TRANSMISSION SCHEDULING IN WIRELESS NETWORKED CONTROL FOR INDUSTRIAL IOT

By

Kang Huang

SUBMITTED IN FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
CENTRE OF EXCELLENCE IN TELECOMMUNICATIONS
SCHOOL OF ELECTRICAL AND INFORMATION ENGINEERING
THE UNIVERSITY OF SYDNEY

JULY 2021

To my beloved parents Yunzao Huang, Xiuping Zhang

Acknowledgements

I've always believed that choosing to become a Ph.D. candidate at the University of Sydney is one of the biggest decisions I've made in my life. I have learned a lot from this experience and made lots of unforgettable memories during the period of research. Here I would like to express my gratitude to those people who were involved in my Ph.D. study.

Firstly, I would like to thank Prof. Yonghui Li. He is a respected and prestigious scholar. He guided me throughout my Ph.D. study. I would also like to thank Prof. Branka Vucetic. Her kindness and passion have inspired me to conduct research.

Secondly, I would like to thank Dr. Wanchun Liu. I could never complete the thesis without her support. I also learned a lot from her professional insights and positive attitude.

Specially, I would like to thank Na Zhao, my girlfriend. You gave me the courage to overcome difficulties and stayed by my side from my undergraduate degree to my Ph.D.

Thanks to everyone in the Telecommunications Lab at the University of Sydney. I experienced many meaningful and happy moments with you during the past years.

Finally, last but by no means least, to my beloved parents who have unconditionally supported me for the past twenty-five years.

Kang Huang

Sydney, NSW

April 2021

Statement of Originality

The work presented in this thesis is the result of original research carried out by myself, in collaboration with my supervisors, while enrolled in the School of Electrical and Information Engineering at the University of Sydney as a Ph.D. candidate.

These studies were conducted under the supervision of Dr. Wanchun Liu, and Prof. Yonghui Li. It has not been submitted for any other degree or award in any other university or educational institution.

Kang Huang
School of Electrical and Information Engineering
The University of Sydney
April 2021

Abstract

Wireless networked control systems (WNCS) consist of spatially distributed sensors, actuators, and controllers communicating through wireless networks. WNCS has recently emerged as a fundamental infrastructure technology to enable reliable control for mission-critical Industrial Internet of Things (IIoT) applications such as factory automation, intelligent transportation systems, tele-medicine and smart grids. The design of WNCS requires the joint design of communications, computing and control. For a large scale deployment, WNCS faces challenges such as unreliable transmission and latency in transmitting control and sensing information due to channel impairment in wireless communications. This can have a significant impact on the stability and performance of WNCS. Most existing works have mainly focused on the design of WNCS from a control perspective rather than communications, or have considered an ideal or simplified wireless model. How to reliably control WNCS in practical wireless channels and design wireless communication scheduling policy to optimize the control performance is a challenging task. This thesis presents the design of practical communication protocols of a general discrete linear time-invariant (LTI) dynamic system in WNCS. We address the transmission scheduling problems in WNCS in three scenarios, which require the development of different strategies. Firstly, to minimize the long-term average remote estimation mean-squared-error (MSE), a hybrid automatic repeat request (HAQR)-based real time estimation framework is proposed. Secondly, a downlink-uplink transmission scheduling policy is developed for half-duplex (FD) controller to optimize the system performance. Finally, a novel controller with adaptive packet length is studied and a variable-length packet-transmission policy is proposed to optimally balance the delay-reliablity tradeoff in WNCS. Numerical results show that our dynamic scheduling policies can significant improve the performance of WNCS in terms of estimation and control costs while maintaining the stability of the system.

Contents

List of Figures

List of Acronyms

ACK	acknowledgment
AoI	age of information
ARQ	automatic repeat request
AWGN	additive white Gaussian noise
CC	chase combining
CSI	channel state information
FD	full-duplex
FDD	frequency division duplexing
HARQ	hybrid automatic repeat request
HD	half-duplex
i.i.d.	independent and identically distributed
IIoT	industrial internet of things
IoT	internet of things
IR	incremental redundancy
LQR	linear quadratic regulator
LTI	linear time-invariant
MAC	medium access control
MAD	Maximally Allowable Delay
MATI	Maximum Allowable Transfer Interval
MDP	Markov decision process
MMSE	minimum mean-squared error

