Channel Access Strategies for Control–Communication Co-Designed Networks

Gourab Ghatak, *Member, IEEE*, Geethu Joseph, *Senior Member, IEEE*, and Chen Quan, *Member, IEEE*

***Abstract*—In this article, we develop a communication–control co-design framework in a wireless networked control system with multiple geographically separated controllers and controlled systems, modeled via a Poisson point process. Each controlled system consists of an actuator, plant, and sensor. Controllers receive state estimates from sensors and design control inputs, which are sent to actuators over a shared wireless channel, causing interference. Our co-design includes control strategies at the controller based on sensor measurements and transmission acknowledgments from the actuators for both rested and restless systems—systems with and without state feedback, respectively. In the restless system, controllability depends on consecutive successful transmissions, while in the rested system, it depends on total successful transmissions. We use both classical and block ALOHA protocols for channel access, optimizing access based on sensor data and acknowledgments. A statistical analysis of control performance is followed by a Thompson sampling-based algorithm to optimize the ALOHA parameter, achieving sublinear regret. We show how the ALOHA parameter influences control performance and transmission success.**

***Index Terms*—ALOHA protocol, poisson point process (PPP), shared wireless channel, Thompson sampling (TS).**

I. Introduction

THE rapid evolution of wireless networks and the rising demand for new applications, such as remote surgery in advanced medical systems, robotics in industrial automation, intelligent buildings, and autonomous driving [1], [2], demand real-time control systems over wireless channels. These networked control systems comprise multiple controllers, sensors, actuators, and plants connected through a shared communication medium. Using shared network resources introduces new design challenges that affect the control performance in a networked control system, including unreliable communication links [3], transmission delays [4], time synchronization issues [5], and network access constraints and conflicts [6]. Furthermore, due to the ad hoc deployment of wireless access points, especially in the unlicensed bands, the interfering signals from co-channel transmissions of different controllers may severely degrade the performance of a wireless control system. These issues underscore the need for a joint communication–control design of a large-scale wireless networked control system, which we address in this article.

Prior research has mainly focused on networked control systems with a single controller and plant with one or more actuators and sensors communicating over wireless links. Current literature discusses two principal approaches to designing these systems: independent control–communication design, which treats control and communication components separately [7], [8], [9], and control–communication co-design, which integrates both for improved performance [10], [11], [12], [13], [14], [15]. This article focuses on the co-design approach, which is better suited for real-time control applications in wireless networks [11]. Existing studies have explored several aspects of co-design problems, such as stabilizing control systems and enhancing network security with communication imperfections [10] and maintaining system reachability and observability under limited communication resources [16]. Other works look at handling packet loss in cloud-controlled systems [11], missing state information from sensors [12], and packet loss in sensor-to-controller and controller-to-actuator channels. Another research direction studies the age-of-information metric to assess the reliability and freshness of information in wireless control systems, with [14] exploring scheduling and power allocation, and [15] optimizing control costs and energy consumption. In short, co-design methods typically involve two optimization types. The first type optimizes control objectives, such as control costs or control performance metrics, such as stabilization and tracking error, under communication constraints, such as time delay, bandwidth, packet loss probability, number of communication channels, and sampling period [17], [18], [19], [20], and the second designs communication protocols to meet control performance goals [18], [21], [22]. While several efforts have been made for joint communication–control designs, large-scale wireless control systems with multiple interfering controllers have not been well studied in the literature. This article addresses this issue by considering interfering signals from the randomized locations of control and actuator pairs.

Our model addresses a wireless networked control system involving multiple geographically separated controllers and

Received 9 June 2025; revised 6 October 2025; accepted 8 October 2025. Date of publication 13 October 2025; date of current version 19 March 2026. This work was supported in part by the IIT Delhi-TU Delft MFIRP under Project MI03219G titled Sensing-Control-Communication Co-Design for Complex Networked Systems. Recommended by Associate Editor J. Wu. *(Corresponding author: Gourab Ghatak.)*

Gourab Ghatak is with the Department of Electrical Engineering, IIT Delhi, Delhi 110016, India (e-mail: gghatak@ee.iitd.ac.in).

Geethu Joseph and Chen Quan are with the Faculty of Electrical Engineering, Mathematics, and Computer Science, TU Delft, 2628 CD Delft, The Netherlands (e-mail: g.joseph@tudelft.nl; C.Quan@tudelft.nl).

Digital Object Identifier 10.1109/TCNS.2025.3621024

controlled systems. Controllers receive state estimates from sensors and transmit control inputs to actuators over a shared, interference-prone channel. A co-design framework integrates control decisions and actuator acknowledgments with a block ALOHA protocol for managing access. This requires accurate interference characterization, combining control and communication theory. We use stochastic geometry, specifically, a homogeneous Poisson point process (PPP), to model controller locations and capture the effect of co-channel transmissions. Our main contributions are as follows.

1) Communication–Control Co-Design: We study two types of control systems: the restless model, where failed transmissions lead to zero control, and the rested model, which uses a local state feedback. For both systems, we present a communication–control co-design with a wireless channel acknowledgment-based control input design and a block ALOHA protocol to ensure that the system state is driven to the desired state. For this design, we define the metric of block controllability: determined by burst length in restless systems and by total successful transmissions in rested systems.

2) Statistical Analysis for the Restless System: For the typical restless system, we derive the conditional success probability of a transmission, using which we analyze the conditional distribution $\bar{F}_{\mathrm{RL,blk}}$ of the burst length. Since deriving the exact meta distribution of the burst length is challenging, we characterize the system's performance based on the first moment of $\bar{F}_{\mathrm{RL,blk}}$ weighted by the channel access probability.

3) Statistical Analysis for the Rested System: For the rested control system, we derive the conditional transmission success probability, compare it with the restless case, and use this to analyze successful transmissions over a time horizon via the meta distribution. We also reconstruct the approximate meta distribution using a finite number of moments.

4) Learning-Based Channel Access: To optimize ALOHA parameters online, we formulate an multiarmed bandit (MAB) problem, solved using a Thompson sampling (TS) variant. For both rested and restless systems, regret grows sublinearly with respect to the number of blocks and the number of slots per block. Simulations confirm that online learning of the optimal ALOHA parameter can improve control performance.

5) System Design Insights: Our work shows that for a rested system, a high channel access probability is beneficial despite increased interference, due to competition for transmission. In contrast, restless systems can reduce access probability after a successful transmission to limit interference. We show how controller density and required control inputs affect performance, providing design guidelines. Finally, we show that classical ALOHA yields higher block controllability in rested systems, while block ALOHA does so in restless systems. Notably, maximizing transmission success does not always optimize controllability, highlighting the need for co-design.

Overall, our work establishes new synergies among networked control systems, stochastic geometry, and learning-based channel access strategies, offering new theoretical and algorithmic insights. Compared to the conference version [23], we extend the restless system analysis, showing sublinear regret of the TS algorithm. We also analyze the success probability of rested systems and show that frequent channel access helps rested systems, while restless systems benefit from reduced access to limit interference.

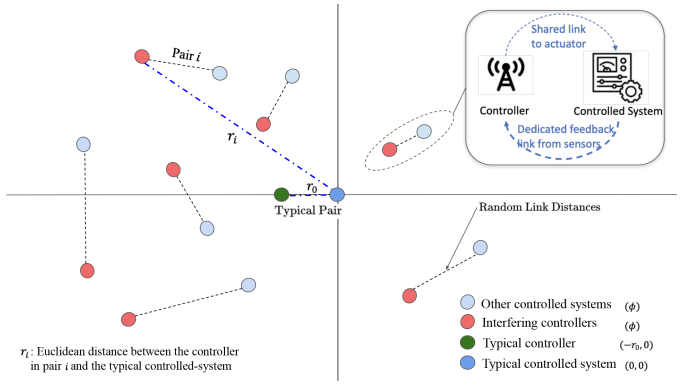


Fig. 1. In Poisson network of controller–controlled system pairs, the actuators receive the control input from the corresponding controller via a shared link, whereas the sensors send their observations to the controller via a dedicated link.

II. Networked Control System Model

Fig. 1, Consider a large network comprising multiple processes and systems, each with its own independent control mechanisms. Each system in the network is represented by a controller–controlled system pair, where the controlled system consists of an actuator, plant, and sensor. We model our network using a Poisson bipolar point process consisting of controller–controlled system pairs in the 2-D Euclidean plane $\mathbb{R}^2$ [24]. The locations of the controllers are modeled as a homogeneous PPP Φ of intensity λ. Recall that PPP is a random collection of points in the Euclidean plane, wherein the number of points in a compact closed set is Poisson distributed with parameter λ times the Lebesgue measure of the set. The controller–controlled system pairs are indexed as $i = 0, 1, 2, \ldots$, with r_i denoting the Euclidean distance between the ith controller and the typical controlled system. The controller–controlled system pair index $i = 0$ refers to the typical pair. Without loss of generality, we assume a typical controlled system at the origin, and the typical controller is at a distance of r_0 away from the origin.

In this setup, the controllers periodically receive the state of their respective controlled systems through a dedicated channel. Based on these system states, each controller computes the control inputs and communicates them to the actuator via a shared wireless link. Since all controller–actuator pairs utilize the same communication resources (time and frequency), interference can occur, resulting in control performance degradation. Our aim is to design a probabilistic channel access protocol that learns and optimizes the control performance.

We emphasize that the Poisson bipolar process assumption stems from the one-to-one pairing between controllers and controlled systems. Also, Slivnyak's theorem [25] guarantees that the process Φ conditioned on the location of the typical pair is a PPP with the same statistics as Φ. In other words, conditioning the PPP Φ on the location of this typical pair leaves the rest of the process statistically unchanged. This allows us to

analyze the network from the perspective of the typical pair while treating the interfering controllers as an independent PPP with the same intensity λ, simplifying the analysis without any loss of generality. The following sections elaborate on our system model.

