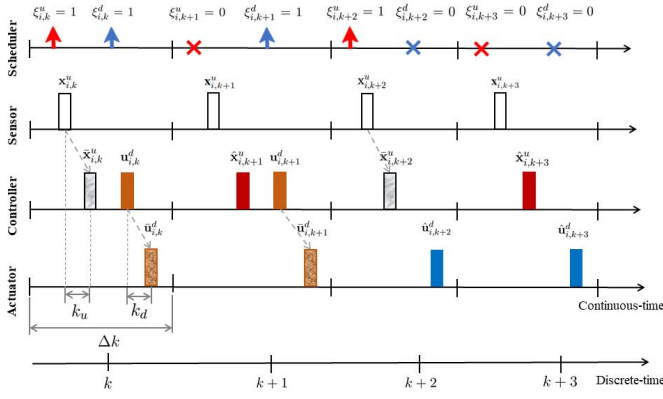
Fig. 2. Timing diagram of the control system. First diagram illustrates the UL and DL transmission indicator variables generated by scheduler, second diagram illustrates uniform sampling by a sensor, third diagram illustrates the received/predicted state and calculated action at a controller. Fourth diagram illustrates the received/predicted action at an actuator.

state of the control system is monotonically decreasing along trajectories, as

$$\mathbb{E}\left[\mathcal{L}(\mathbf{x}_{i,k+1}^u)|\hat{\mathbf{x}}_{i,k}^u, \hat{\mathbf{u}}_{i,k}^d, \mathbf{H}_{i,k}^u, \mathbf{H}_{i,k}^d, P_{i,k}^u, P_{i,k}^d\right]$$
$$\leq \zeta_i \mathbb{E}\left[\mathcal{L}(\mathbf{x}_{i,k}^u)|\hat{\mathbf{x}}_{i,k}^u\right], \quad (25)$$

where the expectation in the right-hand side of (25) is with respect to the plant noise $\mathbf{w}_k$ in (1), the signal estimation error $\mathbf{v}_{i,k}^l$ defined in (7), and the signal prediction error $\mathbf{e}_{i,k}^l$ defined in (13). According to the objective function in (22) and the control stability constraint in (25), the control-constrained optimization problem can be formulated as follows:

$$(\mathcal{P}1) \quad \underset{\mathbf{a}_k^l, \mathbf{P}_k^l}{\text{Minimize}} \; \mathcal{C}\left(\{\bar{\beta}_i^l\}, \{\bar{\hat{P}}_i^l\}\right) \quad (26a)$$

$$\text{subject to: } 0 \leq P_{i,k}^l \leq P_{max}^l, \quad (26b)$$

$$\|\mathbf{H}_{i,k}^l\|^2 P_{i,k}^l / N_0\,\omega \geq \text{SNR}_{th}^l, \quad (26c)$$

$$\alpha_{i,k}^l \in \{0,1\}, \quad (26d)$$

$$\sum_{i=1}^M \alpha_{i,k}^l \leq 1, \quad (26e)$$

$$(25),$$

$\forall l \in \{u,d\}, i \in \mathcal{M}, k$, where $\mathbf{a}_k^l = \{\alpha_{i,k}^l : \forall l \in \{u,d\}, \; i \in \mathcal{M}\}$ and $\mathbf{P}_k^l = \{P_{i,k}^l : \forall l \in \{u,d\}, \; i \in \mathcal{M}\}$ are the UL-DL scheduling vector at time k, and the UL-DL transmission power vector at time k, respectively. The constraint in (26b) bounds the UL-DL transmission power allocation of a control system i at time k by the total transmission power P_{max}^l, while the constraint in (26c) ensures the communication reliability that is based on SNR or SDR in analog uncoded communications. The constraints in (26d)-(26e) ensure at most one transmitter-receiver pair of a control system i is scheduled at time k. The constraint in (25) ensures the state is decreasing along trajectories to satisfy the control stability. It is noted that the control stability constraint

in (25) is independent of communication constraints in (26b)-(26e). However, the control stability constraint is affected and determined by the communication and scheduling variables, hence the original control-constrained problem $\mathcal{P}1$ is rewritten after directly reflecting the communication control relationship of the constraint (25) in the following lemma.

Lemma 2: Given predicted state $\hat{\mathbf{x}}_{i,k}^u$, state prediction error covariance matrix $\mathcal{J}_{i,k}^u$, the predicted action $\hat{\mathbf{u}}_{i,k}^d$, the action prediction error covariance matrix $\mathcal{J}_{i,k}^d$, the channel between the sensor-controller pair $\mathbf{H}_{i,k}^u$, the channel between the controller-actuator pair $\mathbf{H}_{i,k}^d$, the UL transmission power $P_{i,k}^u$, and the DL transmission power $P_{i,k}^d$, the control stability constraint in (25) is satisfied IFF the following conditions on the transmission indicator variables hold, i.e.,

$$\limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \xi_{i,k}^u \geq \limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \frac{\mathbb{A}_{i,k} + \mathbb{B}_{i,k} + \mathbb{M}_{i,k} + Tr[\mathcal{Z}\mathbf{W}]}{\mathbb{J}_{i,k} - \mathbb{K}_{i,k}}, \quad (27)$$

$$\limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \xi_{i,k}^d \geq \limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \frac{\mathbb{A}_{i,k} + \mathbb{B}_{i,k} + \mathbb{M}_{i,k} + Tr[\mathcal{Z}\mathbf{W}]}{\mathbb{M}_{i,k} - \mathbb{N}_{i,k}}, \quad (28)$$

$$\limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \xi_{i,k}^u \xi_{i,k}^d$$

$$\geq \limsup_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \frac{\mathbb{A}_{i,k} + \mathbb{B}_{i,k} + \mathbb{M}_{i,k} + Tr[\mathcal{Z}\mathbf{W}]}{[\mathbb{J}_{i,k} - \mathbb{K}_{i,k}] + [\mathbb{M}_{i,k} - \mathbb{N}_{i,k}]}. \quad (29)$$

where we have further defined the following terms $\mathbb{A}_{i,k} := \|(\mathbf{A}_i^c - \zeta_i\mathbf{I}_D)\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2$, $\mathbb{B}_{i,k} := Tr\left[(\mathbf{A}_i^T\mathcal{Z}\mathbf{A}_i - \zeta_i\mathcal{Z})\mathcal{J}_{i,k}^u\right]$, $\mathbb{J}_{i,k} := Tr\left[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\right]$, $\mathbb{K}_{i,k} := Tr\left[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\right]$, $\mathbb{M}_{i,k} := Tr\left[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\right]$, $\mathbb{N}_{i,k} := Tr\left[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathbf{V}_{i,k}^d\right]$.

Proof: Please refer to Appendix. B ∎

Note that the conditions on the UL and DL transmission indicator variables in (27) and (28), respectively, ensure the control stability constraint in (25) in a decoupled scheduling between the UL and DL communications based on the current predicted control and channel states, while the condition on both the UL and DL scheduling variables in (29) ensures control stability in a coupled scheduling between the UL and DL communications. Intuitively, the growth of AoI at a controller/actuator leads to an increasing in the state/action prediction error due to an outdated training set. Therefore, the transmitter-receiver pair of a control system should be scheduled when it has a reliable communication and the state/action prediction error is greater than the state/action estimation error to ensure control stability.

The actuator is physically decoupled from the centralized scheduler and the controller which is co-located at BS, and the DL indicator variable at the centralized scheduler relies on the action prediction error at the actuator. Hence, the controller leverages another GPR, where the input of this GPR is the discrete-time associated with the generated action by LQR plus the action estimation error as $\bar{\mathbf{u}}_{i,k'}^{d_c} = \mathbf{u}_{i,k'}^d + \mathbf{v}_{i,k'}^d$. As a result of the applied input to this GPR that yields a training set similar to the one at the actuator as $\mathcal{D}_{i,n_d}^{d_c} =$

$\{(k', \xi_{i,k'}^d, \bar{\mathbf{u}}_{i,k'}^{d_c})| \, k' = 1, \cdots, n_d, i = 1, \cdots, M\}$, we obtain the action prediction error similar to the one generated at actuator side.

B. Joint Communication and Control Problem

According to the UL and DL transmission indicator variables constraints in (27)-(29) that result from the control stability constraint in (25) in problem $\mathcal{P}1$, problem $\mathcal{P}1$ is rewritten as

$$(\mathcal{P}2) \quad \underset{\mathbf{a}_k^l, \mathbf{P}_k^l}{\text{Minimize}} \; \mathcal{C}\left(\{\bar{\beta}_i^l\}, \{\hat{\bar{P}}_i^l\}\right) \tag{30a}$$

$$\text{subject to:} \quad \bar{\alpha}_i^u \geq \bar{\mathcal{G}}_{lb}\left(\mathfrak{m}_{i,k}^u\right), \tag{30b}$$

$$\bar{\alpha}_i^d \geq \bar{\mathcal{G}}_{lb}\left(\mathfrak{m}_{i,k}^d\right), \tag{30c}$$

$$\overline{\alpha_i^u \alpha_i^d} \geq \bar{\mathcal{G}}_{lb}\left(\mathfrak{m}_{i,k}\right), \tag{30d}$$

$$(26b) - (26e),$$

$\forall i \in \mathcal{M}$, where $\bar{\alpha}_i^u$ and $\bar{\alpha}_i^d$ are the time-averaged of the UL and DL scheduling variables, respectively, $\mathfrak{m}_{i,k}^u$, $\mathfrak{m}_{i,k}^d$, $\mathfrak{m}_{i,k}$ are the lower-bound stability of the uplink, downlink, and coupling transmission indicator variables of a control system i at time k in (27), (28), and (29) respectively. $\bar{\mathcal{G}}_{lb}$ is the time-averaged of the lower-bound function $\mathcal{G}_{lb}$ that is defined as $\mathcal{G}_{lb}(.) = \max[\min(., 1), 0]$ to ensure the feasibility of the scheduling constraints. The transmission indicator variables in (27)-(29) are only written as a function of the scheduling variables since the SNR indicator function is satisfied in (26c). The stochastic problem $\mathcal{P}2$ is a mixed-integer non-convex problem where the source of stochasticity is due to the observed channel and predicted state at each time k. Moreover, the scheduling decision constraint in (26d) not only depends on its own decision but on all others scheduling decisions. Hence, to obtain the optimal scheduling decisions and optimal transmission power variables of problem $\mathcal{P}2$ in an interpretable closed-form, the Lyapunov optimization framework is utilized rather than the reinforcement learning (RL) approach.