MPC	model predictive control
MRAE	modified Riccati Algebraic equation
MSE	mean-squared error
NACK	negative-acknowledgment
NCS	networked control system
PPC	packetized predictive control
QoS	quality of service
RL	reinforcement learning
SNR	signal noise ratio
TCP	Transmission Control Protocol
TDD	time division duplexing
VR	virtual reality
WNCS	wireless networked control system

List of Symbols and Notations

$\mathbb{N}$	the sets of positive integers
$\mathbb{N}_0$	the sets of non-negative integers
$\mathbb{E}[\cdot]$	the expectation operator
$(\cdot)^T$	the matrix transpose operator
$\mathrm{Tr}(\cdot)$	the trace operator
$\|\mathbf{v}\|_1$	the sum of the vector $\mathbf{v}$'s elements
$\mathrm{diag}\{v_1, v_2, ..., v_K\}$	the diagonal matrix with the diagonal elements $\{v_1, v_2, ..., v_K\}$
$[u]_{B\times B}$	the $B \times B$ matrix with identical element u
$\mathbb{1}(\cdot)$	the indicator function.
$\rho(\mathbf{A})$	the spectral radius of the square matrix $\mathbf{A}$
$\mathbb{P}[\mathcal{A}]$	the probability of the event $\mathcal{A}$

List of Publications

The following is a list of publications in refereed journals and conference proceedings produced during my Ph.D. candidature. In some cases, the journal papers contain materials overlapping with the conference publications.

Journal Papers

[J1] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li and B. Vucetic, "Real-Time Remote Estimation With Hybrid ARQ in Wireless Networked Control," in *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3490-3504, May 2020.
[J2] K. Huang, W. Liu, Y. Li, B. Vucetic and A. Savkin, "Optimal Downlink–Uplink Scheduling of Wireless Networked Control for Industrial IoT," in *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1756-1772, March 2020.
[J3] K. Huang, W. Liu, Y. Li, A. Savkin and B. Vucetic, "Wireless Feedback Control With Variable Packet Length for Industrial IoT," in *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1586-1590, Sept. 2020.

Conference Papers

[C1] K. Huang, W. Liu, Y. Li and B. Vucetic, "To Retransmit or Not: Real-Time Remote Estimation in Wireless Networked Control," in *Proc. IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019, pp. 1-7.
[C2] K. Huang, W. Liu, Y. Li and B. Vucetic, "To Sense or to Control: Wireless

Networked Control Using a Half-Duplex Controller for IIoT," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1-6.

Chapter 1

Introduction

In this section, we first introduce the background of our research and describe the foundations of wireless networked control system (WNCS). Then we present a review of existing literature. Finally, we describe our motivations and highlight the main contributions of the thesis.

1.1 Background

Wireless networked control systems (WNCSs) have attracted significant attention from both academia and industry in the past two decades. The rapid development of WNCS has benefited from the rise of wireless networks [1–3], advanced control and cloud computing, mainly driven by emerging Industrial Internet of Things (IIoT) applications in intelligent transportation [4], smart grids [5], building automation [6] and industrial automation [7]. Essentially, a WNCS is a spatially distributed control system consisting of a set of sensors, actuators and a remote controller communicating through wireless networks. The networked sensors measure and report the physical process states of interest; the remote controller collects the sensors' measurements and generates control signals; and the wireless actuators receive the control signals

and control the processes.

Low cost and high flexibility of wireless network implementation make WNCS a promising infrastructure technology for industrial monitoring applications. However, due to the unreliable nature of wireless transmission, existing WNCS technologies face significant challenges in supporting mission-critical industrial control systems. Conventional control system design is based on extremely high reliability and low latency of data transmission, while the use of wireless networks will inevitably introduce message loss and delay at all times. Failed or outdated control commands would degrade the system performance, even resulting in the risk of catastrophe. Robust control algorithms and efficient transmission protocol need to be designed jointly to ensure the reliable, stable and timely control of WNCS. Several industrial standard organizations such as ISA [8], HART [9], WINA [10] and ZigBee [11] have been actively pushing the application of wireless technologies in industrial automation and manufacturing. However, it is challenging to meet the stringent latency and reliability requirements of feedback control in WSANs for the existing wireless protocol. The design of data transmission should guarantee deterministic latency and ultra reliability to ensure the reliable operation of control systems. Furthermore, communications-control co-design is required in WNCSs to address the interaction between communications and control systems, which is based on the optimization of control-centric communications rather than conventional rate-centric ones.