A. Controller–Controlled System Model

Each controlled system in the network is modeled as a discrete-time linear dynamical system with a controllability index v. The typical system is

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{v}(t). \tag{1}$$

Here, $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state of the typical controlled system in the network at discrete time t, and $\boldsymbol{u}(t) \in \mathbb{R}^m$ and $\boldsymbol{v}(t) \in \mathbb{R}^n$ are its input and process noise, respectively, at time $t \in \mathbb{Z}_+$. We assume that $\boldsymbol{v}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$, where $\boldsymbol{Q}$ is the covariance of $\boldsymbol{v}(t)$. Also, $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times m}$ denote the state and input matrices of the system, respectively. The system aims to drive and retain the system state at a desired state $\boldsymbol{x}_{\text{des}}$ with a suitable choice of control inputs $\boldsymbol{u}(t)$. We emphasize that our work assumes only that all systems share the same controllability index v, not necessarily identical dynamics or desired states.

We assume the actuator computing capabilities are limited, with all computations being offloaded to the controller. For this, the sensors observe the system state $\boldsymbol{x}(t)$ periodically at time $t = kT$ for $k \in \mathbb{Z}_+$, where the block length T is the delay between two consecutive observations. Here, k is referred to as the block index, and the controller gets the state estimate at the beginning of each block. For simplicity, we assume that the state $\boldsymbol{x}(t)$ is communicated to the controller via a reliable link with dedicated resources, ensuring a good system state estimate $\hat{\boldsymbol{x}}(kT)$ at the controller at time kT for $k \in \mathbb{Z}_+$. Since this communication occurs only once per block (i.e., sporadically), we assume that the state $\boldsymbol{x}(t)$ is communicated to the controller via a reliable link with dedicated resources, ensuring that the controller knows $\boldsymbol{x}(t)$ at time kT for $k \in \mathbb{Z}_+$.

The controller estimates the control inputs needed to drive the system to the desired state $\boldsymbol{x}_{\text{des}}$, which are sequentially communicated to the actuator via a shared link. The controller transmissions are synchronized with the controlled system evolution in (1). Therefore, the discrete-time index t also denotes the transmission slot index. Due to noise and interference from the other transmitting controllers, the controlled system may not always correctly decode the received signal, leading to an unsuccessful transmission. The success of the transmission depends on the wireless channel, modeled next.

B. Channel Model and Transmission Success

Each wireless link experiences fast fading, assumed to be Rayleigh distributed with parameter 1. The fast fading is independent across time and spatially independent across all links. Also, ρ and α denote the path-loss constant and the path-loss exponent of the channel, respectively [26]. We denote the channel noise power as N_0 and the transmit power as η.

A controller's decision to transmit in a time slot t depends on the channel access protocol. Let the channel access state $C_i(t) \in \{0, 1\}$ be the indicator variable representing whether the ith controller transmits at time t, for $i = 0, 1, \ldots$. Given that the typical controller transmits, the signal-to-interference-plus-noise ratio (SINR) at the typical actuator is

$$\xi(t) = \frac{\eta\rho|h_0(t)|^2 r_0^{-\alpha}}{N_0 + \sum_{i\in\phi} C_i(t)\eta\rho|h_i(t)|^2 r_i^{-\alpha}} \tag{2}$$

where $h_i(t)$ is the channel fading between the ith controller and the typical actuator [27], [28].

If the SINR at time t exceeds a threshold $\gamma > 0$, then the actuator correctly decodes the received signal, thereby implying a successful transmission. This threshold γ depends on the application's data rate requirements and the receiver's sensitivity, and is set by the network based on these factors rather than being estimated [27]. The successful transmission by the typical controller at time t is indicated by $S(t) \in \{0, 1\}$

$$S(t) = \begin{cases} 1, & \text{if } C_0(t)\xi(t) > \gamma \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Also, the controller receives an acknowledgment of transmission $S(t)$ from the actuator. Since the feedback is one bit in size, we assume dedicated resources for the feedback channel for each controller–controlled system pair. Furthermore, the small-sized feedback data (low payload) also ensures the possibility of packing a high fraction of redundant bits in the transmit packets to take advantage of advanced forward error correcting codes. Accordingly, we assume that the feedback channel is interference-free and noiseless. The controller uses $S(t)$ to design and transmit the control inputs, discussed next.

III. Communication–Control Co-Design

The co-design involves two key components: a control input design based on periodic sensor measurements and transmission acknowledgments, and a random access data transmission scheme that operates independently, without centralized scheduling. We present a co-design approach for two system types: *restless* (without direct state feedback) and *rested systems* (with direct state feedback).

A. Restless System

The restless system lacks a direct state-dependent feedback loop, and its actuator applies only the control inputs that are either sent by the controller or are predefined and stored. The control input design at the controller is based on the system state $\boldsymbol{x}(kT)$ obtained at the beginning of block k. Since $\boldsymbol{v}(t) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$, it follows from (1) that $\boldsymbol{x}(t)$ is a Gaussian random variable, given $\boldsymbol{x}(kT)$. Its mean $\hat{\boldsymbol{x}}(t)$ depends on the control inputs, but its covariance $\sum_{\tau=kT}^{t-1} \boldsymbol{A}^{t-\tau-1}\boldsymbol{Q}[\boldsymbol{A}^{t-\tau-1}]^\top$ is independent of the control inputs, where

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{A}^{t-kT}\hat{\boldsymbol{x}}(kT) + \sum_{\tau=kT}^{t-1} \boldsymbol{A}^{t-\tau-1} S(\tau)\boldsymbol{B}\boldsymbol{u}(\tau). \tag{4}$$

Here, we note that $\hat{\boldsymbol{x}}(kT) = \boldsymbol{x}(kT)$. Since control inputs $\boldsymbol{u}(\tau)$ only affect the mean $\hat{\boldsymbol{x}}(t)$, the controller designs v control inputs

$\{\hat{\boldsymbol{u}}(t+\tau), \tau = 0, 1, \ldots, v-1\}$ at time $t = kT$ such that the mean is driven to $\boldsymbol{x}_{\text{des}}$, i.e.,

$$\boldsymbol{x}_{\text{des}} = \boldsymbol{A}^v \hat{\boldsymbol{x}}(t) + \sum_{\tau=0}^{v-1} \boldsymbol{A}^{v-1-\tau} \boldsymbol{B} \hat{\boldsymbol{u}}(t+\tau) \tag{5}$$

where v is the controllability index of the linear dynamical system in (1). So, the controller solves for control inputs using the least squares solution to obtain

$$\begin{aligned} &\left[\hat{\boldsymbol{u}}(t)^\top \quad \hat{\boldsymbol{u}}(t+1)^\top \ldots \hat{\boldsymbol{u}}(t+v-1)^\top\right]^\top \\ &= \boldsymbol{\Psi}^\dagger \left[\boldsymbol{x}_{\text{des}} - \boldsymbol{A}^v \hat{\boldsymbol{x}}(t)\right] \end{aligned} \tag{6}$$

with $t = kT$, and we define

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{A}^{v-1}\boldsymbol{B} & \boldsymbol{A}^{v-2}\boldsymbol{B} & \ldots & \boldsymbol{B} \end{bmatrix}.$$

If the transmission of $\hat{\boldsymbol{u}}(kT)$ is successful, i.e., $S(kT) = 1$, then it continues to transmit the next control inputs $\hat{\boldsymbol{u}}(kT+1), \hat{\boldsymbol{u}}(kT+2), \ldots, \hat{\boldsymbol{u}}(kT+L-1)$ until either $S(kT+L-1) = 0$ or $L = v$. If $S(kT+L-1) = 0$ for some $L \leq v$, then the controller recomputes the next set of v inputs using (6) with $t = kT + L$ and repeats the above steps. If $S(kT+L-1) = 1$ for $L = 1, \ldots, v$, then all the v control inputs designed by the controller have reached the actuator. Here, the burst length L is the number of consecutive successful transmissions.

Once all the v control inputs are applied, we get $\hat{\boldsymbol{x}}(t) = \boldsymbol{x}_{\text{des}}$. Then, the input required to retain $\hat{\boldsymbol{x}}(t)$ at the desired state $\boldsymbol{x}_{\text{des}}$ can be computed by solving

$$\boldsymbol{x}_{\text{des}} = \boldsymbol{A}^{\bar{v}} \boldsymbol{x}_{\text{des}} + \sum_{\tau=0}^{\bar{v}-1} \boldsymbol{A}^{\bar{v}-\tau-1} \boldsymbol{B} \bar{\boldsymbol{u}}(\tau).$$

Here, $\bar{v} \leq v$ is the smallest integer for which $(\boldsymbol{I} - \boldsymbol{A}^{\bar{v}})\boldsymbol{x}_{\text{des}} \in \text{Range}[(\boldsymbol{A}^{\bar{v}-1}\boldsymbol{B} \quad \boldsymbol{A}^{\bar{v}-2}\boldsymbol{B} \quad \ldots \quad \boldsymbol{B})]$. Thus, we derive

$$\left[\bar{\boldsymbol{u}}(0)^\top \quad \bar{\boldsymbol{u}}(1)^\top \ldots \bar{\boldsymbol{u}}(\bar{v}-1)^\top\right]^\top = \bar{\boldsymbol{\Psi}}^\dagger \left[\boldsymbol{I} - \boldsymbol{A}^{\bar{v}}\right] \boldsymbol{x}_{\text{des}} \tag{7}$$

where $\bar{\boldsymbol{\Psi}} = [\boldsymbol{A}^{\bar{v}-1}\boldsymbol{B} \quad \boldsymbol{A}^{\bar{v}-2}\boldsymbol{B} \quad \ldots \quad \boldsymbol{B}]$. The aforementioned control inputs ensure that $\hat{\boldsymbol{x}}(t)$ is driven to $\boldsymbol{x}_{\text{des}}$. Further, the control input in (7) is state-independent and can be computed offline. Thus, it can be prestored at the actuator and reused after receiving and applying v inputs from the controller, without further computation. This continues till the end of the block when the sensor sends the state information to the controller at time $(k+1)T$. Overall, in block k, the actuator applies the following control inputs:

$$\boldsymbol{u}(t) = \begin{cases} S(t)\hat{\boldsymbol{u}}(t), & \text{if } \sum_{\tau=kT}^{t} S(\tau) \leq v \\ \bar{\boldsymbol{u}}((t-\bar{t}) \bmod \bar{v}), & \text{otherwise} \end{cases}$$

where $\bar{t}$ is the smallest time t for which $\sum_{\tau=kT}^{t} S(\tau) > v$. Note that $\boldsymbol{u}(t)$ is the control input received by the actuator, while $\hat{\boldsymbol{u}}(t)$ is the control input sent by the controller. The algorithm for data transmission is provided in Algorithm 1.