IV. DYNAMIC CONTROL ALGORITHM USING LYAPUNOV OPTIMIZATION

In this section, we propose a dynamic control algorithm using the stochastic Lyapunov optimization framework to solve problem $\mathcal{P}2$. However, the problem involves minimizing a weighted sum of non-decreasing concave functions of the time-averaged AoI and transmission power. Based on the dynamic stochastic optimization theory [36], it can be transformed into an equivalent problem that involves minimizing a time-averaged cost function of instantaneous AoI and transmission power. This transformation is achieved through the use of non-negative auxiliary variables $\gamma_{i,k}^{\beta^l}$ and $\gamma_{i,k}^{P^l}$ and corresponding virtual queues $Q_{i,k}^{\beta^l}$ and $Q_{i,k}^{P^l}$ with queue dynamics as

$$Q_{i,k+1}^{\beta^l} = \max\left\{Q_{i,k}^{\beta^l} - \gamma_{i,k}^{\beta^l}, 0\right\} + \beta_{i,k}^l, \tag{31a}$$

$$Q_{i,k+1}^{P^l} = \max\left\{Q_{i,k}^{P^l} - \gamma_{i,k}^{P^l}, 0\right\} + \hat{P}_{i,k}^l, \tag{31b}$$

$\forall l \in \{u, d\}, i \in \mathcal{M}, k$, where $\hat{P}_{i,k}^l$ will be optimized at each time k. Then, the transformed problem is given as

$$(\mathcal{P}3) \quad \underset{\mathbf{a}_k^l, \mathbf{P}_k^l, \mathbf{r}^{\beta^l}, \mathbf{r}^{P^l}}{\text{Minimize}} \; \overline{\mathcal{C}\left(\{\gamma_{i,k}^{\beta^l}\}, \{\gamma_{i,k}^{P^l}\}\right)} \tag{32a}$$

$$\text{subject to:} \quad \bar{\beta}_i^l \leq \bar{\gamma}_i^{\beta^l}, \tag{32b}$$

$$\hat{\bar{P}}_i^l \leq \bar{\gamma}_i^{P^l}, \tag{32c}$$

$$1 \leq \gamma_{i,k}^{\beta^l} \leq B_{max}, \tag{32d}$$

$$0 \leq \gamma_{i,k}^{P^f} \leq P_{max}^l, \tag{32e}$$

$$(26b) - (26e), \; (30b) - (30d),$$

$\forall l \in \{u, d\}, i \in \mathcal{M}$, where $\mathbf{r}_k^{\beta^l} = \{\gamma_{i,k}^{\beta^l} : l \in \{u, d\}, i \in \mathcal{M}\}$ and $\mathbf{r}_k^{P^l} = \{\gamma_{i,k}^{P^l} : l \in \{u, d\}, i \in \mathcal{M}\}$ are the vectors of the introduced auxiliary variables. The constraints in (32d) and (32e) are introduced to bound the auxiliary variables. These constraints can be satisfied by ensuring the stability of their virtual queues since the lower bound of these constraints can be viewed as the arrival rate of their virtual queues, while the upper bound can be viewed as the service rate of such virtual queues [36]. Following [36], the problem $\mathcal{P}2$ and the transformed problem $\mathcal{P}3$ are equivalent in which the optimal solution of $\mathcal{P}3$ can be directly turned into an optimal solution of $\mathcal{P}2$.

To handle UL and DL scheduling variables constraints in (30b)−(30d) associated with the control stability constraint in (25), the virtual queues $Q_{i,k}^{C^l}$ and $Q_{i,k}^{C}$ are introduced for all control systems whose dynamics are

$$Q_{i,k+1}^{C^l} = \max\{Q_{i,k}^{C^l} - \alpha_{i,k}^l, 0\} + \mathcal{G}_{lb}(m_{i,k}^l), \tag{33a}$$

$$Q_{i,k+1}^{C} = \max\{Q_{i,k}^{C} - \alpha_{i,k}^u \alpha_{i,k}^d, 0\} + \mathcal{G}_{lb}(m_{i,k}), \tag{33b}$$

$\forall l \in \{u, d\}, i \in \mathcal{M}, k$, where $\alpha_{i,k}^l$ will be optimized at each time k. The constraints in (30b)−(30d) can be satisfied, if their virtual queues are mean-rate stable, i.e., their time-averaged arrival rate is not larger than its time-averaged service rate [36]. At this point, the dynamic stochastic optimization is applied to solve the transformed problem $\mathcal{P}3$, which minimizes a weighted sum of the time-averaged cost function of instantaneous AoI and transmission power subject to the virtual queues stability constraints and the original problem constraints in (26b)−(26e). In this regard, we define $\mathbf{Q}_k^{\beta^l}$, $\mathbf{Q}_k^{P^l}$, $\mathbf{Q}_k^{C}$, and $\mathbf{Q}_k^{C^l}$ as a vector of all virtual queues $Q_{i,k}^{\beta^l}$, $Q_{i,k}^{P^l}$, $Q_{i,k}^{C}$, and $Q_{i,k}^{C^l}$ for all control systems, respectively. We denote the combined queue vector of all virtual queues at time k by $\mathcal{X}_k = \left[\mathbf{Q}_k^{\beta^l}, \mathbf{Q}_k^{P^l}, \mathbf{Q}_k^{C}, \mathbf{Q}_k^{C^l}\right]$, and express the conditional Lyapunov drift-plus-penalty as

$$\Delta(\mathcal{X}_k)$$
$$= \mathbb{E}\left[\mathcal{L}(\mathcal{X}_{k+1}) - \mathcal{L}(\mathcal{X}_k) + V\mathcal{C}\left(\{\gamma_{i,k}^{\beta^l}\}, \{\gamma_{i,k}^{P^l}\}\right) \Big| \mathcal{X}_k\right], \tag{34}$$

where $\mathcal{L}(\mathcal{X}_k)$ is the quadratic Lyapunov function of $\mathcal{X}_k$ that measures the virtual queues congestion in a scalar metric and is defined as $\mathcal{L}(\mathcal{X}_k) = \frac{1}{2}\sum_{i=1}^{M}\left[(Q_{i,k}^{\beta^l})^2 + (Q_{i,k}^{P^l})^2 + (Q_{i,k}^{C})^2 + (Q_{i,k}^{C^l})^2\right]$. $V \geq 0$ controls the trade-off between minimizing

the objective function and stabilizing the virtual queues. Subsequently, plugging the inequalities $(\max[a-b,0]+c)^2 \leq a^2+b^2+c^2-2\,a\,(b-c)$, $\forall a,b,c \geq 0$, $(\max(a,0))^2 \leq a^2$, and all virtual queue dynamics into (34), we derive

$$(34) \leq B + \mathbb{E}\Big[\sum_{i=1}^{M}\Big(V\omega_{\beta_l}\mathcal{G}_\beta\left(\gamma_{i,k}^{\beta^l}\right) - Q_{i,k}^{\beta^l}\gamma_{i,k}^{\beta^l}\Big)\,|\mathcal{X}_k\Big]$$
$$+\mathbb{E}\Big[\sum_{i=1}^{M}\Big(V\omega_{P_l}\mathcal{G}_P\left(\gamma_{i,k}^{P^l}\right) - Q_{i,k}^{P^l}\gamma_{i,k}^{P^l}\Big)\,|\mathcal{X}_k\Big]$$
$$+\mathbb{E}\Big[\sum_{i=1}^{M}Q_{i,k}^{\beta^l}\beta_{i,k}^l|\mathcal{X}_k\Big] + \mathbb{E}\Big[\sum_{i=1}^{M}Q_{i,k}^{P^l}\hat{P}_{i,k}^l|\mathcal{X}_k\Big]$$
$$-\mathbb{E}\Big[\sum_{i=1}^{M}Q_{i,k}^{C}\left(\alpha_{i,k}^u\alpha_{i,k}^d - \mathcal{G}_{lb}(m_{i,k})\right)|\mathcal{X}_k\Big]$$
$$-\mathbb{E}\Big[\sum_{i=1}^{M}Q_{i,k}^{C^l}\left(\alpha_{i,k}^l - \mathcal{G}_{lb}(m_{i,k}^l)\right)|\mathcal{X}_k\Big]. \quad (35)$$

The constant B details in (35) are omitted since it does not affect the system performance in the Lyapunov optimization. A solution to $\mathcal{P}3$ can be obtained by minimizing the upper-bound (35) at each time as

$$(\mathcal{P}4) \; \underset{\mathbf{a}^l, \mathbf{P}^l, \mathbf{r}^{\beta^l}, \mathbf{r}^{P^l}}{\text{Minimize}} \sum_{i=1}^{M}\Bigg[\Big(V\omega_{\beta_l}\mathcal{G}_\beta\left(\gamma_{i,k}^{\beta^l}\right) - Q_{i,k}^{\beta^l}\gamma_{i,k}^{\beta^l}\Big)$$
$$+ \Big(V\omega_{P_l}\mathcal{G}_P\left(\gamma_{i,k}^{P^l}\right) - Q_{i,k}^{P^l}\gamma_{i,k}^{P^l}\Big)$$
$$+ Q_{i,k}^{P^l}\hat{P}_{i,k}^l - Q_{i,k}^{C^l}\left(\alpha_{i,k}^l - \mathcal{G}_{lb}(m_{i,k}^l)\right)$$
$$+ Q_{i,k}^{\beta^l}\beta_{i,k}^l - Q_{i,k}^{C}\left(\alpha_{i,k}^u\alpha_{i,k}^d - \mathcal{G}_{lb}(m_{i,k})\right)\Bigg] (36a)$$
$$\text{subject to:} \quad (26b) - (26e) \text{ and } (32d) - (32e).$$

The optimality of problem $\mathcal{P}4$ is asymptotically approached by increasing V [36]. The problem $\mathcal{P}4$ is of separable structure, which motivates us to determine the AoI auxiliary vector $\mathbf{r}^{\beta^l}$, transmission power auxiliary vector $\mathbf{r}^{P^l}$, scheduling vector $\mathbf{a}^l$, and transmission power vector $\mathbf{P}^l$ in an alternative optimization form. Hence, the overall minimization problem $\mathcal{P}4$ can be decomposed into two separate sub-problems that can be solved concurrently with the observation of the virtual queues, control, and channel states.