1.2 Foundations of WNCS

In this section, we will introduce some basic concepts and elements of WNCS in detail.

1.2.1 Separation Principle

In WNCS, a separation principle states that under some assumptions, the problem of designing an optimal feedback controller for a stochastic system can be broken into two separate ones, designing an optimal observer and an optimal deterministic controller. In a networked control system, if the communication protocol supports the acknowledgement of the packet transmission, it is proved that the separation principle holds, i.e. the optimal estimator and controller can be designed separately. In contrast, the separation principle does not hold if the acknowledgement is not supported [12].

1.2.2 Kalman Filter

The Kalman filter is one of the most popular approaches to estimate the plant state in WNCS [13], also known as linear quadratic estimator. It uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces the estimates of unknown variables based on their joint probability distribution for each timeframe. The Kalman filter has numerous applications in autopilot, navigation system, objects tracking and so on. It has been proved to be the optimal linear estimator in minimum mean-square-error (MMSE) sense with known process and measurement covariances [14]. Some modified Kalman filters have also been proposed to address different models of network loss and delay [12, 15].

1.2.3 Stability

Stability is the fundamental requirement for WNCS. There are two conceptional stability types, called input-output stability and internal stability [16]. Input-output stability describes the ability of the system to generate a bounded output for any

bounded input, while internal stability describes the system's ability to return to equilibrium after a perturbation. Internal stability is usually measured by the magnitude of system state, i.e. the system is said to be stable if the system's steady state is bounded and vice versa. The analysis of stability condition in this thesis refers to internal stability.

1.2.4 Controllability and Observability

Controllability and observability are dual aspects and important properties of a control system. Roughly, the concept of controllability denotes the ability to move a system around in its entire configuration space using only certain admissible manipulations, while observability is a measure of how well the internal states of a system can be inferred from knowledge of its external outputs. The detailed definition is given in our subsequent research.

1.2.5 Representative Controllers

The controller is designed to ensure that WNCS has desirable steady state dynamics and response characteristics. Here we briefly describe three representative controllers: Deadbeat controller [17], Linear Quadratic Regulator (LQR) [18] control and Model Predictive Control (MPC) [19].

A deadbeat controller is also known as time-optimal controller that constructs a control sequence to drive the discrete system state to the origin with a minimum number of time steps. It is often used in process control due to its good dynamic properties, including zero steady-state error, minimum rise time and minimum settling time.

LQR control aims to minimize a quadratic cost function subject to plant state

and control effort. The LQR algorithm is essentially an automated way of finding an appropriate state-feedback controller. It serves as a solution to Linear Quadratic Gaussian (LQG) problems, which deal with uncertain linear systems disturbed by additive Gaussian noise.

MPC is an advanced method of process control that optimizes a linear quadratic control problem over a finite horizon. It optimizes the performance of the current time slot while taking into account the performance of future ones, and it thus has the ability to anticipate future events.

1.2.6 Performance Evaluation

Different from conventional rate-centric performance evaluation, the general performance metric in WNCS is usually quantified by the quadratic cost of plant state and control cost. A general regulation control goal is to reset plant state derivations to zero with minimum effort of control actions. LQR and MPC are typical controller design approaches using the quadratic cost function, defined as the sum of quadratic functions of state derivations and control effort. On the other hand, from the perspective of remote state estimation, the average mean square error (MSE) is normally adopted as the estimation performance criterion. Bounded average MSE serves as the estimation stability condition in most of the existing literature [12, 13, 20–25].