Along with the control input design in (6), the controller also needs to devise a channel access policy. Within the Poisson network model, we aim to optimize this policy based on actuator acknowledgments $S(t)$ in (3) to maximize its probability of successfully steering its state estimate in (4) to the desired state $\boldsymbol{x}_{\text{des}}$. For the restless system, the controller can drive its state estimate to the desired state $\boldsymbol{x}_{\text{des}}$ only if v consecutive transmissions are successful. This notation of control performance is captured by the following event.

Algorithm 1: Control Design and Data Transmission of Typical Controller for the Restless System.

Parameters: Block k, system matrices $\boldsymbol{A}$, $\boldsymbol{B}$, desired state $\boldsymbol{x}_{\text{des}}$, channel access $\{C_0(kT+\tau), \tau = 0, 1, \ldots, T-1\}$
Initialization: Time $t = kT$, burst length $L = 0$
Define v as the controllability index of the system
Receive $\hat{\boldsymbol{x}}(kT)$ from the sensor
for $t = kT, kT+1, \ldots, (k+1)T-1$ **do**
  if $C_0(t) = 1$ **then**
// if the last tx fails, redesign inputs
    if $L = 0$ **then**
      Compute $\hat{\boldsymbol{u}}(t), \hat{\boldsymbol{u}}(t+1), \ldots, \hat{\boldsymbol{u}}(t+v-1)$ via (6)
// if block controllable, transmit dummy data
    else if $L = v$
      Set $\hat{\boldsymbol{u}}(t) = \boldsymbol{1}$
    end if
    Transmit $\hat{\boldsymbol{u}}(t)$ and receive $S(t)$
// if tx fails, reset burst length to 0; otherwise increment
    if $L < v$ **then**
      $L \leftarrow S(t)(L+1)$
    end if
  end if
else for

Definition 1: Consider the typical restless system $(\boldsymbol{A}, \boldsymbol{B})$ with v being the controllability index of the system. The system is said to *block controllable* in a given block k if there is a run of at least v ones in the sequence $S(kT), S(kT+1), \ldots, S((k+1)T-1)$, where $S(t) \in \{0, 1\}$ given by (3) denotes the success of transmission from the typical controller to the typical actuator at time t.

We note that the above notion assumes $T \geq v$. If $T < v$, then the controller can still compute inputs to steer the state close to $\boldsymbol{x}_{\text{des}}$, although controllability is not guaranteed; performance can then be measured by achieving at least v_{mod} successful transmissions, for some $v_{\text{mod}} < T < v$ where our framework still applies. Exploring other success metrics lies beyond this article's scope and is left for future co-design work.

Having presented the restless system, its control strategy, and its controllability, we now describe the rested system before presenting the channel access policy.

B. Rested System

The rested system uses a direct state-dependent feedback loop to handle interruptions from missing control signals due to unsuccessful transmission from the controller. It is a special class of control system where we assume that the column space

of $\boldsymbol{I}-\boldsymbol{A}$ is contained within the column space of $\boldsymbol{B}$. This assumption holds, for example, when $\boldsymbol{B}$ has full row rank, as in fully actuated or overactuated systems. The actuators switch between controller inputs and the local feedback loop based on the success of the transmission, as elaborated below.

At the start of each block k, the controller computes v control inputs using $\hat{\boldsymbol{x}}(kT)=\boldsymbol{x}(kT)$, as in the restless system. These inputs, $\{\hat{\boldsymbol{u}}(kT+\tau)\}_{\tau=0}^{v-1}$, are given by (6) with $t=kT$ and transmitted sequentially to the actuator. If the transmission is successful, then the corresponding input is applied by the actuator; otherwise, the actuator switches to a state-dependent feedback loop instead of applying zero input. When transmission fails, i.e., $S(t)=0$, $\boldsymbol{u}(t)=\boldsymbol{B}^{\dagger}(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}(t)$. The above input ensures that $\hat{\boldsymbol{x}}(t+1)=\hat{\boldsymbol{x}}(t)$ when $S(t)=0$, as the column space of $\boldsymbol{I}-\boldsymbol{A}$ is contained within the column space of $\boldsymbol{B}$. Thus, the control input design at the controller need not be recalculated if the transmission fails. Consequently, the controller keeps attempting to send the same control input until it succeeds. Once all v control inputs from the controller are successfully received by the actuator and applied, the actuator switches to the feedback loop till the end of the block. To summarize, let $\Lambda(k,t)=\sum_{\tau=kT}^{t-1}S(\tau)$ be the number of successful transmissions up to time $t-1$ in block k. Then, the control input at time t in block k is given by

$$\boldsymbol{u}(t)=\begin{cases}\hat{\boldsymbol{u}}(kT+\Lambda(k,t)), & \text{if } S(t)=1 \text{ and } \Lambda(k,t)<v\\ \boldsymbol{B}^{\dagger}\left(\boldsymbol{I}-\boldsymbol{A}\right)\boldsymbol{x}(t), & \text{otherwise.}\end{cases}$$

Thus, the effective model for $\hat{\boldsymbol{x}}(t)$ resembles (4) for the restless system, but it evolves only over $\Lambda(k,t)$ steps, yielding

$$\hat{\boldsymbol{x}}(t)=\boldsymbol{A}^{\Lambda(k,t)}\hat{\boldsymbol{x}}(kT)+\sum_{\tau=0}^{\Lambda(k,t)-1}\boldsymbol{A}^{\Lambda(k,t)-1-\tau}\boldsymbol{B}\hat{\boldsymbol{u}}(kT+\tau).$$

We emphasize that while local feedback helps keep the state near $\boldsymbol{x}_{\text{des}}$ over short durations, process noise can gradually push the state away, making remote control essential for actively steering it back with bounded error.

Furthermore, we reiterate that the key difference between rested and restless systems is that the actuator in the rested system does not use predefined stored inputs if the transmission fails or once all v inputs are received; instead, it switches to feedback-based control. In addition, the controller computes the control inputs only once per block, with no recomputation after transmission failures. This approach simplifies the control input design process compared to the restless system. The algorithm for data transmission is provided in Algorithm 2.

In the rested system, due to the feedback loop at the actuator, the controller can drive its state estimate to the desired state x_{des} if there are v successful (not necessarily consecutive) transmissions in a block of $T\geq v$ slots. Consequently, the notation of control performance is captured as follows.

Definition 2: Consider the typical rested system $(\boldsymbol{A},\boldsymbol{B})$ with v being the controllability index of the system. The system is said to *block controllable* in a given block k if the sequence $S(kT),S(kT+1),\ldots,S((k+1)T-1)$ satisfies $\sum_{\tau=0}^{T-1}S(kT+\tau)\geq v$, where $S(t)\in\{0,1\}$ given by (3) denotes the success of transmission from the typical controller to the typical actuator at time t.

Algorithm 2: Control Design and Data Transmission of Typical Controller for the Rested System.

Parameters: Block k, system matrices $\boldsymbol{A}$, $\boldsymbol{B}$, desired state $\boldsymbol{x}_{\text{des}}$, channel access $\{C_0(kT+\tau),\tau=0,1,\ldots,T-1\}$
Initialization: Time $t=kT$, Number of successes $\Lambda=0$
Define v as the controllability index of the system
Receive $\hat{\boldsymbol{x}}(kT)$ from the sensor
Compute $\hat{\boldsymbol{u}}(kT),\hat{\boldsymbol{u}}(kT+1),\ldots,\hat{\boldsymbol{u}}(kT+v-1)$ via (6)
for $t=kT,kT+1,\ldots,(k+1)T-1$ **do**
// if not block controllable, transmit the next control input
  if $\Lambda<v$ and $C_0(t)=1$ **then**
    Transmit $\hat{\boldsymbol{u}}(kT+\Lambda)$ and receive $S(t)$
    Update $\Lambda\leftarrow\Lambda+S(t)$
  // if block controllable, transmit dummy data
  else if $\Lambda=v$ and $C_0(t)=1$
    Transmit $\hat{\boldsymbol{u}}(t)=\boldsymbol{1}$
  end if
else for

C. ALOHA-Based Channel Access Policy

After designing the control inputs and transmission scheme, the next crucial step is to develop the channel access state $C_i(t)$, without coordination among the controllers. We note that the controllers lack knowledge of the spatial configuration (density and locations) of interfering controller–system pairs in the network, as well as their transmission states. Therefore, we explore a random channel access strategy, specifically, the ALOHA protocol—a widely used multiple-access method for transmitting data over a shared network channel without centralized coordination [29].