1) Auxiliary Variable Sub-Problems: The first decomposed sub-problem is the AoI auxiliary sub-problem, while the second decomposed sub-problem is the transmission power auxiliary sub-problem. Since the auxiliary variables of such problems are separated and independent among different control systems, their minimization sub-problems can be decoupled to be computed for each control system separately as the following convex problems

$$(\mathcal{P}4.1)\underset{\gamma_{i,k}^{\beta^l}}{\text{Minimize}} \; V\omega_{\beta_l}\mathcal{G}_\beta\left(\gamma_{i,k}^{\beta^l}\right) - Q_{i,k}^{\beta^l}\gamma_{i,k}^{\beta^l} \quad (37a)$$
$$\text{subject to:} \; 1 \leq \gamma_{i,k}^{\beta^l} \leq B_{max}, \quad (37b)$$
$$(\mathcal{P}4.2)\underset{\gamma_{i,k}^{P^l}}{\text{Minimize}} \; V\omega_{P_l}\mathcal{G}_P\left(\gamma_{i,k}^{P^l}\right) - Q_{i,k}^{P^l}\gamma_{i,k}^{P^l} \quad (38a)$$
$$\text{subject to:} \; 0 \leq \gamma_{i,k}^{P^l} \leq P_{max}^l. \quad (38b)$$

The optimal auxiliary variables are obtained by differentiating the objective functions of these problems. Let $\mathcal{A}(\gamma^{\beta^l}) = V\omega_{\beta_l}\log(1+\gamma^{\beta^l}) - Q_{i,k}^{\beta^l}\gamma_{i,k}^{\beta^l}$ and $\gamma_{i,k}^{\beta^{l*}}$ denotes the solution of $\mathcal{A}(\gamma^{\beta^l})$ as $\acute{\mathcal{A}}(\gamma^{\beta^l}) = \frac{V\omega_{\beta_l}}{\left(1+\gamma_{i,k}^{\beta^l}\right)} - Q_{i,k}^{\beta^l} = 0$, the optimal AoI auxiliary variable of $\mathcal{P}4.1$ is given as

$$\gamma_{i,k}^{\beta^{l*}} = \min\left\{\max\left\{\frac{V\omega_{\beta_l} - Q_{i,k}^{\beta^l}}{Q_{i,k}^{\beta^l}}, 1\right\}, B_{max}\right\}, \quad (39)$$

$\forall l \in \{u,d\}, i \in \mathcal{M}, k$. Similarly, by letting $\mathcal{A}(\gamma^{P^l}) = V\omega_{P_l}\log(1+\gamma^{P^l}) - Q_{i,k}^{P^l}\gamma_{i,k}^{P^l}$ and $\gamma_{i,k}^{P^{l*}}$ denotes the solution of $\mathcal{A}(\gamma^{P^l})$ as $\acute{\mathcal{A}}(\gamma^{P^l}) = \frac{V\omega_{P_l}}{\left(1+\gamma_{i,k}^{P^l}\right)} - Q_{i,k}^{P^l} = 0$, the optimal transmission power auxiliary variable of $\mathcal{P}4.2$ is given as

$$\gamma_{i,k}^{P^{l*}} = \min\left\{\max\left\{\frac{V\omega_{P_l} - Q_{i,k}^{P^l}}{Q_{i,k}^{P^l}}, 0\right\}, P_{max}^l\right\}, \quad (40)$$

$\forall l \in \{u,d\}, i \in \mathcal{M}, k$.

2) Scheduling Decision and Transmission Power Sub-Problems: The optimal UL-DL scheduling variables and the optimal UL-DL transmission power variables are obtained by minimizing the remaining terms of the objective function of problem $\mathcal{P}4$ at each time subject to the scheduling and transmission power constraints in (26b)−(26e), which is given

$$(\mathcal{P}4.3)\underset{\mathbf{a}_{i,k}^l, \mathbf{P}_{i,k}^l}{\text{Minimize}} \sum_{i=1}^{M} Q_{i,k}^{\beta^l}\beta_{i,k}^l + Q_{i,k}^{P^l}\hat{P}_{i,k}^l$$
$$-Q_{i,k}^{C^l}\left[\alpha_{i,k}^l - \mathcal{G}_{lb}(m_{i,k}^l)\right]$$
$$-Q_{i,k}^{C}\left[\alpha_{i,k}^u\alpha_{i,k}^d - \mathcal{G}_{lb}(m_{i,k})\right] \quad (41a)$$
$$\text{subject to:} \quad (26b) - (26e),$$

which is a mixed-integer non-convex problem. Due to the complexity of the exhaustive search for finding the optimal solution, we propose a low-complexity two-stage sequential optimization strategy to find a sub-optimal solution to the joint power allocation and scheduling assignment problem. This strategy firstly obtains the UL and DL transmission power, followed by the UL and DL scheduling variables. The optimal UL and DL transmission power for each control system, determined by solving the following power allocation problem

$$(\mathcal{P}4.4)\underset{\mathbf{P}_k^u, \mathbf{P}_k^d}{\text{Minimize}} \sum_{i=1}^{M}\alpha_{i,k}^u\left[\mathcal{Q}_{i,k}^{S_1} + Q_{i,k}^{P^u}P_{i,k}^u\right] + \mathcal{Q}_{i,k}^{l_1} \; (42a)$$
$$+ \alpha_{i,k}^d\left[\mathcal{Q}_{i,k}^{C_1} + Q_{i,k}^{P^d}P_{i,k}^d\right] - \alpha_{i,k}^u\alpha_{i,k}^d Q_{i,k}^C,$$
$$\text{subject to:} \; \frac{\text{SNR}_{th}^u N_0\,\omega}{\|\mathbf{H}_{i,k}^u\|^2} \leq P_{i,k}^u \leq P_{max}^u,$$
$$(42b)$$
$$\frac{\text{SNR}_{th}^d N_0\,\omega}{\|\mathbf{H}_{i,k}^d\|^2} \leq P_{i,k}^d \leq P_{max}^d, \quad (42c)$$

$\forall i \in \mathcal{M}, k$, where $\mathcal{Q}_{i,k}^{S_1} = -Q_{i,k}^{\beta^u}\beta_{i,k-1}^u - Q_{i,k}^{C^u}$, $\mathcal{Q}_{i,k}^{C_1} = -Q_{i,k}^{\beta^d}\beta_{i,k-1}^d - Q_{i,k}^{C^d}$, and $\mathcal{Q}_{i,k}^{l_1} = Q_{i,k}^{C^u}\mathcal{G}_{lb}(m_{i,k}^u) + Q_{i,k}^{C^d}\mathcal{G}_{lb}(m_{i,k}^d) + Q_{i,k}^{C}\mathcal{G}_{lb}(m_{i,k}) + Q_{i,k}^{\beta^u}\left(1+\beta_{i,k-1}^u\right) +$

$Q_{i,k}^{\beta^d} \left(1 + \beta_{i,k-1}^d \right)$ are the constant terms defined in the objective function of $\mathcal{P}4.4$. The above problem $\mathcal{P}4.4$ is a generalized min-weight problem, which can be decoupled into a series of independent sub-problems for each control system separately. Hence, the optimal UL-DL transmission power variables are given as

$$P_{i,k}^{l^*} = \begin{cases} \dfrac{\text{SNR}_{th}^l N_0 \, \omega}{\|\mathbf{H}_{i,k}^l\|^2}, & \text{if } Q_{i,k}^{P^l} \geq 0 \\ P_{max}^l, & \text{if } Q_{i,k}^{P^l} < 0. \end{cases} \tag{43}$$

Given the optimal UL and DL transmission power variables in (43), the optimal UL and DL scheduling variables for each control system that has a control state/action to transmit are obtained by solving the following scheduling assignment problem

$$(\mathcal{P}4.5) \operatorname*{Minimize}_{\mathbf{a}_k^u, \mathbf{a}_k^d} \sum_{i=1}^{M} \alpha_{i,k}^u \left[\mathcal{Q}_{i,k}^{S_1} + Q_{i,k}^{P^u} P_{i,k}^{u^*} \right] + \mathcal{Q}_{i,k}^{l_1}$$

$$+ \alpha_{i,k}^d \left[\mathcal{Q}_{i,k}^{C_1} + Q_{i,k}^{P^d} P_{i,k}^{d^*} \right] - \alpha_{i,k}^u \alpha_{i,k}^d Q_{i,k}^C, \tag{44a}$$

$$\text{subject to: } (26d) - (26e).$$

The optimal UL and DL scheduling variables are obtained as follows:

$$\alpha_{i,k}^{u^*} \& \alpha_{i,k}^{d^*}$$

$$= \begin{cases} \alpha_{j_1,k}^u = 1, \ \alpha_{j_2,k}^d = 1, & \text{if } \mathbb{Q}_{j_1,k}^1 + \mathbb{Q}_{j_2,k}^2 < \mathbb{Q}_{j_3,k} \\ \alpha_{j_3,k}^u = 1, \ \alpha_{j_3,k}^d = 1, & \text{if } \mathbb{Q}_{j_1,k}^1 + \mathbb{Q}_{j_2,k}^2 > \mathbb{Q}_{j_3,k} \\ \alpha_{j,k}^u = 0, \ \alpha_{j,k}^d = 0, & \forall j \notin \{j_1 \& j_2 || j_3 \}, \end{cases} \tag{45}$$

where $\mathbb{Q}_{i,k}^1 = \mathcal{Q}_{i,k}^{S_1} + Q_{i,k}^{P^u} P_{i,k}^{u^*}$, $\mathbb{Q}_{i,k}^2 = \mathcal{Q}_{i,k}^{C_1} + Q_{i,k}^{P^d} P_{i,k}^{d^*}$, $\mathbb{Q}_{i,k}^3 = -Q_{i,k}^C$, and $\mathbb{Q}_{j_3,k} = \mathbb{Q}_{j_3,k}^1 + \mathbb{Q}_{j_3,k}^2 + \mathbb{Q}_{j_3,k}^3$ are the terms defined in the objective function of problem $\mathcal{P}4.5$. Moreover, $j_1 = \arg\min_{i \in \mathcal{M}} \mathbb{Q}_{i,k}^1$, $j_2 = \arg\min_{i \in \mathcal{M}} \mathbb{Q}_{i,k}^2$, and $j_3 = \arg\min_{i \in \mathcal{M}} (\mathbb{Q}_{i,k}^1 + \mathbb{Q}_{i,k}^2 + \mathbb{Q}_{i,k}^3)$ are control system indices. Finally, the procedures of the proposed stability-aware scheduling approach to obtain the optimal solution of problem $\mathcal{P}3$ is outlined in Algorithm 1.