1.2.7 General System Diagram

The general system model of WNCS considered in this thesis is illustrated in Fig 1.1, where the plant and controller are communicating via wireless channel with a scheduler to regulate the transmission between them. Our objective aims to optimize the overall control system performance by properly designing the transmission

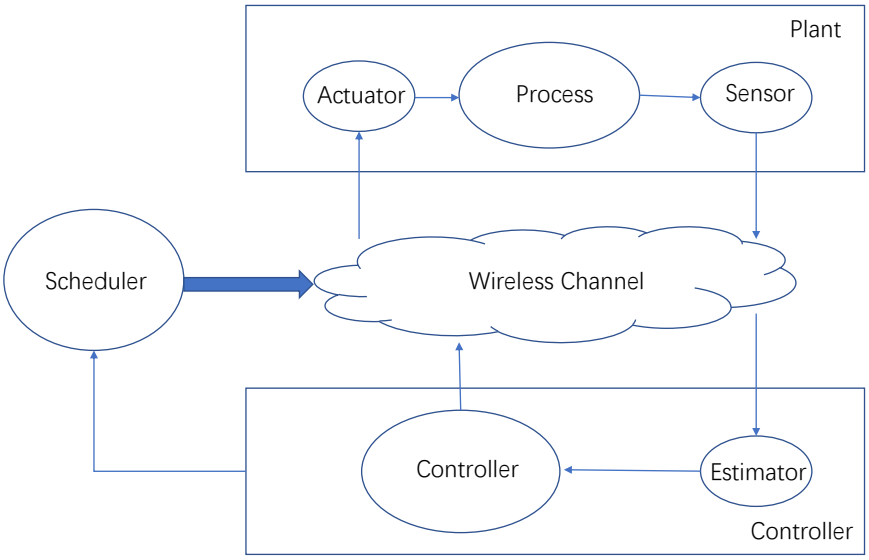


Figure 1.1: General system diagram.

scheduler. We classify the scheduling problems in WNCS into three aspects including sensor's transmission scheduling, downlink-uplink transmission scheduling and controller's transmission scheduling. The comprehensive studies of above three topics are given in Chapter 2-4 respectively.

1.3 Literature Review

In this section, a literature review of related works in WNCS is presented. These works are classified into three categories: communications and control tradeoffs, constrained wireless control, and transmission scheduling and medium access control.

1.3.1 Communications and Control Tradeoffs

The communications and control are highly coupled in WNCS, motivating a codesign approach that integrates wireless networks and control design. The interactions between WNCS communications and control systems include sampling rate, transmission latency and reliability, and power consumption.

Sampling Rate Tradeoff: A low sampling rate usually degrades control performance [26] while high sampling rate may increase transmission traffic in bandwidth-constrained WNCSs, which eventually results in lower system performance [27]. In order to balance such tradeoff, an event-triggered schedule [28] and an event-based control scheme [29] were proposed to reduce the transmission congestion with minor degradation of estimation MSE and system performance. The transmissions were triggered only when the pre-defined state-dependent function exceeded the threshold. More advanced triggered control schemes, such as self-triggered control [29] and the combination of event-triggered and self-triggered control [30] were investigated to schedule transmissions based on frequent monitoring and occasional transmitting.

Lantency vs Sampling Rate Tradeoff: The tradeoff between time-varying sampling period and message delays has been dicussed in [31]. By denoting the Maximum Allowable Transfer Interval (MATI) and the Maximally Allowable Delay (MAD), the tradeoff curves between MATI and MAD were derived to guarantee the control stability of NCS. Such curves provided quantitative information to select network requirements for the given desired control performance.

Transmission Latency vs Reliablity Tradeoff: Latency and reliability are the fundamental parameters in conventional wireless communications. The tradeoffs between them have been examined for dynamic state estimation over wireless communication

channels [32], where a cross-layer design methodology was developed–i.e., the code length was selected depending on the plant dynamics to optimize estimation performance. The stability region of a WNCS in terms of the channel-coding blocklength (latency-dependent) and data rate (reliability-dependent) was derived in [24]. The study concluded that the plant can be stabilized with an arbitrarily large blocklength as long as the SNR of the wireless channel is greater than a system-dependent threshold.

Power Consumption Tradeoff: Motivated by the removal of power supply, limited power consumption has been considered in designing a WNCS, which brings new challenges to balance control performance and power consumption [33–35]. High transmitted power would help increase the receiving SNR and thus improve the performance of the control system, which apparently introduces the tradeoff between plant performance and power consumption. A dynamic power control policy was proposed in [36] to minimize the combination of power expenses and control cost by adapting the transmission power to channel and plant states.

1.3.2 Constrained Wireless Control

Apart from achieving dynamic balance between communication and control cost, many scholars focus on the problems in WNCS from controller's perspective, i.e. constrained wireless control. To overcome the communication constraints, numerous studies have been conducted to design a proper control algorithm.