We consider two versions of the protocol: classical ALOHA and a modified variant known as *block ALOHA*. In classical ALOHA, each controller independently accesses the channel in every time slot with a fixed probability $q\in\mathcal{P}$ selected from a finite set of D access probabilities $\mathcal{P}=\{p_1,p_2,\ldots,p_D\}$. The channel access probability q is called the ALOHA parameter. As a result, the set of controllers transmitting simultaneously may vary from slot to slot within a block. We refer the readers to [30] for a detailed discussion on this. Clearly, this strategy is not suitable for restless systems as it requires consecutive successful transmission, motivating the alternative approach of block ALOHA. In the block ALOHA channel access protocol, each controller is either active or idle during an entire block k with a probability $q\in\mathcal{P}$. Therefore, the channel access state $C_i(t)$ for $i=0,1,\ldots$ remains the same for all values of $t=kT,kT+1,\ldots,(k+1)T-1$ within a given block k. This reduces slot-to-slot interference variability and aligns better with block controllability, suiting our block controllability notion.

To the best of the authors' knowledge, this is the first formal treatment of block ALOHA, which arises naturally in control–communication co-design but is not typically considered in standard wireless networks.

Under these protocols, the controller continues to transmit even after v successful transmissions, which are consecutive for the restless system but not necessarily consecutive for the rested system. Following these v transmissions, the controller sends dummy data to the actuator. While the actuators do not use these inputs from a controllability perspective, they continue to send acknowledgments $S(t)$ upon successfully receiving data from the controller. From a control perspective, continued transmissions after achieving block controllability result in resource inefficiency and increased interference. However, from a communication-learning standpoint, they facilitate accurate estimation of controller density, which is crucial for adaptive tuning of the ALOHA parameter, and allow for a more tractable analysis. While our framework can be extended to model the case where controllers stop transmitting upon achieving controllability (as discussed in the next section), continuous transmissions provide ongoing feedback that helps each controller infer interference levels and adapt accordingly. If transmissions cease prematurely, then controllers may underestimate the network density, leading to suboptimal ALOHA tuning. Moreover, assuming that all active controllers transmit throughout the block yields a conservative lower bound on the probability of achieving block controllability.

The rest of this article studies the optimal ALOHA parameter that maximizes the success probability for a typical controller, beginning with a statistical analysis of block controllability. Before proceeding with our statistical analysis, we highlight that a key assumption is that the controllability index of each controlled system is the same, although individual control inputs at each slot may differ. We focus on integrating controllability into channel access, revealing that maximizing transmission success alone may not optimize system performance.

IV. Statistical Analysis of Controllability

This section first characterizes the probability of a successful transmission at a given time for a typical pair under the two ALOHA protocols. Building on the above success probabilities, we then characterize the statistics of the controllability metrics: the burst length for the restless system and the total number of successful transmissions for the rested system.

A. Success Analysis of ALOHA Protocols

We first look at the block ALOHA protocol with a fixed ALOHA parameter q for a given block, starting with the transmission success probability at time t for a typical pair.

Proposition 1 (See [23, Prop. 1]): Consider a network following a given PPP ϕ, whose ith controller is at a distance of r_i from the typical actuator, and $C_i(k) \in \{0,1\}$ indicates if it transmits in block k of T slots. Given that the typical controller transmits in block k, the conditional success probability P_{blk} of the typical controller at time $kT \le t \le (k+1)T-1$ is

$$P_{\text{blk}} = e^{-\frac{\gamma N_0}{\eta \rho r_0^{-\alpha}}} \prod_{i\in\phi:\, C_i(k)=1} \frac{r_0^{-\alpha}}{r_0^{-\alpha}+\gamma r_i^{-\alpha}}$$

where the parameters η, ρ, N_0, and α are defined in (2) and γ is the SINR threshold for successful transmission in (3).

For the block ALOHA case, the set of interfering controllers remains constant throughout the block, so successful transmissions across slots within a block are conditionally independent given the active transmitters. Thus, Proposition 1 accounts for the channel access states C_i. For classical ALOHA, given $C_i(t)$, Proposition 1 holds as well. However, in classical ALOHA, successful transmissions across slots within the same block are independent and identically distributed for a given access probability. Therefore, we average out the randomness in the set of interfering controllers within a block, leading to the following success probability.

Proposition 2: Consider a network following a given PPP realization ϕ, whose ith controller is at a distance of r_i from the typical actuator, and q indicates the ALOHA parameter in a slot of the given block k. Given that the typical controller transmits in block k, the conditional success probability P_{cls} of the typical controller at a given time t within the block is

$$P_{\text{cls}} = e^{-\frac{\gamma N_0}{\eta \rho r_0^{-\alpha}}} \prod_{i\in\phi} \left[q \frac{r_0^{-\alpha}}{r_0^{-\alpha}+\gamma r_i^{-\alpha}} + 1 - q \right]$$

where the parameters η, ρ, N_0, and α are defined in (2) and γ is the SINR threshold for successful transmission in (3).

Proof: From Proposition 1, the success probability is

$$P_{\text{cls}} = \mathbb{E}\left[e^{-\frac{\gamma N_0}{\eta \rho r_0^{-\alpha}}} \prod_{i\in\phi:\, C_i(t)=1} \frac{r_0^{-\alpha}}{r_0^{-\alpha}+\gamma r_i^{-\alpha}} \right].$$

Since each of the $C_i(t)$ is one with probability q and zero with probability $1-q$, the result follows. □

B. Statistical Analysis of Restless System

We first look at the restless system with the block ALOHA protocol, where conditioned on Φ, the success event in Proposition 1 is independent across the time slots when the typical controller transmits. Also, the probability P_{blk} changes across blocks as C_i changes. We first compute the probability of block controllability.

Proposition 3: Consider a network of restless systems following a given realization ϕ of PPP Φ described in Proposition 1. Given that the typical controller transmits in block k, the probability that the system is block controllable is $\bar{F}_{\text{RL,blk}}(v) = \mathbb{P}(L \ge v \mid \Phi)$

$$\bar{F}_{\text{RL,blk}}(v) = \sum_{l=1}^{\lfloor \frac{T+1}{v+1} \rfloor} (-1)^{l+1} \left[P_{\text{blk}} + \frac{T-lv+1}{l}(1-P_{\text{blk}}) \right] \times \binom{T-lv}{l-1} P_{\text{blk}}^{lv} (1-P_{\text{blk}})^{l-1}.$$

Proof: The result follows from the de Moivre's solution [31, Sect. 22.6]. □

The above result establishes the probability of block controllability for a given realization ϕ of PPP Φ with a given value of C_i for the controllers. As mentioned in the previous section, in case the controllers cease transmitting after achieving controllability, the intensity of the interference process remains $q\lambda$ until the first v slots. For slots beyond v, the process further undergoes a second level of thinning, wherein each active transmitter is removed with a probability $\bar{F}_{\mathrm{RL,blk}}(v,t)$ on each slot $\in [v+1,T]$, where

$$\bar{F}_{\mathrm{RL,blk}}(v,t) = \sum_{l=1}^{\lfloor \frac{t+1}{v+1} \rfloor} (-1)^{l+1} \left[P_{\mathrm{blk}} + \frac{t-lv+1}{l}(1-P_{\mathrm{blk}}) \right] \times \binom{t-lv}{l-1} P_{\mathrm{blk}}^{lv} (1-P_{\mathrm{blk}})^{l-1}.$$

This formulation leads to a more complex analysis since the conditional success probability changes in each slot.

Next, we look at the control performance averaged over the network realization using the moments of the conditional success probability.

Theorem 1 (See [23, Thm. 1]): Consider a network of restless systems following a PPP Φ with density λ and block ALOHA parameter q. The probability of block controllability for the typical pair is

$$P_{\mathrm{RL,blk}} = q\mathbb{E}\left[\bar{F}_{\mathrm{RL,blk}}(v)\right] = q \sum_{l=1}^{\lfloor \frac{T+1}{v+1} \rfloor} (-1)^{l+1} \binom{T-lv}{l-1} \times \left(\sum_{\ell=0}^{l-1} \binom{l-1}{\ell} (-1)^{\ell} \zeta(lv+1+\ell) + \frac{T-lv+1}{l} \sum_{\ell=0}^{l} \binom{l}{\ell} (-1)^{\ell} \zeta(lv+\ell) \right).$$

Here, with the parameters η, ρ, N_0, and α in (2), and γ being the SINR threshold for successful transmission in (3), we define

$$\zeta(l) = e^{-\frac{\gamma N_0 l}{\eta \rho r_0^{-\alpha}}} \exp\left(2\pi\lambda q I(l)\right)$$

with the function $I(l)$ given as

$$I(l) = \sum_{\ell=1}^{l} \binom{l}{\ell} \int_0^{\infty} \left(\frac{-\gamma z^{-\alpha}}{r_0^{-\alpha} + \gamma z^{-\alpha}} \right)^{\ell} \mathrm{d}z.$$

Next, we look at the classical ALOHA protocol, where the ALOHA parameter corresponds to the probability of transmitting in each slot. We obtain a result similar to Proposition 3 by replacing P_{blk} with qP_{cls}, as follows.