V. SIMULATION RESULTS AND DISCUSSIONS

In this section, the performance of the proposed stability-aware scheduling algorithm is investigated in an inverted-pendulum on a cart system with $M = 2$, and $M = 20$ inverted-pendulums, respectively. Each inverted-pendulum system is described by a four-dimensional state vector as $\mathbf{x}_{i,k}^u = \left[x_{i,k}, \dot{x}_{i,k}, \theta_{i,k}, \dot{\theta}_{i,k} \right]$, where $x_{i,k}$ represents the cart's position along the horizontal axis, $\dot{x}_{i,k}$ represents the cart's velocity, $\theta_{i,k}$ represents the pendulum angle along the vertical axis, and $\dot{\theta}_{i,k}$ represents the pendulum's angular velocity. The initial state of the control systems i is $\mathbf{x}_{i,0}^u = [0 \quad 0 \quad 0.1 \quad 0]^T$. The action $\mathbf{u}_{i,k}^a$ is the horizontal force applied on the linear cart. By applying a zeroth-order with a state sampling of 10ms on the continuous dynamics of the inverted-pendulum system and linearizing around the pendulum up-position, i.e., $\theta_{i,k} = 0$,

we obtain the following discrete-time linear dynamics matrices [22],

$$\mathbf{A}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2.055 & -0.722 & 4.828 \\ 0 & 0.023 & 0.91 & 0.037 \\ 0 & 0.677 & -0.453 & 2.055 \end{bmatrix}, \mathbf{B}_i = \begin{bmatrix} 0.034 \\ 0.168 \\ 0.019 \\ 0.105 \end{bmatrix}, \tag{46}$$

Algorithm 1 Stability-Aware Scheduling Algorithm

Initialize V, ω_{β_u}, ω_{β_d}, ω_{P_u}, ω_{P_d} and $\mathcal{X}_{i,0} = \mathbf{0}$, $\forall i \in \mathcal{M}$;
while $k \in [0, K]$ **do**
  for $i = 1$ to M **do**
    Observe the UL and DL channel states $H_{i,k}^u$ and $H_{i,k}^d$;
    Obtain the predicted state at the controller $\hat{\mathbf{x}}_{i,k}^u$ in (13);
    Obtain the predicted action at the controller $\hat{\mathbf{u}}_{i,k}^{d_c}$ in (13);
    Obtain $\gamma_{i,k}^{\beta^u}$, $\gamma_{i,k}^{\beta^d}$, $\gamma_{i,k}^{P^u}$, and $\gamma_{i,k}^{P^d}$ by solving (39) and (40);
    Obtain UL and DL scheduling $\alpha_{i,k}^u$ and $\alpha_{i,k}^d$ in (45);
    Obtain the UL and DL transmission power $P_{i,k}^l$ in (43);
    if $\alpha_{n,k}^u = 1$ & $\alpha_{n,k}^d = 1$ **then**
      Control system n-th transmits both its state and action;
      $\forall i \left((i \in \mathcal{M} \wedge i \neq n) \longrightarrow \mathbf{x}_{i,k}^u = \hat{\mathbf{x}}_{i,k}^u \ \& \ \mathbf{u}_{i,k}^a = \hat{\mathbf{u}}_{i,k}^d \right);$
    else if $\alpha_{n,k}^u = 1$ & $\alpha_{m,k}^d = 1$ **then**
      Control system n-th only transmits its state;
      Control system m-th only transmits its action;
      $\forall i \left((i \in \mathcal{M} \wedge i \neq n) \longrightarrow \mathbf{x}_{i,k}^u = \hat{\mathbf{x}}_{i,k}^u \right);$
      $\forall j \left((j \in \mathcal{M} \wedge j \neq m) \longrightarrow \mathbf{u}_{j,k}^a = \hat{\mathbf{u}}_{j,k}^d \right)$
    else
      $\forall i \in \mathcal{M} \longrightarrow \mathbf{x}_{i,k}^u = \hat{\mathbf{x}}_{i,k}^u \ \& \ \mathbf{u}_{i,k}^a = \hat{\mathbf{u}}_{i,k}^d;$
    end if
    Update the virtual queues vector $\mathcal{X}_{i,k+1}$ according to (31a)−(31b) and (33a)−(33b)
  end for
end while

Since $\mathbf{A}_i$'s largest eigenvalues $\{3.85, 0.42, 0.92, 1.00\}$ is greater than unity, the inverted-pendulum is unstable without an appropriate control action [33]. To stabilize the control system, the feedback gain matrix $\mathbf{\Phi}_i$ is calculated at the controller based on LQR in (16). The rest of the simulation parameters are $P_{max}^l = 20$ dBm/Hz, $N_0 = -20$ dBm, $\zeta_i = 0.01$, $\omega = 100$kHz, $V = 1000$, $\omega_\beta = 1$, $\omega_P = 1$, $h_k = 1$, $h_q = 1$, $\mu = 1$, and $\sigma_n^2 = 0.01$.

The performance of the proposed stability-aware scheduling method is compared versus five scheduling baselines. In *Baseline* 1. (**Round-Robin Scheduling**), each sensor/controller periodically transmits its state/action over a wireless channel with fixed transmission power and a predefined repeating order [8], [9]. In *Baseline* 2. (**Opportunistic Scheduling**), the sensor/controller is scheduled under favorable channel conditions. Otherwise, the controller/actuator applies the last received state/action [10], [11]. In *Baseline* 3, (**Event-triggered Scheduling without FDMA**), one control system with the largest state discrepancy, i.e., the difference between the current predicted state using Kalman filtering and the previous received/predicted state is larger than a predefined threshold, is scheduled at each time to transmit the state/action with fixed transmission power [12], [13]. In *Baseline* 4. (**Event-triggered Scheduling with FDMA**), each sensor/controller transmits its state/action with fixed transmission power based on its stability
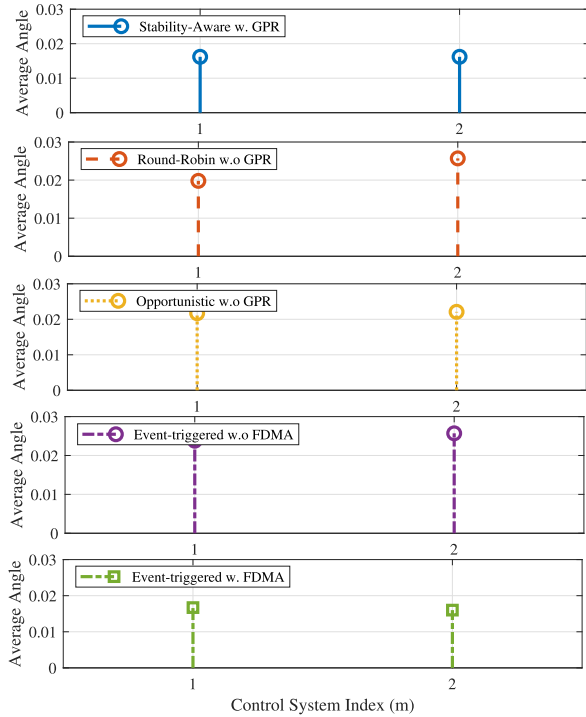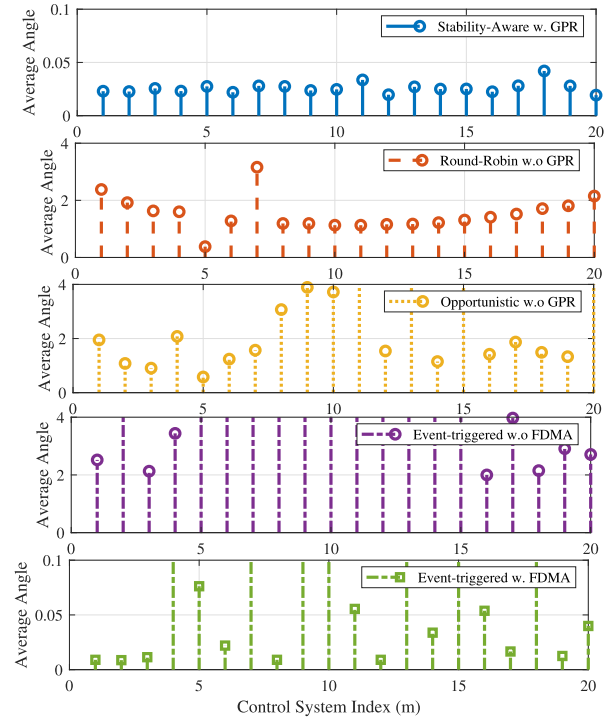
(a) Average Pendulum Angle with $M = 2$.

(b) Average Pendulum Angle with $M = 20$.

Fig. 3. Average pendulum angle to the vertical center, i.e., average control error with $M = 2$ and $M = 20$ pendulum angle systems using the proposed stability-aware, round-robin, opportunistic, event-triggered with and without FDMA.

condition, i.e., the difference between the current and previous states is greater than a predefined threshold, using FDMA [37]. In *Baseline* 5. (**Ideal Control Scheduling**), all control systems simultaneously transmit their states/actions with ultra-low latency and high reliability over perfect channels [9]. Results are collected over ten independent simulation runs, and each simulation is run for $K = 90$ discrete-time steps.

A. Average Control Error vs. Control Systems

Fig. 3 illustrates the average pendulum angle to the vertical center of each control system, i.e., the average control error of each control system, during 90 control time steps. As shown in Fig. 3(a), the proposed and baseline scheduling methods, assuming a low number of control systems ($M = 2$), can keep all the pendulums upright. Moreover, the proposed stability-Aware with GPR and the event-triggered with FDMA keep both pendulums close to zero ensuring both pendulums have the same control performance. This is because the proposed solution adapts to both channel and control states, i.e., the control system is scheduled if it has a favorable channel condition and an unstable control state needs to be stabilized. Meanwhile, the event-triggered with FDMA has approximately the same performance at the cost of wasting wireless communication resources by transmitting with fixed transmission power and high communication rate. The opportunistic scheduler without GPR has better control performance compared to the event-triggered without FDMA and the round-robin methods leveraging channel state in scheduling compared to the scheduling baselines. This in turn reflects the connection between the state estimation stability and control stability.

Fig. 3(b) plots the average control error of each control system for the proposed and baseline scheduling methods. In large numbers of control systems, the proposed stability-aware with GPR and the event-triggered with FDMA scheduling methods keep all pendulum upright, unlike the baselines. Unlike the scheduling baselines except the event-triggered with FDMA scheduling wherein at most one control system is scheduled each time due to the limited bandwidth, our proposed scheduling allows all control systems to operate simultaneously even without receiving either the current state or action, highlighting the effectiveness of GPRs at the controller and actuator thereby improving communication efficiency and control stability. Moreover, the proposed scheme applies the AoI-Aware scheduling to maintain the GPR prediction credibility and hence achieving control stability. Meanwhile, the event-triggered with FDMA scheduling keeps some pendulums upright at the cost of frequent transmissions by equally dividing the available bandwidth between the control systems, such that each control system receives a fixed fraction f_i of the total capacity $f_i = \omega/M$, affecting transmission latency.