The modified LQG controller over lossy network under TCP-like protocol was given in [12], which was computed from a Modified Riccati Algebraic Equation (MRAE). A MPC-based controllers were proved to be able to handle wireless message

missing to some extent, by computing the control based on the system's predictive future states [37, 38]. A data-based predictive controller was proposed in [39] to actively compensate for the random round-trip time delay. The controller lumped a sequence of predicted control into one packet together with the timestamp and transmitted it to the actuator, where a network delay compensator was designed and allocated on the plant side. Similarly, a packetized predictive control (PPC) was proposed to increase WNCS robustness with the aid of a buffer attached to the actuator [40]. A sequence of predicted control commands were packed into a long packet and transmitted to the actuator buffer, where old data was overwritten by new arrived data. Quantized PPC was studied in [41] with a proposed vector quantizer to reduce the impact over error-prone digital channel. A more general fading channel was considered in [23], where the coding-free control method was investigated to achieve ultralow latency communications in WNCS. Besides, computation constraints were considered in [42, 43], where an anytime-control algorithm was proposed for a multi-input linear system with time-varying processor availability.

1.3.3 Transmission Scheduling and Medium Access Control

In addition to communications and control tradeoffs (focus on single system), the transsmion in multiple sensors/plants scenarios will introduce extra schuduling problems.

Multiple sensors attached to a single plant may independently transmit their measurements to the controller, which introduces transmission scheduling problems [44]. A multi-sensor event-based scheduler was proposed in [45] to reduce the traffic while keeping remote estimation error covariance bounded to a certain level. In [46], an optimal transmission scheduling policy over shared channels was obtained as a result

of a Markov decision process (MDP). In addition, the power consumption was considered in the sensor's scheduling in [47] and a dynamic sensor's transmission scheduler was proposed based on current state of local covariance. The multi-sensors stability condition for remote estimation with intermittent measurements was derived in [48]. A more effective stability condition over Markov fading channels was derived in [22].

The transmission scheduling over multiple plants in WNCS has also been investigated. [49] proposed an explicit optimal periodic transmission schedule for simple two-multidimensional systems over an infinite horizon. A transmission scheduling policy of multiple systems was designed in [50], which aimed to minimize overall power consumption while guaranteeing a desired expected decreasing rates for Lyapunov stability. The generally internal stability condition of multiple remote estimators for any transmission scheduling over Markov fading channels was given in [25].

Apart from transmission scheduling, various medium access control (MAC) policies were proposed for multiple plants to reduce or avoid transmission collision, such as different waiting time [51], error dependent measure priority [52] and tunable queue parameter [53]. Random access policies were also developed in [54], where the total transmit power of the sensors was minimized while the desired control performance was guaranteed for each involved control loop.

1.4 Motivation

Although numerous works on WNCS have been reported in the literature to address the interaction between control and wireless system for maximum overall system performance and efficiency, the real-time transmission scheduling problem in practical communication protocols is rarely discussed in WNCS. The characteristics of wireless

network such as retransmission mechanism, half-duplex (HD) communication constraints and variable-length packet transmission have not been fully utilized in the design of WNCS. In this thesis, we aim to design a practical communication transmission scheduling in WNCS to achieve optimal control performance. We describe our research motivations according to the subsequent chapters.

Our first research problem (Chapter 2) is how to schedule sensor's transmission scheduling based on HARQ protocol to minimize the remote estimation MSE. Retransmission is always required by conventional communication systems with non-real-time backlogged data to the receivers. Energy-constrained remote estimation systems and systems with a low sampling rate can also benefit from retransmission; see e.g. [55] and [56]. However, it was shown in [57] that retransmissions may not be an effective strategy for a mission-critical real-time remote estimation system as it is a waste of a transmission opportunity, transmitting an out-of-date measurement instead of the current one. Nevertheless, this is true only when a retransmission has the same success probability as a new transmission, e.g., with the standard automatic repeat request (ARQ) protocol. Note that a hybrid ARQ (HARQ) protocol, e.g., with a chase combining (CC) or incremental redundancy (IR) scheme, is able to effectively increase the successful detection probability of a retransmission by combining multiple copies from previously failed transmissions [58]. Therefore, a HARQ protocol has the potential to improve the performance of real-time remote estimation. Although HARQ has been widely investigated in general wireless communication systems for enhancing transmission reliabilities (see [15] and references therein), it has rarely been considered in the literature of real-time remote estimation of time-correlated

dynamic process. Based on HARQ protocol, we naturally ask: does a sensor need retransmission or not for remote estimation in WNCS for IIoT? Therefore, we propose a HARQ-based real-time remote estimation framework.