Theorem 2: Consider a network of restless systems following a PPP Φ with density λ and classical ALOHA parameter q. The probability of block controllability for the typical pair is

$$P_{\mathrm{RL,cls}} = \sum_{l=1}^{\lfloor \frac{T+1}{v+1} \rfloor} (-1)^{l+1} \binom{T-lv}{l-1} \times \left(\sum_{\ell=0}^{l-1} \binom{l-1}{\ell} (-1)^{\ell} \zeta'(lv+1+\ell) + \frac{T-lv+1}{l} \sum_{\ell=0}^{l} \binom{l}{\ell} (-1)^{\ell} \zeta'(lv+\ell) \right).$$

Here, with the parameters η, ρ, N_0, and α in (2), and γ being the SINR threshold for successful transmission in (3), we define

$$\zeta'(l) = \mathbb{E}\left[(qP_{\mathrm{cls}})^l\right] = q^l e^{-\frac{\gamma N_0 l}{\eta \rho r_0^{-\alpha}}} \exp\left(2\pi\lambda q I(l)\right)$$

$$I'(l) = \sum_{\ell=1}^{l} \binom{l}{\ell} \int_0^{\infty} \left(\frac{-q\gamma z^{-\alpha}}{r_0^{-\alpha} + \gamma z^{-\alpha}} + q - 1 \right)^{\ell} \mathrm{d}z.$$

We skip the proof, as it is similar to that of Theorem 1. Further, we can define the meta distribution of the burst length L as the distribution of $\bar{F}_{\mathrm{RL,blk}}(v) = \mathbb{P}(L \geq v \mid \Phi)$ defined in (3) as $\mathcal{M}_{\mathrm{RL}}(v,\beta) = \mathbb{P}(\bar{F}_{\mathrm{RL,blk}}(v) \geq \beta)$. To clarify, $\mathcal{M}_{\mathrm{RL}}(v,\beta)$ represents the fraction of control systems in the network that experience a burst length of at least v in a sequence of T transmissions in at least β fraction of the network realization (or equivalently, due to ergodicity, in at least β fraction of transmissions episodes). Unlike the meta distribution of the SINR in wireless networks, we observe that the meta distribution of the burst length is challenging to derive even indirectly via its moments. However, if all the controllers have the same desired burst length v, then the first moment of the distribution is given by $P_{\mathrm{RL,blk}}(v)$ in Theorem 1.

C. Statistical Analysis of Rested System

For the rested system, the controllability metric is Λ, which is the total number of successes within a block k, which is relatively easy to compute. The complementary cumulative distribution functions (CCDFs) $\bar{F}_{\mathrm{RD,blk}}(v)$ and $\bar{F}_{\mathrm{RD,cls}}(v)$ of Λ for block and classical ALOHA, respectively, are

$$\bar{F}_{\mathrm{RD,blk}}(v) = \mathbb{P}(\Lambda \geq v | \Phi) = q \sum_{l=v}^{T} \binom{T}{l} (P_{\mathrm{blk}})^l (1-P_{\mathrm{blk}})^{T-l}$$

$$\bar{F}_{\mathrm{RD,cls}}(v) = \mathbb{P}(\Lambda \geq v | \Phi) = \sum_{l=v}^{T} \binom{T}{l} (qP_{\mathrm{cls}})^l (1-qP_{\mathrm{cls}})^{T-l}.$$

Then, we can determine the probability that the typical rested control system is block controllable in at least a specified fraction of network realizations. This probability is denoted as the meta distribution of the number of successes Λ, as discussed below. We start with the block ALOHA protocol.

Theorem 3: Consider a network of rested systems following a PPP Φ with density λ and block ALOHA parameter q. The typical controlled system is block controllable in at least β fraction of the network realizations with probability

$$\mathcal{M}_{\mathrm{RD,blk}}(v,\beta) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\Im\left(e^{-js\log(P_{\mathrm{blk}}(q))} \zeta_{\mathcal{I}}(s)\right)}{s} \mathrm{d}s$$

where $\Im(\cdot)$ represents the imaginary part of the argument and

$$P_{\mathrm{blk}}(q) = \min\left\{ p \in [0,1] : \; q \sum_{l=v}^{T} \binom{T}{l} p^l (1-p)^{T-l} \geq \beta \right\}$$

$$\zeta_{\mathcal{I}}(s) = e^{-\frac{js\gamma N_0}{\eta \rho r_0^{-\alpha}}} \exp\left[-2\pi q\lambda \int_0^{\infty} 1 - \left(\frac{r_0^{-\alpha}}{r_0^{-\alpha} + \gamma z^{-\alpha}} \right)^{js} \mathrm{d}z \right].$$

Proof: See Appendix A. □

The next theorem characterizes classical ALOHA.

Theorem 4: Consider a network of restless systems following a PPP Φ with density λ and classical ALOHA parameter q. The probability that the typical pair is block controllable in at least β fraction of the network realizations is

$$\mathcal{M}_{\mathrm{RD,cls}}(v,\beta)=\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\frac{\Im\left(e^{-js\log(P_{\mathrm{cls}}(q))}\zeta'_{\mathcal{I}}(s)\right)}{s}\mathrm{d}s$$

where we define

$$P_{\mathrm{cls}}(q)=\min\left\{p\in[0,1]:\sum_{l=v}^{T}\binom{T}{l}(qp)^l(1-qp)^{T-l}\geq\beta\right\}$$

$$\zeta'_{\mathcal{I}}(s)=e^{-\frac{js\gamma N_0}{\eta\rho r_0^{-\alpha}}}\times\exp\left[-2\pi\lambda\int_0^{\infty}1-\left(q\frac{r_0^{-\alpha}}{r_0^{-\alpha}+\gamma z^{-\alpha}}+1-q\right)^{js}\mathrm{d}z\right].$$

In the rested controlled system, unlike the restless case, the meta distribution of the conditional success probability directly characterizes controllability. However, analytically evaluating the integrals in $\zeta_{\mathcal{I}}(s)$ and $\zeta'_{\mathcal{I}}(s)$ from the Gil–Pelaez inversion theorem is generally intractable, and the numerical approximation is often computationally expensive. Recently, the Chebyshev–Markov method was introduced to reconstruct the meta distribution based on a finite sequence of moments [32]. Since the moments of P_{blk} and P_{cls} have already been calculated as an intermediate step in the proof of Theorems 3 and 4, respectively, the Chebyshev–Markov method can be formulated as an Hausdorff moment problem (HMP) [33]. This formulation utilizes the property that if an infinite sequence of moments is monotonic, then the distribution of a random variable exists and is unique. While the exact distribution is based on infinite moments, using finite moments results in an HMP reconstruction. For more details, we refer the reader to [32].

This concludes our statistical analysis averaged over PPP. However, for a given realization, the optimal ALOHA parameter depends on the locations of the controllers, which is impractical for each controller to know globally. Therefore, the controllers can only learn the optimum value of the ALOHA parameter through the acknowledgments sent by the actuators. In the next section, we discuss a TS-based learning algorithm to optimize the ALOHA parameter selection.

V. TS-Based ALOHA Parameter Selection

This section presents a centralized online channel access learning policy based on statistical analysis. A central decision-maker broadcasts the block parameter $q(k)\in\mathcal{P}$ to all the controllers in the network at the beginning of the block k. Then, the individual controllers set their channel access states $C_i(t)$ probabilistically based on this parameter. In block ALOHA, $C_i(t)$ remains constant, while in classical ALOHA, $C_i(t)$ varies across the slots. The central decision-maker receives the T acknowledgments for each transmitting (active) controller and updates the channel access probability centrally. The goal of the decision-maker is to sequentially select channel access probabilities $q^{(K)}=\{q(1),q(2),\ldots,q(K)\}$ to maximize the probability of block controllability of the typical pair, where K is the number of blocks.

A. Multiarm Bandit Formulation for the Restless System

We formulate the decision-maker's task of sequentially selecting channel access probabilities across blocks as a MAB problem. Here, the set of D access probabilities $\mathcal{P}:=\{p_1,p_2,\ldots,p_D\}$ represents the arms or actions. We observe the corresponding block controllability for every choice of arm, and the probability of block controllability can be learned over different blocks. Although the ALOHA parameter developed in the previous section is theoretically continuous, in this section, we select it from a discrete set for online implementation. This choice reflects practical constraints as the implementation of wireless parameters is governed by standardized protocols that impose discrete selections [34].

We next define the reward of MAB to reflect block controllability. A naive approach is to assign a reward of one if the typical pair is block controllable, and zero otherwise. However, in block ALOHA, controllability depends on consecutive successful transmissions, while in classical ALOHA, it depends on the total number of successful transmissions. Since the probability of consecutive successes increases with the total number of successes, and the rewards are independent and identically distributed across slots given ϕ and $q(k)$, the reward for both protocols can be based on the number of successful transmissions within a block. Consequently, an alternative reward formulation can set the reward to one if the transmission is successful, i.e., $S(t)=1$, and zero otherwise. As the formulation has multiple rewards per block, it leads to faster learning than the naive scheme.

To complete the MAB formulation, we need to define the regret. For a given choice of ALOHA parameters $q^{(K)}$ and the optimal ALOHA parameter q_*, we define the Bayesian regret over K blocks averaged over the realizations of Φ and the reward sample path $S(t)$ of the algorithm as

$$\begin{aligned}\mathcal{R}(q^{(K)})&=\mathbb{E}_{\Phi,S(t)}\left[\sum_{k=0}^{K-1}\sum_{t=kT}^{(k+1)T-1}S(t)\middle|q(k)=q_*\right]\\&\quad-\mathbb{E}_{\Phi,S(t)}\left[\sum_{k=0}^{K-1}\sum_{t=kT}^{(k+1)T-1}S(t)\middle|q(k)=q_k\right]\\&=\mathbb{E}_{\Phi}\left[\sum_{k=0}^{K-1}T(q_*\bar{S}(q_*)-q_k\bar{S}(q_k))\right]\end{aligned}\tag{8}$$

where $\bar{S}(q)=\mathbb{E}(S(t)|q(k)=q)$, for $t=kT,kT+1,\ldots,(k+1)T-1$. Naturally, due to the randomness of Φ (technically, the bandit environment), the optimal ALOHA parameter q_* is a random variable depending on Φ. We next study a Bayesian approach for the MAB problem, based on TS.

B. Block TS for the Restless System

In the TS framework, the decision-maker chooses the ALOHA parameter based on their belief for the expected reward

Algorithm 3: TS for Block ALOHA Parameter Selection.