B. Communication Rate vs. Number of Control Systems

Fig. 4 presents a histogram of the achieved communication rates for the sensing and actuating links for $M = 20$ during 90 control time steps. The sensing/actuating communication rate is defined as the number of times the sensing/actuating link of a control system is scheduled divided by the time interval as n_l/K. The proposed method achieves sensing communication rates concentrated in the range from 0 to

(a) Communication rate of sensing link.
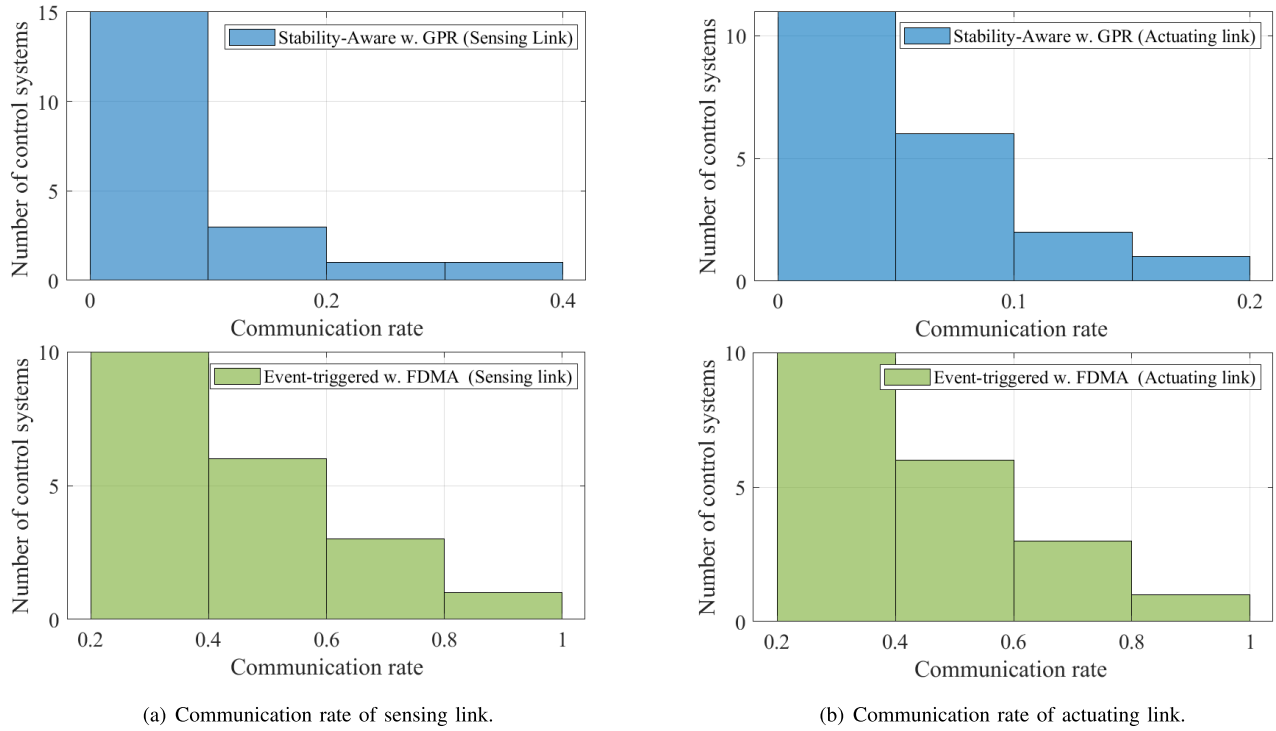
(b) Communication rate of actuating link.

Fig. 4.   Histogram of achieved communication rate in a large number of control system ($M = 20$) of the proposed stability-aware and event-triggered with FDMA throughout 90 control time steps.

0.4 and actuating communication rates ranging from 0 to 0.2. On the other hand, the event-triggered with FDMA scheduling achieves wide sensing and actuating communication rates ranging from 0 to 1. The reason behind this result is that the stability-aware with GPR scheduling is adapting to both the channel and control states, and GPRs at the controller and actuator sides compensate for the missing received observations, hence improving the communication efficiency. Meanwhile, each control system, in the event-triggered with FDMA scheduling, only transmits its control state/action based on its control stability condition without taking into account channel states that result in increasing state/action estimation uncertainty from the adverse channel states, and in turn, affecting control stability. Hence, each control system in the event-triggered with FDMA scheduling requires frequent transmissions to ensure control stability by applying appropriate action based on low communication uncertainty. Note that the range of the sensing communication rate is larger than that of the actuating communication rate since some control systems fail to transmit their states at the beginning affecting prediction credibility. However, the number of control systems that require frequent scheduling in the actuating link is larger than that of the sensing link which stems from the fast dynamics related to the inverted-pendulum system that requires quick appropriate action.

C. State Trajectory and Controller/Actuator AoI vs. Time

To dive deeper into the benefits of the proposed stability-aware with GPR scheduling, we present in Fig. 5 the state trajectory, controller AoI, and actuator AoI of a randomly chosen control system in low and large number control systems

regimes. In Fig 5(a), the state trajectory of the proposed stability-aware with GPR scheduling and event-triggered with FDMA scheduling, in a low number of control systems, are extremely close to that of the ideal control system where their pendulums remain upright over time. Meanwhile, the state trajectory of the opportunistic without GPR scheduling is slightly better than that of the event-triggered without FDMA scheduling by keeping the pendulum upright over time due to the scheduled control system with a favorable channel state. Finally, the state trajectory of the round-robin without GPR scheduling slightly matches the desired state at the cost of a high communication rate and fixed transmission power.

As shown in Fig. 5(a), the controller and actuator AoI of the proposed stability-aware scheduling equals one, i.e., the sensing and actuating links of a control system are synchronously scheduled, until 15 control time steps. This is to guarantee that the received actions to action GPR has a low state/action estimation uncertainty that affects the GPR state/action prediction credibility and control stability. Then, the actuator AoI starts increasing compared to the controller AoI that remains at value one until 45 control time steps, i.e., the sensing link of a control system is scheduled until 45 control time steps while the actuating link is not scheduled. The rationale behind this result is to schedule either the sensing link or/and actuating link of a control system with favorable channel conditions to ensure the state/action prediction credibility. Hence, the transmitted action, when the sensing link of a control system is not scheduled, depends on a credible predicted state. This shows the impact of the decoupled scheduling between the UL and DL communications compared to the coupled scheduling in terms of improving communication efficiency by reducing sensing and actuating communication rates while
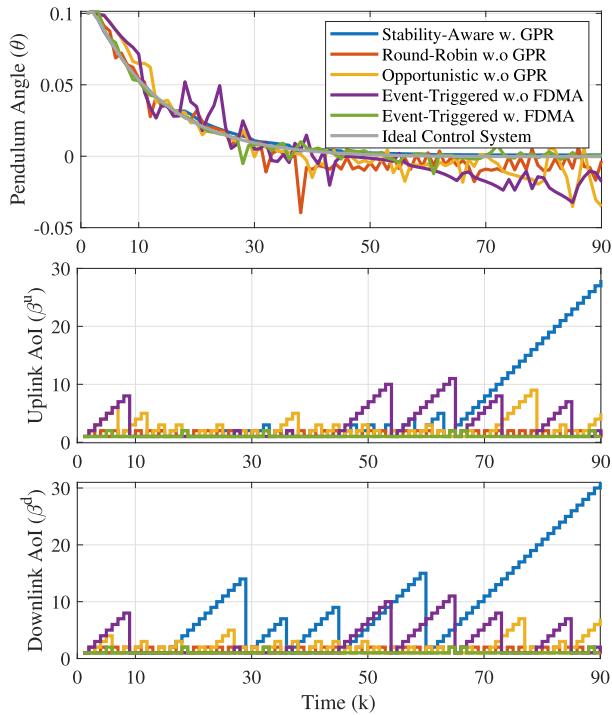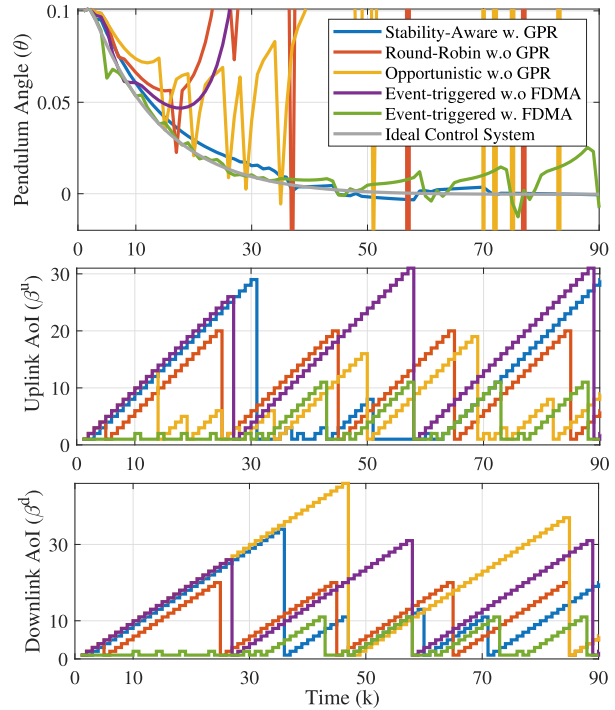
(a) Single control system out of $M = 2$.

(b) Single control system out of $M = 20$.

Fig. 5. Comparison of the state trajectory, controller AoI, and actuator AoI of a randomly chosen control system between the proposed stability-aware, round-robin, Opportunistic, event-triggered with and without FDMA, and ideal control system.

guaranteeing control stability. For large numbers of control systems shown in Fig. 5(b), the control error of the stability-aware scheduling and event-triggered with FDMA scheduling exponentially decay over time compared to the scheduling baselines. Meanwhile, the other scheduling baselines are exponentially growing over time due to the accumulated control error in the absence of appropriate action.

D. Digital Transmissions vs. Analog Transmissions

Thus far, we have considered analog uncoded communication (AUC). To demonstrate its effectiveness, Fig. 6 compares AUC and its digital coded communication (DCC) counterpart, in terms of the average pendulum angle, i.e., average control error, of a single control system for different values of SNR_{th}. Here, we assume the DL channels are ideal and focus only on the UL with the 180 kHz channel bandwidth of NB-IoT systems [38]. Following [39], we consider that the signaling rate of the communication system is 10x higher than the sampling rate of the control system, with the following reasoning. To convey 12 symbols, it requires 12 orthogonal sub-carriers that correspond to 1ms in NB-IoT systems [38]. By contrast, to convey the 256 bits of BCH(7, 4) coded control states using the 16-QAM modulation, it requires transmitting 112 modulated symbols that result in around a 10ms delay. Given this setting, for high SNR_{th} ($\text{SNR}_{th} > 40$), Fig. 6 shows that AUC achieves the lowest average pendulum angle that can be achieved by DCC using the BCH(7, 4) code with 256-QAM modulation. The rationale behind this result is that DCC requires a high modulation order with a low coding
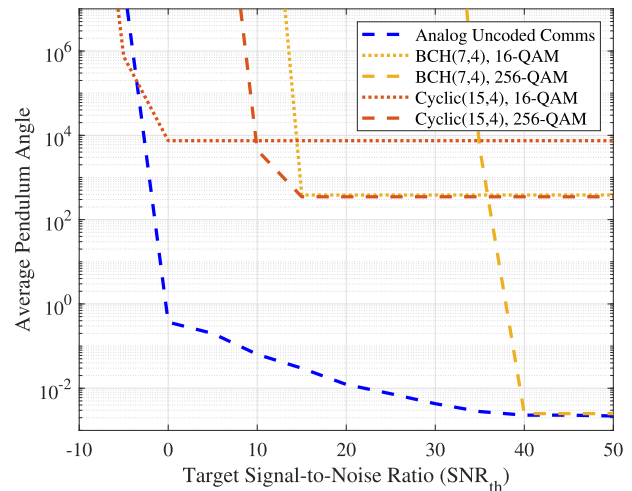

Fig. 6. Average control error of a single control system with analog uncoded and digital communications vs. different SNR_{th}.