Parameters: Beta distribution parameters $\{a_d, b_d\}_{d=1}^{D}$
Initialization: $a_d = b_d = 1$, **for** $all d \in \{1, 2, \dots D\}$
for $k = 0, 1, 2, \dots, K$ **do**
  Sample $\theta_d \sim \mathcal{B}(a_d, b_d)$ **for** $all d \in \{1, 2, \dots D\}$
  Set parameter $q(k) = p_{d^*}$, where $d^* = \arg\max_d \theta_d$
  Observe acknowledgments $S(kT + \tau)$ for $\tau = 0, 1, \dots, T-1$
  Update $a_{d^*} \leftarrow a_{d^*} + \sum_{\tau=0}^{T-1} S(kT + \tau)$
  Update $b_{d^*} \leftarrow b_{d^*} + T - \sum_{\tau=0}^{T-1} S(kT + \tau)$.
else for

with each access probability in $\mathcal{P}$. Specifically, the decision-maker starts with a prior probability θ_d of obtaining a reward of one when the ALOHA parameter $p_d \in \mathcal{P}$ is chosen, and a reward of zero with probability $1 - \theta_d$ [35]. Each θ_k is an action's success probability or mean reward. In the Bayesian model, θ_d is modeled as a Beta distribution (conjugate prior for the Bernoulli rewards) with parameters a_d and b_d. We initialize both parameters as $a_d = b_d = 1$, which corresponds to a uniform prior distribution, since we do not assume any prior knowledge of the interference conditions at the typical transmitter. Then, for a selected ALOHA parameter $q(k) = p_d$ in a block k, if the typical controller is active, then it experiences a reward of $S(t)$ for a given t in the block k. Based on this reward, the posterior distribution of the chosen access probability p_d is updated using the Bayes' rule [35]. The Bernoulli rewards facilitate a simple update rule: the parameter a_d is incremented by 1 if $S(t) = 1$, while the parameter b_d is decreased by one if $S(t) = 0$. Nonetheless, the ALOHA parameter does not change until the next block begins. Therefore, we can do a batch update of the parameters at the end of a block. Particularly, at the end of block k, the parameter a_d is incremented by $\sum_{t=kT}^{(k+1)T-1} S(t)$, while the parameter b_d is incremented by $T - \sum_{t=kT}^{(k+1)T-1} S(t)$. Since the typical pair is randomly selected from the distribution of the pairs in the network, the central transmitter can either randomize the pair's selection for observation or consider the average successes across all the pairs in the network. The overall algorithm is summarized in Algorithm 3.

Proposition 4: The Bayesian regret of the block TS algorithm in Algorithm 3 after K blocks, defined in (8), is bounded as $\mathcal{R}_{\mathrm{TS}}(K) \leq \mathcal{O}(\sqrt{TKD\log(K)})$, where T is the number of time slots per block and D is the number of choices for access probabilities.

Proof: See Appendix B. □

The result aligns with the classical TS, highlighting the dependence of regret on the parameters K, T, and D.

VI. Numerical Results and Discussion

In this section, we assume a transmit power of $\eta = 24$ dBm and a path-loss exponent of $\alpha = 2$ considering line-of-sight propagation. The carrier frequency is assumed to be 3.2 GHz with an operating bandwidth of 200 MHz. The distance between the typical controller–controlled system pair is set as 10 m.

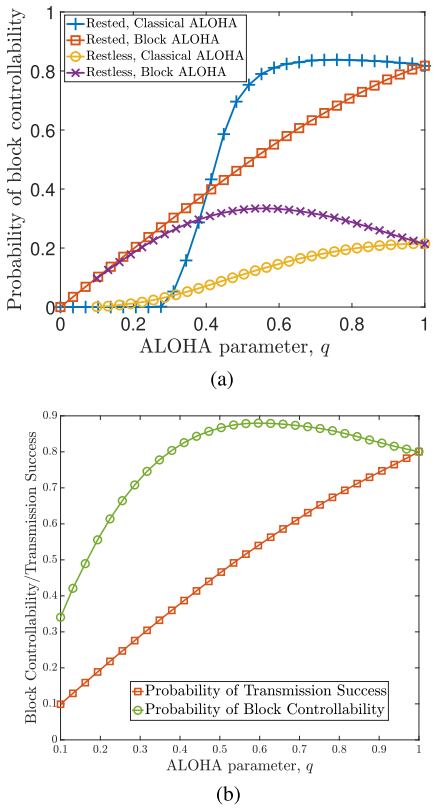


Fig. 2. Comparisons of different systems and success metrics. (a) $v = 4$, $T = 20$, and $\lambda = 5e-3\mathrm{m}^{-2}$. (b) $v = 6$, $T = 20$, and $\lambda = 1e-3\mathrm{m}^{-2}$.

We empirically study the effect of classical and block ALOHA parameters on the probability of block controllability for the rested and restless systems. Fig. 2(a) shows the probability of block controllability for the rested and restless systems for a single block ($K = 1$) averaged across 10 000 realizations of PPP. With classical ALOHA, both the rested and restless systems have a lower probability of controllability when the ALOHA parameter q is low ($q \leq 0.3$), as a lower channel access frequency limits controllability performance. As q increases, controllability improves but declines at high q due to interference, reducing successful transmission probability. The block ALOHA also has a similar trend. However, the increase in the probability of block controllability here is sharper as compared to the classical ALOHA, as a single successful access guarantees channel access across all block slots. We note that although for the selected values of the system parameters, the controllability of the rested system with block ALOHA channel access shows a monotonic increase, for a denser deployment of control systems (higher λ), a high q may lead to a decrease in the probability of block controllability. At $q = 1$, both protocols yield similar performance.

For the restless system, block ALOHA outperforms classical ALOHA across all q values. This observation is intuitive, as block ALOHA enables consecutive channel access, which increases the likelihood of successive successful transmissions, thereby leading to a higher probability of achieving block controllability in the restless system. Conversely, for the rested

system, the choice between block and classical ALOHA is nontrivial. For example, in Fig. 2(a), we observe that for the rested system with classical ALOHA (blue plot), the probability of block controllability increases with q, reaching a peak above 0.8 at $q \approx 0.75$ (shown with the blue circle). Then, it decreases to 0.8 at $q = 1$, which is the maximum value of the probability of block controllability for the block ALOHA scheme. Notably, classical ALOHA introduces randomness by varying the set of transmitting controllers in each slot, whereas block ALOHA fixes the set of transmitting controllers for the entire block. So, for different values of q, the expectation of the block controllability over the transmitting set may be greater or lower than the block controllability for an expected transmission set. Moreover, rested systems generally exhibit a higher probability of block controllability at any given q value when using the same ALOHA protocol, as they are more flexible and incorporate an additional feedback loop.

In Fig. 2(b), we plot the transmission success metric, which is the product of the transmission success probability in one slot and the channel access probability, i.e., qP_{cls}, and compare it with the controllability metric, which is the probability of block controllability in at least one block given that the controller attempts to access in a sequence of $K = 5$ blocks, i.e., $1 - (1 - qP_{\mathrm{blk}})^5$. With a low interferer density ($\lambda = 10^{-4}\mathrm{m}^{-2}$), the transmission success probability qP_{cls} increases monotonically with q, peaking at $q = 1$. This observation suggests that from a communication-only perspective, full channel access ($q = 1$) is optimal. However, from a control-co-design perspective, block controllability is maximized at a lower q, revealing a tradeoff.

Further, recall that in the block ALOHA access scheme, once a controller gets access in the block, it is allowed to transmit in all the slots of that block without any further access contention. This permits reduced channel access probability, particularly advantageous for long blocks, because even a single access may suffice over the entire time horizon. This decision, however, is tightly linked to the number and length of control blocks, which are functions of the controllability parameter v. In contrast, classical (slot) ALOHA enforces per-slot competition, which may enhance transmission success rates but is misaligned with block-controllability goals. We next study the impact of the block and classical ALOHA parameters on block controllability for a large value of $K = 5000$. Here, unlike the empirical study in Fig. 2(a), we use the theoretical expressions derived in Theorems 1–4. Since the analysis of block and classical ALOHA are similar, we focus on illustrating the impact of block ALOHA parameter on the block controllability in the restless system and the impact of classical ALOHA parameter on block controllability in the rested system. These results are summarized in Figs. 3 and 4.

Fig. 3 shows the theoretical expression for block controllability probability $P_{\mathrm{RL,blk}}$ of a restless system in Theorem 1 for two different network densities with $\lambda = 10^{-4}$ m^{-2} and $\lambda = 5 \times 10^{-4}$ m^{-2}. As in Fig. 2(a), low channel access probability q results in low $P_{\mathrm{RL,blk}}$ despite low interference. Also, $P_{\mathrm{RL,blk}}$ increases with q until a threshold is reached, beyond which higher q leads to increased interference, degrading success probability despite more frequent channel access. Further, higher density of the network λ leads to lower $P_{\mathrm{RL,blk}}$ due to an

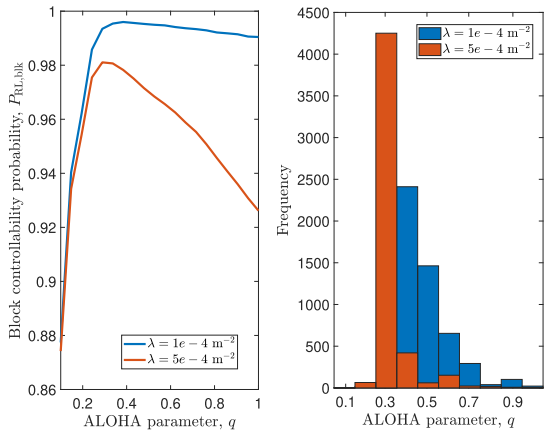


Fig. 3. Block controllability of restless system and parameter selection by block TS with $v = 5$, $T = 20$, and $K = 5000$.

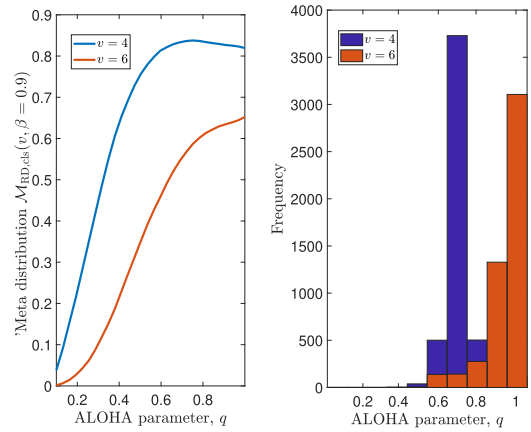


Fig. 4. Meta distribution of rested system with classical ALOHA and the parameter selection by TS with $K = 5000$.

increase in interference and steep drop-off with q, as the intensity of interfering controllers is $q\lambda$. Furthermore, the optimal q also shifts with network density: a high value of approximately 0.35 for a low value of $\lambda = 10^{-4}$ m^{-2} and a lower value of 0.25 when $\lambda = 5 \times 10^{-4}$ m^{-2}.

The right panel of Fig. 3 shows TS performance with $D = 10$ and $\mathcal{P} = \{0.1, 0.2, \ldots, 1\}$. TS quickly identifies the optimal ALOHA parameters, 0.4 and 0.3, for $\lambda = 10^{-4}$ m^{-2} and 5×10^{-4} m^{-2}, respectively, as shown in the left panel of Fig. 3. In particular, with $\lambda = 5 \times 10^{-4}$ m^{-2}, the optimum ALOHA parameter is selected in more than 4000 blocks over $K = 5000$ blocks. On the contrary, for $\lambda = 10^{-4}$ m^{-2}, performance is slightly lower due to smaller differences between rewards of different choices.

Fig. 4 shows the meta distribution $\mathcal{M}_{\mathrm{RD,cls}}(v, \beta)$ of the rested system with classical ALOHA in Theorem 4 for two different controllability indices $v = 4$ and 6 with a reliability threshold of $\beta = 0.9$. The higher value of v makes the system more demanding, resulting in a lower meta distribution. For $v = 6$, the optimal ALOHA parameter is 1, prioritizing channel access over increased interference. The optimal parameter decreases as q decreases, indicating that the rested system should reduce its

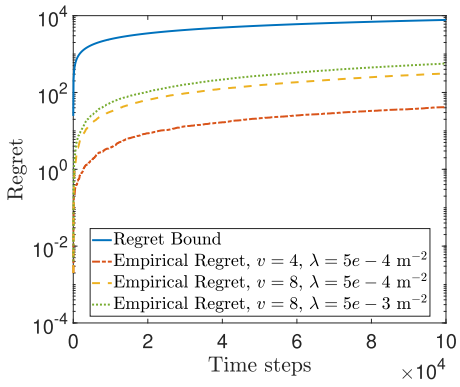


Fig. 5. Regret for the restless system.

channel access probability for lower v to achieve the required successful transmissions while minimizing interference. The right panel of Fig. 4 shows the performance of the classical TS algorithm in determining the optimal classical ALOHA parameter with respect to different v. Here, we consider blocks of 20 slots and observe the system over 250 blocks. Similar to the restless case, we consider $\mathcal{P} = \{0.1, 0.2, \ldots, 1\}$. For $v = 4$, our analytical framework predicts that the optimal access probability is $q = 0.7$, while for $v = 6$, the optimum scheme is to transmit in all slots. This is reflected in the frequency with which the TS framework selects the channel access probability. However, the performance of the rested system using classical TS is inferior to that of block TS for the restless system, indicating that block TS consistently outperforms classical TS for all systems and both protocols.

Finally, we illustrate the regret of TS for the restless system, defined in Fig. 5. Higher λ and v lead to higher regret due to intensified interference and decreased controllability, respectively. Nonetheless, the regret evolves sublinearly in both cases, aligning with the bound in Proposition 4.

VII. Conclusion

In this article, we designed a control strategy and communication protocol for two Poisson networks, restless and rested, and analyzed control performance under interference. To mitigate interference, we introduced block and classical ALOHA protocols for channel access. Our study identified optimal ALOHA parameters based on system characteristics, such as controller density. A TS-based algorithm was developed for online parameter tuning, achieving sublinear regret. Future directions include decentralized parameter selection and its analysis.

Appendix A
Proof of Theorem 3

From the definition of $\mathcal{M}_{\mathrm{RD,blk}}(v,\beta)$, we have

$$\mathcal{M}_{\mathrm{RD,blk}}(v,\beta) = \mathbb{P}\left(q\sum_{l=v}^{T}\binom{T}{l}P_{\mathrm{blk}}^{l}(1-P_{\mathrm{blk}})^{T-l} \geq \beta\right)$$
$$= \mathbb{P}\left(P_{\mathrm{blk}} \geq P_{\mathrm{blk}}(q)\right)$$

which is the CCDF of P_{blk}. Using the Gil–Pelaez theorem [36]

$$\mathcal{M}_{\mathrm{RD,blk}}(v,\beta) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\Im\left(e^{-js\log(\beta)}\mathbb{E}\left[P_{\mathrm{blk}}^{js}\right]\right)}{s}\mathrm{d}s.$$

Now, using Proposition 1, the moments $\zeta_{\mathcal{I}}(s) = \mathbb{E}[P_{\mathrm{blk}}^{js}] = e^{-\frac{js\gamma N_0}{\eta\rho r_0^{-\alpha}}}\mathbb{E}[\prod_{i\in\phi}(\frac{r_0^{-\alpha}}{r_0^{-\alpha}+\gamma r_i^{-\alpha}})^{js}]$, and the result follows as the transmitting controllers form a PPP with density $q\lambda$.

Appendix B
Proof of Proposition 4

The proof is based on the analysis of TS for finite-armed bandits [37, Ch. 36]. We define $\mu_* = q_* T\bar{S}(q_*)$ as the average reward in a block with the optimal channel access probability q_*, and $\mu(p_d) = p_d T\bar{S}(p_d)$ as the average reward obtained when the channel access probability is p_d, for $d = 1, 2, \ldots, D$. Then, from (8), the reward is

$$\mathcal{R}(K) = \mathbb{E}_\Phi\left[\sum_{k=0}^{K-1}\mathbb{E}\left(\mu_* - \mu(q(k))\right)\right]$$

where $q(k)$ is the channel access probability in block k. Now, by the law of total expectation, for any random event $\mathcal{E}$

$$\mathcal{R}(K) = \mathbb{E}_\Phi\left[\mathbb{P}(\mathcal{E})\mathbb{E}\left[\sum_{k=0}^{K-1}\mathbb{E}\left(\mu_* - \mu(q(k))|\mathcal{E}\right)\right]\right]$$
$$+ \mathbb{E}_\Phi\left[\mathbb{P}(\mathcal{E}^{\complement})\mathbb{E}\left[\sum_{k=0}^{K-1}\mathbb{E}\left(\mu_* - \mu(q(k))|\mathcal{E}^{\complement}\right)\right]\right]$$
$$\leq \mathcal{R}_1(K) + KT\mathbb{P}(\mathcal{E}^{\complement}) \quad (9)$$

where $\mathcal{R}_1(K) = \mathbb{E}_\Phi[\sum_{k=0}^{K-1}\mathbb{E}(\mu_* - \mu(q(k))|\mathcal{E})]$ and we use the property that $\mu_* - \mu(q(k)) \leq \mu_* \leq T$.

We next define the event $\mathcal{E}$ using the empirical estimate of the reward computed by the TS algorithm. Let the empirical estimate of the reward corresponding to the access probability p_d after the kth block be

$$\hat{\mu}_d(k) = \frac{1}{\Gamma(k,d)}\sum_{\kappa=0}^{k-1}\mathbb{1}(q(k) = p_d)p_d\sum_{t=\kappa T}^{(\kappa+1)T-1}S(t)$$

where $\Gamma(k,d)$ denotes the number of blocks until block k in which channel access probability is chosen as p_d. If $\Gamma(k,d) = 0$, then we define $\hat{\mu}_d(k) = 0$. Here, given the channel access probability sequence $q^{(K)} = \{q(1), q(2), \ldots, q(K)\}$, the Bernoulli random variable $S(t)$ has mean $P_{\mathrm{blk}}(p_d)$ defined in Proposition 1 with $q = p_d$. Hence, $\mathbb{E}(\hat{\mu}_d(k)|q^{(K)}) = \mu(p_d)$. We define $\mathcal{E} = \cap_{k=0}^{K-1}\cap_{d=1}^{D}\mathcal{E}(k,d)$, where the event $\mathcal{E}(k,d)$ is that for a given $k = 0, 1, \ldots, K-1$ and $d = 1, 2 \ldots, D$

$$|\hat{\mu}_d(k-1) - \mu(p_d)| < \epsilon(k,d) = \sqrt{\frac{2\log(1/\delta)}{\max\{1, T\Gamma(k,d)\}}}$$

for some $0 < \delta < 1$. Now, to compute the bound in (9), we first compute $\mathbb{P}(\mathcal{E}^{\complement})$ using the union bound as follows:

$$\mathbb{P}(\mathcal{E}^{\complement}) \leq \sum_{k=0}^{K-1}\sum_{d=1}^{D}\mathbb{P}\left(\mathcal{E}(k,d)^{\complement}\right) \leq 2KD\delta$$

where we also use Hoeffding's inequality. So, (9) implies

$$\mathcal{R}(K) \leq \mathcal{R}_1(K) + 4K^2TD\delta. \tag{10}$$

Now, we bound the term $\mathcal{R}_1(K)$ in (10) by defining the σ-algebra generated by the ALOHA parameters and the corresponding rewards by the end of the kth slot as $\mathcal{F}_k = \sigma(q^{(k)}, S(0), S(1), \ldots, S((k+1)T-1))$. Then, we have

$$\mathcal{R}_1(K) = \mathbb{E}_\Phi\left[\sum_{k=0}^{K-1} \mathbb{E}\left(\mu_* - \mu(q(k))|\mathcal{F}_{k-1}, \mathcal{E}\right)\right].$$