rate to achieve low transmission latency for a given SNR_{th}, in contrast to AUC whose transmission latency depends only on the bandwidth. In the practical range of SNR_{th} ($0 < \text{SNR}_{th} < 40$), AUC achieves the lower average pendulum angle for all DCC counterparts. The rationale can be explained as follows. First, DCC has higher transmission latency than the sampling period of the inverted pendulum system, affecting the control stability. Second, DCC with BCH(7, 4) and 256-QAM achieves the same transmission latency as that of AUC, yet it requires either a high SNR_{th} to achieve the bit error rate (BER) of a designed point or a low coding rate to achieve the BER
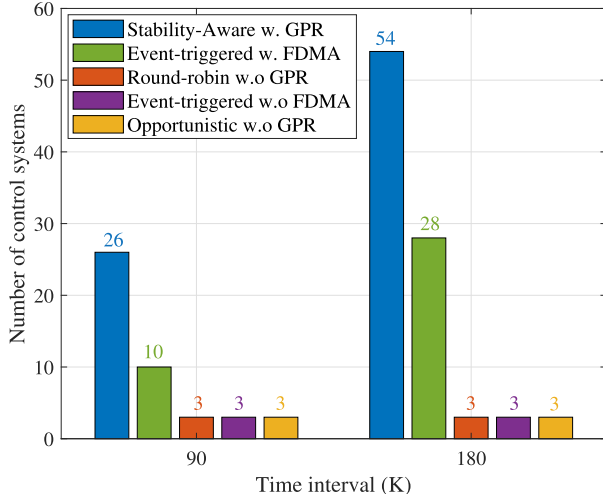
Fig. 7. Total number of supported control systems using stability-aware with GPR, event-triggered with FDMA, round-robin, event-triggered without FDMA, and opportunistic scheduling.

of the designed point at the cost of increasing the transmission latency. Otherwise, the degradation of the channel conditions in the conventional DCC leads to a system breakdown, and the BER would diverge from the designed point. For low SNR_{th} ($\mathrm{SNR}_{th} < 0$), DCC with Cyclic$(15, 4)$ and 16-QAM achieves a slightly lower average pendulum angle than AUC, which in turn concludes that DCC outperforms AUC only under low SNR_{th}.

E. Number of Served Control Systems vs. Time Interval

Fig. 7 presents the final capacity of the scheduled control systems over two different time intervals. We assume that a scheduling method successfully controls several control systems for all control systems within $|\theta_{i,k}| \leq 0.05$ error region for 10 independent simulation runs. As observed in Fig. 7, the proposed approach has a significant impact compared to the baselines in terms of supporting a large number of control systems over the time interval. The rationale behind this result is due to exploiting two GPRs at the controller and actuator sides and obtaining a sufficient number of observations in the GPR training sets. This, in turn, enhances the prediction credibility, communication efficiency, and control stability performance while supporting a large number of control systems.

VI. CONCLUSION

In this work, we proposed a GPR-based predictive WNCS architecture and its UL-DL decoupled scheduling algorithm that ensures control stability, communication reliability, and GPR prediction credibility. Extensive simulations corroborated the effectiveness of the proposed framework in supporting a larger number of inverted pendulum control systems without compromising control stability, compared to existing communication and/or control aware schedulers without GPR prediction, such as event-triggered and opportunistic schedulers. To further improve the scalability, a possible extension of this work is to incorporate advanced multiple access schemes such

as non-orthogonal and rate-splitting multiple access schemes. For unknown control system dynamics, applying distributed multi-agent reinforcement learning to multiple control systems could be another interesting topic for future research.

APPENDIX A
PROOF OF LEMMA 1

Given the predicted state $\hat{\mathbf{x}}_{i,k}^{u}$ in (13) and the state prediction error covariance matrix $\mathcal{J}_{i,k}^{u}$ in (14), it holds that the expected current Lyapunov value of the system is

$$
\begin{aligned}
&\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k}^{u})|\hat{\mathbf{x}}_{i,k}^{u}\big] \\
&= \mathbb{E}\big[\mathbf{x}_{i,k}^{u^{T}} \mathcal{Z}\, \mathbf{x}_{i,k}^{u}|\hat{\mathbf{x}}_{i,k}^{u}\big] \\
&= \mathbb{E}\big[\left(\hat{\mathbf{x}}_{i,k}^{u}-\mathbf{e}_{i,k}^{u}\right)^{T} \mathcal{Z}\left(\hat{\mathbf{x}}_{i,k}^{u}-\mathbf{e}_{i,k}^{u}\right)\big] \\
&= \mathbb{E}\big[\hat{\mathbf{x}}_{i,k}^{u^{T}} \mathcal{Z}\hat{\mathbf{x}}_{i,k}^{u} - \hat{\mathbf{x}}_{i,k}^{u^{T}} \mathcal{Z}\mathbf{e}_{i,k}^{u} - \mathbf{e}_{i,k}^{u^{T}} \mathcal{Z}\hat{\mathbf{x}}_{i,k}^{u} + \mathbf{e}_{i,k}^{u^{T}} \mathcal{Z}\mathbf{e}_{i,k}^{u}\big],
\end{aligned}
\tag{47}
$$

$\forall \mathcal{Z} \in \mathbb{S}_{++}^{D}$. In (47), the first term is a constant as the expectation is taken w.r.t. the state prediction error $\mathbf{e}_{i,k}^{u}$, while the cross-terms such as $\mathbb{E}[\hat{\mathbf{x}}_{i,k}^{u^{T}} \mathcal{Z}\mathbf{e}_{i,k}^{u}]$ and $\mathbb{E}[\mathbf{e}_{i,k}^{u^{T}} \mathcal{Z}\hat{\mathbf{x}}_{i,k}^{u}]$ can be canceled since the predicted state $\hat{\mathbf{x}}_{i,k}^{u}$ and the state prediction error $\mathbf{e}_{i,k}^{u}$ are uncorrelated. Hence, the expected current value of $\mathcal{L}(\mathbf{x}_{i,k})$ is given as

$$
\begin{aligned}
\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k}^{u})|\hat{\mathbf{x}}_{i,k}^{u}\big] &= \hat{\mathbf{x}}_{i,k}^{u^{T}} \mathcal{Z}\hat{\mathbf{x}}_{i,k}^{u} + \mathrm{Tr}\Big[\mathcal{Z}\,\mathbb{E}[\mathbf{e}_{i,k}^{u^{T}}\mathbf{e}_{i,k}^{u}]\Big] \\
&= \|\hat{\mathbf{x}}_{i,k}^{u}\|_{\mathcal{Z}^{\frac{1}{2}}}^{2} + \mathrm{Tr}\big[\mathcal{Z}\,\mathcal{J}_{i,k}^{u}\big],
\end{aligned}
\tag{48}
$$

$\forall \mathcal{Z} \in \mathbb{S}_{++}^{D}$, where the last term in (48) is obtained via the expectation of the quadratic form $\mathbb{E}[\mathbf{e}_{i,k}^{u^{T}}\mathcal{Z}\mathbf{e}_{i,k}^{u}]$ w.r.t the state prediction error, i.e, $\mathbb{E}[\mathbf{e}_{i,k}^{u^{T}}\mathcal{Z}\mathbf{e}_{i,k}^{u}]=(\mathbb{E}[\mathbf{e}_{i,k}^{u}])^{T}\mathcal{Z}(\mathbb{E}[\mathbf{e}_{i,k}^{u}])+\mathrm{Tr}[\mathcal{Z}\,\mathcal{J}_{i,k}^{u}]$. It is observed in (48) that the expected current Lyapunov value grows larger as the predicted state $\hat{\mathbf{x}}_{i,k}^{u}$ is near instability and/or the prediction error covariance matrix $\mathcal{J}_{i,k}^{u}$ is larger due to lack of sufficient observations. ∎

APPENDIX B
PROOF OF LEMMA 2

As a result of the remote sensing-loop state evolution in (19) and the open-loop state evolution in (18), the expected future value of $\mathcal{L}(\mathbf{x}_{i,k+1}^{u})$ in (25) of the UL transmission is given as

$$
\begin{aligned}
&\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^{u})|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big] \\
&= \xi_{i,k}^{u}\, \mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^{s})|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big] \\
&\quad + \left(1-\xi_{i,k}^{u}\right) \mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^{o})|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big],
\end{aligned}
\tag{49}
$$

For a given predicted state $\hat{\mathbf{x}}_{i,k}^{u}$, state prediction error covariance matrix $\mathcal{J}_{i,k}^{u}$, predicted action $\hat{\mathbf{u}}_{i,k}^{d}$, action prediction error covariance matrix $\mathcal{J}_{i,k}^{d}$, wireless UL channel $\mathbf{H}_{i,k}^{u}$, and UL transmission power $P_{i,k}^{u}$, the expected future value of $\mathcal{L}(\mathbf{x}_{i,k+1}^{u})$ of the UL transmission in (49) is

$$
\begin{aligned}
&\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^{u})|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big] \\
&\overset{(1)}{=} \xi_{i,k}^{u}\mathbb{E}\big[\mathbf{x}_{i,k+1}^{s^{T}}\mathcal{Z}\mathbf{x}_{i,k+1}^{s}|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big] \\
&\quad + \left(1-\xi_{i,k}^{u}\right)\mathbb{E}\big[\mathbf{x}_{i,k+1}^{o^{T}}\mathcal{Z}\mathbf{x}_{i,k+1}^{o}|\hat{\mathbf{x}}_{i,k}^{u}, \hat{\mathbf{u}}_{i,k}^{d}, \mathbf{H}_{i,k}^{u}, P_{i,k}^{u}\big]
\end{aligned}
$$

$$\overset{(2)}{=} \xi_{i,k}^u \mathbb{E}\Big[\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{v}_{i,k}^u - \mathbf{B}_i\mathbf{e}_{i,k}^d + \mathbf{w}_k\big)^T$$
$$\times \mathcal{Z}\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{v}_{i,k}^u - \mathbf{B}_i\mathbf{e}_{i,k}^d + \mathbf{w}_k\big)\Big]$$
$$+ (1-\xi_{i,k}^u)\mathbb{E}\Big[\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{e}_{i,k}^d$$
$$+ \mathbf{w}_k\big)^T$$
$$\times \mathcal{Z}\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{e}_{i,k}^d + \mathbf{w}_k\big)\Big]$$
$$\overset{(3)}{=} \xi_{i,k}^u\Big[\|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^{c^T}\mathcal{Z}\mathbf{A}_i^c\mathcal{J}_{i,k}^u\big] + \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big]$$
$$+ \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\big] + \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]\Big]$$
$$+ (1-\xi_{i,k}^u)\Big[\|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^{c^T}\mathcal{Z}\mathbf{A}_i^c\mathcal{J}_{i,k}^u\big]$$
$$+ \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\big] + \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]\Big]$$
$$+ \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big]. \tag{50}$$