However, the TS algorithm implies that the conditional distribution of q_* is the same as $q(k)$ [37], i.e., $\mathbb{E}[\nu(\mu_*)|\mathcal{F}_{k-1}] = \mathbb{E}[\nu(\mu(q(k)))|\mathcal{F}_{k-1}]$, for any function ν. Consequently

$$\mathcal{R}_1(K) = \mathbb{E}_\Phi\left[\sum_{k=0}^{K-1} \mu_* - \nu(\mu_*) + \nu(\mu(q(k))) - \mu(q(k))|\mathcal{E}\right]$$

using the law of total expectation. Furthermore, we choose

$$\nu(\hat{\mu}_d(k-1)) = \min\left\{1, \max\left\{0, \hat{\mu}_d(k-1) + \epsilon(k,d)\right\}\right\}.$$

Therefore, under $\mathcal{E}$, we have $\mu_* < \nu(\mu_*)$, leading to

$$\begin{aligned}
\mathcal{R}_1(K) &\leq \mathbb{E}_\Phi\left[\sum_{k=0}^{K-1} \nu(\mu(q(k))) - \mu(q(k))|\mathcal{E}\right] \\
&\leq \mathbb{E}_\Phi\left[\sum_{k=0}^{K-1}\sum_{d=1}^{D} \mathbb{1}(q(k) = p_d) 2\epsilon(k,d)\right] \\
&= \mathbb{E}_\Phi\left[\sum_{d=1}^{D} \int_0^{\Gamma(K-1,d)} 2\sqrt{\frac{2\log(1/\delta)}{z}}\mathrm{d}z\right] \\
&= \mathbb{E}_\Phi\left[\sum_{d=1}^{D} \sqrt{32\Gamma(K-1,d)\log(1/\delta)}\right].
\end{aligned}$$

So, we get $\mathcal{R}_1(K) \leq \sqrt{32KD\log(1/\delta)}$. Hence, with $\delta = 1/\sqrt{K}$ and using (10), we $\mathcal{R}(K) \leq \sqrt{64KD\log(K)} + 4TD$, arriving at the desired result.

REFERENCES

[1] L. Ding, Q.-L. Han, L. Y. Wang, and E. Sindi, "Distributed cooperative optimal control of DC microgrids with communication delays," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 3924–3935, Sep. 2018.

[2] Z. Lu and G. Guo, "Control and communication scheduling co-design for networked control systems: A survey," *Int. J. Syst. Sci.*, vol. 54, no. 1, pp. 189–203, 2023.

[3] J. Hu, Z. Wang, G.-P. Liu, H. Zhang, and R. Navaratne, "A prediction-based approach to distributed filtering with missing measurements and communication delays through sensor networks," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 51, no. 11, pp. 7063–7074, Nov. 2021.

[4] B. Ning, Q.-L. Han, and L. Ding, "Distributed finite-time secondary frequency and voltage control for islanded microgrids with communication delays and switching topologies," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 3988–3999, Aug. 2021.

[5] D. Baumann, F. Mager, M. Zimmerling, and S. Trimpe, "Control-guided communication: Efficient resource arbitration and allocation in multi-hop wireless control systems," *IEEE Control Syst. Lett.*, vol. 4, no. 1, pp. 127–132, Jan. 2020.

[6] 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.

[7] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *IEEE Wirel. Commun.*, vol. 12, no. 1, pp. 3–11, Feb. 2005.

[8] M. S. Mahmoud, *Control and Estimation Methods Over Communication Networks*. New York, NY, USA: Springer, 2014.

[9] 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.

[10] D. Zhang, Q.-L. Han, and X. Jia, "Network-based output tracking control for a class of TS fuzzy systems that can not be stabilized by nondelayed output feedback controllers," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1511–1524, Aug. 2015.

[11] Y. Qiao, Y. Fu, and M. Yuan, "Communication-control co-design in wireless networks: A cloud control AGV example," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2346–2359, Feb. 2023.

[12] S. Hu and W.-Y. Yan, "Stability robustness of networked control systems with respect to packet loss," *Automatica*, vol. 43, no. 7, pp. 1243–1248, 2007.

[13] X. Cao, P. Cheng, J. Chen, and Y. Sun, "An online optimization approach for control and communication codesign in networked cyber-physical systems," *IEEE Trans. Ind. Inform.*, vol. 9, no. 1, pp. 439–450, Feb. 2013.

[14] A. M. Girgis, J. Park, M. Bennis, and M. Debbah, "Predictive control and communication co-design via two-way Gaussian process regression and AoI-aware scheduling," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7077–7093, Oct. 2021.

[15] X. Wang, C. Chen, J. He, S. Zhu, and X. Guan, "AoI-aware control and communication co-design for Industrial IoT systems," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8464–8473, May 2021.

[16] L. Zhang and D. Hristu-Varsakelis, "Communication and control co-design for networked control systems," *Automatica*, vol. 42, no. 6, pp. 953–958, 2006.

[17] B. Chang, G. Zhao, Z. Chen, L. Li, and M. A. Imran, "Packet-drop design in URLLC for real-time wireless control systems," *IEEE Access*, vol. 7, pp. 183081–1830 90, 2019.

[18] M. Eisen, M. M. Rashid, K. Gatsis, D. Cavalcanti, N. Himayat, and A. Ribeiro, "Control aware radio resource allocation in low latency wireless control systems," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7878–7890, Oct. 2019.

[19] K. Gatsis, A. Ribeiro, and G. J. Pappas, "State-based communication design for wireless control systems," in *Proc. IEEE Conf. Decis. Control*, 2016, pp. 129–134.

[20] L. Scheuvens, T. Hößler, A. N. Barreto, and G. P. Fettweis, "Wireless control communications co-design via application-adaptive resource management," in *Proc. IEEE 5G World Forum*, 2019, pp. 298–303.

[21] P. Park, J. Araújo, and K. H. Johansson, "Wireless networked control system co-design," in *Proc. Int. Conf. Netw. Sens. Control*, 2011, pp. 486–491.

[22] B. Chang, L. Zhang, L. Li, G. Zhao, and Z. Chen, "Optimizing resource allocation in URLLC for real-time wireless control systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8916–8927, Sep. 2019.

[23] G. Ghatak, G. Joseph, and C. Quan, "Poisson networked control systems: Statistical analysis and online learning for channel access," in *Proc. WiOpt Workshop RAWNET*, 2024, pp. 62–69.

[24] M. Haenggi, "The meta distribution of the SIR in Poisson bipolar and cellular networks," *IEEE Trans. Wirel. Commun*, vol. 15, no. 4, pp. 2577–2589, Apr. 2016.

[25] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*. Hoboken, NJ, USA: Wiley, 2013.

[26] *3GPP*, "Technical specification group radio access network; study on channel model for frequencies from 0.5 to 100 GHz (release 16), document TR," ETSI, Sophia Antipolis, France, Tech. Rep. 3GPP TR 38.901, version 16.0.0., Oct. 2019.

[27] J. G. Andrews, A. K. Gupta, and H. S. Dhillon, "A primer on cellular network analysis using stochastic geometry," 2016, *arXiv:1604.03183*.

[28] X. Lu, M. Salehi, M. Haenggi, E. Hossain, and H. Jiang, "Stochastic geometry analysis of spatial-temporal performance in wireless networks: A tutorial," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2753–2801, Fourthquarter 2021.

[29] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 421–436, Feb. 2006.

[30] R. T. Ma, V. Misra, and D. Rubenstein, "An analysis of generalized slotted-Aloha protocols," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 936–949, Jun. 2009.

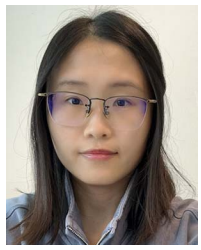
[31] A. Hald, *A History of Probability and Statistics and Their Applications Before 1750*. Hoboken, NJ, USA: Wiley, 2005.
[32] X. Wang and M. Haenggi, "Fast Hausdorff moment transforms for meta distributions in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 2607–2621, Apr. 2024.
[33] R. M. Mnatsakanov, "Hausdorff moment problem: Reconstruction of probability density functions," *Statist. Probab. Lett.*, vol. 78, no. 13, pp. 1869–1877, 2008.
[34] 3rd Generation Partnership Project (3GPP), "5G; NR; physical layer procedures for control," ETSI, Sophia Antipolis, France, Tech. Rep. 3GPP TS 38.213, Version 16.2.0, Release 16, Jul. 2020.
[35] D. J. Russo et al., "A tutorial on Thompson sampling," *Found. Trends Mach. Learn.*, vol. 11, no. 1, pp. 1–96, 2018.
[36] J. Gil-Pelaez, "Note on the inversion theorem," *Biometrika*, vol. 38, no. 3-4, pp. 481–482, 1951.
[37] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2020.

Gourab Ghatak (Member, IEEE) received the Ph.D. degree in wireless communications from Telecom Paris Tech (University of Paris Saclay), Paris, France, in 2019.

He is currently an Assistant Professor with the Department of Electrical Engineering, Indian Institute of Technology Delhi (IIT Delhi), Delhi, India. His research interests include stochastic geometry, MACPHY cross-layer issues in beyond 5G systems, and machine learning for wireless communications.

Geethu Joseph (Senior Member, IEEE) received the Ph.D. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 2019.

She is currently an Assistant Professor with the Signal Processing Systems Group, Delft University of Technology, Delft, The Netherlands. Her research interests include statistical signal processing and network control.



Chen Quan (Member, IEEE) received the Ph.D. degree in electrical engineering from Syracuse University, Syracuse, NY, USA, in 2023.

She is currently a Postdoctoral Researcher with the Signal Processing Systems Group, Delft University of Technology, Delft, The Netherlands. Her research interests include signal processing and optimization.