Step (1) is a result of using the quadratic Lyapunov function. The step (2) holds when applying the remote sensing-loop and open-loop state evolution in (19) and (18), respectively. The step (3) holds using the expectation in (50) with respect to the state estimation error $\mathbf{e}_{i,k}^u$, the action prediction error $\mathbf{e}_{i,k}^d$, and the plant noise $\mathbf{w}_k$. As a consequence of obtaining the expected current Lyapunov value in (48) and the expected future Lyapunov value of the UL transmission in (50), the control stability constraint in (25) for the UL transmission is

$$\xi_{i,k}^u\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^s)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u\big]$$
$$+ (1-\xi_{i,k}^u)\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^o)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u\big]$$
$$\leq \zeta_i\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k}^u)|\hat{\mathbf{x}}_{i,k}^u\big]$$
$$= \xi_{i,k}^u\Big[\mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\big]$$
$$- \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\big]\Big]$$
$$+ \|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^T\mathcal{Z}\mathbf{A}_i\mathcal{J}_{i,k}^u\big] + \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]$$
$$+ \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big] \leq \zeta_i\|\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \zeta_i\mathrm{Tr}\big[\mathcal{Z}\mathcal{J}_{i,k}^u\big] \tag{51}$$

where we have defined the following terms $\mathbb{J}_{i,k} := \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\big]$, $\mathbb{M}_{i,k} := \big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]$, $\mathbb{K}_{i,k} := \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\big]$. After rearranging the terms in (51), the constraint on the UL transmission indicator variable is given as

$$\xi_{i,k}^u \geq \frac{\mathbb{A}_{i,k} + \mathbb{B}_{i,k} + \mathbb{M}_{i,k} + \mathrm{Tr}[\mathcal{Z}\mathbf{W}]}{\mathbb{J}_{i,k} - \mathbb{K}_{i,k}}. \tag{52}$$

where we have further defined the following terms $\mathbb{A}_{i,k} := \|(\mathbf{A}_i^c - \zeta_i\mathbf{I}_D)\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2$ and $\mathbb{B}_{i,k} := \mathrm{Tr}\big[(\mathbf{A}_i^T\mathcal{Z}\mathbf{A}_i - \zeta_i\mathcal{Z})\mathcal{J}_{i,k}^u\big]$. To capture the overall state evolution of each control system in the UL over time interval of length K, according to the time-averaged Lyapunov [40], (53) is summed over time $k \in \{0,\cdots,K-1\}$, then the result is divided by K and taking the limit as time tends to infinity. This yields the UL transmission indicator variable constraint in (27).

Similarly, we obtain the DL transmission indicator variable constraint in (28) according to the remote actuating-loop state evolution in (20) and open-loop state evolution in (18). Finally, the expected future Lyapunov value of the UL-DL coupled transmission is obtained using the closed-loop state evolution in (21) and the open-loop state evolution in (18) where

$$\mathbb{E}\Big[\mathcal{L}(\mathbf{x}_{i,k+1}^u)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,\mathbf{H}_{i,k}^d,P_{i,k}^d\Big]$$
$$= \xi_{i,k}^u\,\xi_{i,k}^d\,\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^c)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,\mathbf{H}_{i,k}^d,P_{i,k}^d\big]$$
$$+ \big(1-\xi_{i,k}^u\,\xi_{i,k}^d\big)\,\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^o)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,$$
$$\mathbf{H}_{i,k}^d,P_{i,k}^d\big], \tag{53}$$

For a given predicted state $\hat{\mathbf{x}}_{i,k}^u$, the state prediction error covariance matrix $\mathcal{J}_{i,k}^u$, the predicted action $\hat{\mathbf{u}}_{i,k}^d$, the action prediction error covariance matrix $\mathcal{J}_{i,k}^d$, the wireless UL channel $\mathbf{H}_{i,k}^u$, the wireless DL channel $\mathbf{H}_{i,k}^d$, the UL transmission power $P_{i,k}^u$, and the DL transmission power $P_{i,k}^d$, the expected future value of $\mathcal{L}(\mathbf{x}_{i,k+1}^u)$ of the UL-DL coupling transmission in (53) is given as

$$\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k+1}^u)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,\mathbf{H}_{i,k}^d,P_{i,k}^d\big]$$
$$= \xi_{i,k}^u\,\xi_{i,k}^d\mathbb{E}\big[\mathbf{x}_{i,k+1}^{c^T}\mathcal{Z}\mathbf{x}_{i,k+1}^c|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,\mathbf{H}_{i,k}^d,P_{i,k}^d\big]$$
$$+ \big(1-\xi_{i,k}^u\xi_{i,k}^d\big)\mathbb{E}\big[\mathbf{x}_{i,k+1}^{o^T}\mathcal{Z}\mathbf{x}_{i,k+1}^o|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,$$
$$\mathbf{H}_{i,k}^d,P_{i,k}^d\big]$$
$$= \xi_{i,k}^u\xi_{i,k}^d\mathbb{E}\Big[\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{v}_{i,k}^c - \mathbf{B}_i\mathbf{v}_{i,k}^d + \mathbf{w}_k\big)^T$$
$$\times \mathcal{Z}\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{v}_{i,k}^c - \mathbf{B}_i\mathbf{v}_{i,k}^d + \mathbf{w}_k\big)\Big]$$
$$+ (1-\xi_{i,k}^u\xi_{i,k}^d)\mathbb{E}\Big[\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{e}_{i,k}^u$$
$$- \mathbf{B}_i\mathbf{e}_{i,k}^d + \mathbf{w}_k\big)^T$$
$$\times \mathcal{Z}\big(\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u - \mathbf{A}_i^c\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{\Phi}_i\mathbf{e}_{i,k}^u - \mathbf{B}_i\mathbf{e}_{i,k}^d + \mathbf{w}_k\big)\Big]$$
$$= \xi_{i,k}^u\xi_{i,k}^d\Big[\|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^{c^T}\mathcal{Z}\mathbf{A}_i^c\mathcal{J}_{i,k}^u\big] + \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big]\Big]$$
$$+ \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\big] + \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathbf{V}_{i,k}^d\big]$$
$$+ (1-\xi_{i,k}^u\,\xi_{i,k}^d)\Big[\|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^{c^T}\mathcal{Z}\mathbf{A}_i^c\mathcal{J}_{i,k}^u\big]$$
$$+ \mathrm{Tr}\big[(\mathbf{B}_i^T\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i^T\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\big] + \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]$$
$$+ \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big]\Big] \tag{54}$$

Given the expected current value of $\mathcal{L}(\mathbf{x}_{i,k}^u)$ in (48) and future of $\mathcal{L}(\mathbf{x}_{i,k+1}^u)$ of the UL-DL coupled transmission in (54), the control stability constraint in (25) for UL-DL coupled transmission is

$$\xi_{i,k}^u\xi_{i,k}^d\mathbb{E}\Big[\mathcal{L}(\mathbf{x}_{i,k+1}^a)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,\mathbf{H}_{i,k}^d,P_{i,k}^d\Big]$$
$$+ (1-\xi_{i,k}^u\xi_{i,k}^d)\mathbb{E}\Big[\mathcal{L}(\mathbf{x}_{i,k+1}^o)|\hat{\mathbf{x}}_{i,k}^u,\hat{\mathbf{u}}_{i,k}^d,\mathbf{H}_{i,k}^u,P_{i,k}^u,$$
$$\mathbf{H}_{i,k}^d,P_{i,k}^d\Big]$$
$$\leq \zeta_i\mathbb{E}\big[\mathcal{L}(\mathbf{x}_{i,k}^u)|\hat{\mathbf{x}}_{i,k}^u\big] = \xi_{i,k}^u\xi_{i,k}^d\Big[\mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathbf{V}_{i,k}^u\big]$$
$$+ \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathbf{V}_{i,k}^d\big] - \mathrm{Tr}\big[(\mathbf{B}_i\mathbf{\Phi}_i)^T\mathcal{Z}(\mathbf{B}_i\mathbf{\Phi}_i)\mathcal{J}_{i,k}^u\big]$$
$$- \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big]\Big] + \|\mathbf{A}_i^c\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \mathrm{Tr}\big[\mathbf{A}_i^T\mathcal{Z}\mathbf{A}_i\mathcal{J}_{i,k}^u\big]$$
$$+ \mathrm{Tr}\big[\mathbf{B}_i^T\mathcal{Z}\mathbf{B}_i\mathcal{J}_{i,k}^d\big] \leq \zeta_i\|\hat{\mathbf{x}}_{i,k}^u\|_{\mathcal{Z}^{\frac{1}{2}}}^2 + \zeta_i\mathrm{Tr}\big[\mathcal{Z}\mathcal{J}_{i,k}^u\big]$$
$$+ \mathrm{Tr}\big[\mathcal{Z}\mathbf{W}\big] \tag{55}$$

The UL-DL transmission indicator variable constraint after rearranging the terms in (55) is given as

$$\xi_{i,k}^u \xi_{i,k}^d \geq \frac{\mathbb{A}_{i,k} + \mathbb{B}_{i,k} + \mathbb{M}_{i,k} + \mathrm{Tr}\left[\mathcal{Z}\mathbf{W}\right]}{\left[\mathbb{J}_{i,k} - \mathbb{K}_{i,k}\right] + \left[\mathbb{M}_{i,k} - \mathbb{N}_{i,k}\right]} \tag{56}$$

where we have further defined the following term $\mathbb{N}_{i,k} := \mathrm{Tr}\left[\mathbf{B}_i^T \mathcal{Z} \mathbf{B}_i \mathbf{V}_{i,k}^d\right]$. At last, we apply the time-averaged Lyapunov [40] to obtain the overall state evolution per control system in the UL-DL coupled transmission and UL-DL coupled transmission indicator variable constraint in (29). ∎

REFERENCES

[1] A. M. Girgis, J. Park, C.-F. Liu, and M. Bennis, "Predictive control and communication co-design: A Gaussian process regression approach," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.

[2] P. Park, S. Ergen, C. Fischione, C. Lu, and K. Johansson, "Wireless network design for control systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 978–1013, Dec. 2017.

[3] J. Park *et al.*, "Extreme URLLC: Vision, challenges, and key enablers," 2020, *arXiv:2001.09683*. [Online]. Available: http://arxiv.org/abs/2001.09683

[4] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[5] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, "Wireless network intelligence at the edge," *Proc. IEEE*, vol. 107, no. 11, pp. 2204–2239, Nov. 2019.

[6] C.-F. Liu and M. Bennis, "Taming the tail of maximal information age in wireless industrial networks," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2442–2446, Dec. 2019.

[7] T. Zeng, O. Semiari, W. Saad, and M. Bennis, "Joint communication and control for wireless autonomous vehicular platoon systems," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7907–7922, Nov. 2019.

[8] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proc. IEEE*, vol. 95, no. 1, pp. 138–162, Jan. 2007.

[9] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proc. IEEE*, vol. 95, no. 1, pp. 163–187, Jan. 2007.

[10] Y. Xu, H. Su, Y.-J. Pan, Z.-G. Wu, and W. Xu, "Stability analysis of networked control systems with round-robin scheduling and packet dropouts," *J. Franklin Inst.*, vol. 350, no. 8, pp. 2013–2027, Oct. 2013.

[11] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Comput. Netw.*, vol. 41, no. 4, pp. 451–474, 2003.

[12] A. Cervin and T. Henningsson, "Scheduling of event-triggered controllers on a shared network," in *Proc. IEEE Conf. Decision Control*, Dec. 2008, pp. 3601–3606.

[13] R. Postoyan, P. Tabuada, D. Nešić, and A. Anta, "Event-triggered and self-triggered stabilization of distributed networked control systems," in *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, Dec. 2011, pp. 2565–2570.

[14] G. Zhao, M. A. Imran, Z. Pang, Z. Chen, and L. Li, "Toward real-time control in future wireless networks: Communication-control co-design," *IEEE Commun. Mag.*, vol. 57, no. 2, pp. 138–144, Feb. 2019.

[15] M. Royyan, M. Vehkapera, T. Charalambous, and R. Wichman, "Throughput-maximum energy-aware rate adaptation in W-NCSs over quasi-static fading channels," in *Proc. 59th IEEE Conf. Decis. Control (CDC)*, Dec. 2020, pp. 1860–1865.

[16] X. Tong, G. Zhao, M. A. Imran, Z. Pang, and Z. Chen, "Minimizing wireless resource consumption for packetized predictive control in real-time cyber physical systems," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.

[17] Q. Wang, S. Xie, G. Zhao, L. Zhang, and Z. Chen, "URLLC packet management for packetized predictive control," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–5.

[18] B. Chang, G. Zhao, M. A. Imran, L. Li, and Z. Chen, "Dynamic QoS allocation for real-time wireless control in tactile internet," in *Proc. IEEE 5G World Forum (GWF)*, Jul. 2018, pp. 273–277.

[19] M. Yu, S. Cai, and V. K. N. Lan, "Event-driven sensor scheduling for mission-critical control applications," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1537–1549, Mar. 2019.

[20] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, "Opportunistic control over shared wireless channels," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3140–3155, Dec. 2015.

[21] Y. Ma *et al.*, "Optimal dynamic scheduling of wireless networked control systems," in *Proc. ACM/IEEE Int. Conf. Cyber-Phys. Syst.*, Apr. 2019, pp. 77–86.

[22] M. Eisen, M. M. Rashid, K. Gatsis, D. Cavalcanti, N. Himayat, and A. Ribeiro, "Control aware communication design for time sensitive wireless systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4584–4588.

[23] J. Li, P. Zeng, X. Zong, M. Zheng, and X. Zhang, "Communication and control co-design for wireless sensor networked control systems," in *Proc. 11th World Congr. Intell. Control Automat. (WCICA)*, Jun. 2014, pp. 156–161.

[24] Y. Sadi, S. C. Ergen, and P. Park, "Minimum energy data transmission for wireless networked control systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2163–2175, Apr. 2014.

[25] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2, no. 3. Cambridge, MA, USA: MIT Press, 2006.

[26] A. Kosta *et al.*, *Age Information: A New Concept, Metric, Tool*, vol. 12, no. 3. Boston, MA, USA: Now, Nov. 2017.

[27] W. Yan, H. Qiu, and Y. Xue, "Gaussian process for long-term time-series forecasting," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2009, pp. 3420–3427.

[28] K. J. Åström and B. Wittenmark, *Adaptive Control*. Mineola, NY, USA: Courier Corporation, 2013.

[29] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Commun. Surveys Tuts.*, vol. 9, no. 2, pp. 18–48, 2nd Quart., 2007.

[30] M. Hassanin, O. Fresnedo, J. Garcia-Frias, and L. Castedo, "Analog joint source channel coding for Gaussian multiple access channels," in *Proc. Int. Workshop Multiple Access Commun.*, 2013, pp. 21–32.

[31] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. Inf. Process. Sensor Netw.*, Apr. 2003, pp. 162–177.

[32] M. S. Mahmoud, *Control and Estimation Methods Over Communication Networks*. Cham, Switzerland: Springer, Mar. 2014.

[33] A. V. Oppenheim and G. C. Verghese, *Introduction to Communication, Control, and Signal Processing*. Cambridge, MA, USA: MIT Press, 2010, ch. 8.

[34] M. A. Alvarez, L. Rosasco, and N. D. Lawrence, "Kernels for vector-valued functions: A review," *Found. Trends Mach. Learn.*, vol. 4, no. 3, pp. 195–266, 2012.

[35] L. Li, M. Pal, and Y. R. Yang, "Proportional fairness in multi-rate wireless LANs," in *Proc. IEEE INFOCOM*, Phoenix, AZ, USA, Apr. 2008, pp. 1004–1012.

[36] M. J. Neely, "Stability and capacity regions or discrete time queueing networks," 2010, *arXiv:1003.3396*. [Online]. Available: http://arxiv.org/abs/1003.3396

[37] M. S. Branicky, S. M. Phillips, and W. Zhang, "Scheduling and feedback co-design for networked control systems," in *Proc. 41st IEEE Conf. Decis. Control*, vol. 2, Dec. 2002, pp. 1211–1217.

[38] M. Kanj, V. Savaux, and M. L. Guen, "A tutorial on NB-IoT physical layer design," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2408–2446, Sep. 2020.

[39] A. Khina, G. M. Pettersson, V. Kostina, and B. Hassibi, "Multi-rate control over AWGN channels via analog joint source-channel coding," in *Proc. IEEE 55th Conf. Decis. Control (CDC)*, Dec. 2016, pp. 5968–5973.

[40] A. N. Michel and L. Hou, "Stability results involving time-averaged Lyapunov function derivatives," *Nonlinear Anal., Hybrid Syst.*, vol. 3, no. 1, pp. 51–64, Feb. 2009.

Abanoub M. Girgis (Student Member, IEEE) received the B.Sc. degree (Hons.) in electronics and communications engineering from Thebes Higher Institute of Engineering, Egypt, in 2013, and the M.Sc. degree in electronics and communications engineering from Ain Shams University, Egypt, in 2018. He is currently pursuing the Ph.D. degree with the Department of Information Technology and Electrical Engineering, University of Oulu, Finland. He was a Teaching Assistant with Thebes Higher Institute of Engineering from 2013 to 2018. His research interests include cyber-physical systems, ultra-reliable and low-latency communications, machine learning, 5G communications, massive MIMO, and signal processing.

Jihong Park (Member, IEEE) received the B.S. and Ph.D. degrees from Yonsei University, South Korea. He is currently a Lecturer with the School of IT, Deakin University, Australia. His research interests include ultra-dense/ultra-reliable/mmWave system designs, and distributed learning/control/ledger technologies and their applications for beyond-5G/6G communication systems. He served as a Conference/Workshop Program Committee Member for IEEE GLOBECOM, ICC, and WCNC, and for NeurIPS, ICML, and IJCAI. He is an Associate Editor of *Frontiers in Data Science for Communications,* and a Review Editor of *Frontiers in Aerial and Space Networks*.

Mehdi Bennis (Fellow, IEEE) is currently an Associate Professor at the Centre for Wireless Communications, University of Oulu, Finland, and the Head of the Intelligent Connectivity and Networks/Systems Group (ICON). He has published more than 200 research papers in international conferences, journals, and book chapters. His main research interests include radio resource management, heterogeneous networks, game theory, and distributed machine learning in 5G networks and beyond. He is a fellow of the Academy of Finland Research. He was a recipient of several prestigious awards including the 2015 Fred W. Ellersick Prize from the IEEE Communications Society, the 2016 Best Tutorial Prize from the IEEE Communications Society, the 2017 EURASIP Best paper Award for the *Journal of Wireless Communications and Networks,* the all-University of Oulu award for research, the 2019 IEEE ComSoc Radio Communications Committee Early Achievement Award, and the 2020 Clarviate Highly Cited Researcher by the Web of Science. He is an editor of IEEE TRANSACTIONS ON COMMUNICATIONS (TCOM) and the Specialty Chief Editor of Data Science for Communications in *Frontiers in Communications and Networks* journal.

Mérouane Debbah (Fellow, IEEE) received the M.Sc. and Ph.D. degrees from the Ecole Normale Supérieure Paris-Saclay, France. He was with Motorola Labs, Saclay, France, from 1999 to 2002, and also with Vienna Research Center for Telecommunications, Vienna, Austria, until 2003. From 2003 to 2007, he was an Assistant Professor with the Mobile Communications Department, Institut Eurecom, Sophia Antipolis, France. In 2007, he was appointed as a Full Professor at CentraleSupelec, Gif-sur-Yvette, France. From 2007 to 2014, he was the Director of the Alcatel-Lucent Chair on Flexible Radio. From 2014 to 2021, he was the Vice-President of the Huawei France Research Center. He was jointly the Director of the Mathematical and Algorithmic Sciences Laboratory, and Lagrange Mathematical and Computing Research Center. Since 2021, he has been the Chief Research Officer with the Technology Innovation Institute, Abu Dhabi. He also leads jointly the AI and telecommunication centers. He has managed 8 EU projects and more than 24 national and international projects. His research interests lie in fundamental mathematics, algorithms, statistics, information, and communication sciences research. He is a WWRF Fellow, a Eurasip Fellow, an Institut Louis Bachelier Fellow, and a Membre émérite SEE. He was a recipient of the ERC Grant MORE (Advanced Mathematical Tools for Complex Network Engineering) from 2012 to 2017, the Mario Boella Award in 2005, the IEEE Glavieux Prize Award in 2011, the Qualcomm Innovation Prize Award in 2012, the 2019 IEEE Radio Communications Committee Technical Recognition Award, and the 2020 SEE Blondel Medal. He received more than 20 best paper awards, among which the 2007 IEEE GLOBECOM Best Paper Award, the Wi-Opt 2009 Best Paper Award, the 2010 Newcom++ Best Paper Award, the WUN CogCom Best Paper 2012 and 2013 Award, the 2014 WCNC Best Paper Award, the 2015 ICC Best Paper Award, the 2015 IEEE Communications Society Leonard G. Abraham Prize, the 2015 IEEE Communications Society Fred W. Ellersick Prize, the 2016 IEEE Communications Society Best Tutorial Paper Award, the 2016 European Wireless Best Paper Award, the 2017 Eurasip Best Paper Award, the 2018 IEEE Marconi Prize Paper Award, the 2019 IEEE Communications Society Young Author Best Paper Award, the 2021 Eurasip Best Paper Award, the 2021 IEEE Marconi Prize Paper Award, and the Valuetools 2007, Valuetools 2008, CrownCom 2009, Valuetools 2012, SAM 2014, and 2017 IEEE Sweden VT-COM-IT Joint Chapter best student paper awards. He is currently the Associate Editor-in-Chief of the journal *Random Matrices: Theory and Applications*. He was an Associate Area Editor and a Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2011 to 2013 and from 2013 to 2014, respectively. Since 2021, he has been serving as an IEEE Signal Processing Society Distinguished Industry Speaker.