Our second research problem (Chapter 3) is how to schedule downlink and uplink transmission in terms of HD controller in WNCS. We notice that most existing estimation and control co-design problems in WCNS always ideally assume that the controller works in a full-duplex (FD) mode, where the controller is able to receive sensor's packet and transmit its own control packet at the same time slot. From a wireless communication perspective, although an FD system can improve the spectrum efficiency, it brings other challenges such as self-interference, high device cost and excessive power consumption[59]. Moreover, the conventional half-duplex, time division duplexing (TDD), has been considered as a more attractive duplexing scheme than FDD in 5G dense deployment environment for higher spectrum flexibility, lower device cost, lower spectrum cost and less overhead in channel estimation [60]. Considering an HD controller, the scheduling of sensor's and controller's transmissions has to be optimized in WNCS for Industrial IoT. A frequent schedule of the sensor's transmission results in a better estimation of plant states and thus a higher quality of the control command. On the other side, a frequent schedule of controller's transmission leads to a more timely plant control. Thus, considering the overall control performance of plant's states, e.g., the average cost function of the plant, there exists a fundamental tradeoff between the sensor's and the controller's transmission. Therefore, we proposed an optimal downlink-uplink scheduling policy in terms of HD controller in WNCS.

Our last research problem (Chapter 4) is how to optimally and dynamically design

the packet length for transmission in WNCS. In most of the existing work on WNCSs, the packet length for transmission is fixed. However, from the channel-encoding theory, if a message is encoded into a longer codeword, its reliability is improved at the expense of a longer delay. There exists a fundamental delay-reliability tradeoff since both delay and reliability have a great impact on the control performance. This raises the question of how to balance the tradeoff between reliability and latency of control, and whether variable packet length feedback control will give a better system performance. Therefore, we developed an optimal variable-length packet-transmission policy to minimize the long-term average control cost of WNCS.

1.5 Contributions

We proposed three novel WNCSs where the sensor/controller makes online transmission or schedule decision based on remote estimations and/ or plant states. These optimal transmission scheduling problems are all formulated into MDP-based problems and corresponding system stability conditions are derived in terms of transmission reliability and system plant characters. The main contributions of the thesis are summarized below.

- Firstly, we propose a novel HARQ-based real-time remote estimation system of time-correlated random processes, where the sensor makes online decision on whether to send a new measurement or retransmit the previously failed one depending on both current estimation quality of the receiver and the current number of retransmissions of the sensor. We formulate the problem to optimize the sensor's transmission control policy so as to maximize the long-term performance of the receiver in terms of the average MSE for both the static

and Markov fading channels. Since it is not clear whether the long-term average MSE can be bounded or not, we derive an elegant sufficient condition in terms of the transmission reliability provided by the HARQ protocol and parameters of the process of interest to ensure that an optimal policy exists and stabilizes the remote estimation system. We derive a structural property of the optimal policy (i.e., the optimal policy is a switching-type policy) and give an easy-to-compute suboptimal policy.

- Secondly, we propose a WNCS with an HD controller, where the controller schedules the sensor's measurement transmission and its own control-command transmission depending on both the estimation quality of current plant states and the current cost function of the plant. We formulate a problem to optimally design the transmission-scheduling policy to optimize the long-term control performance of the WNCS in terms of the average cost function for both the one-step and v-step controllable plants. As the long-term average cost of the plant may not be bounded with high transmission-error probabilities leading to an unstable situation in the static channel scenario, we derive a necessary and sufficient condition in terms of the transmission reliabilities of the sensor-controller and controller-actuator channels and the plant parameters to ensure the existence of an optimal policy that stabilizes the plant. In the fading channel scenario, we derive a necessary condition and a sufficient condition in terms of the uplink and downlink channel qualities under which the optimal transmission scheduling policy exits. We also derive a suboptimal policy with a low computation complexity.

- Finally, We propose another WNCS, where the controller is able to adaptively

change the packet length for control based on the current status of the physical process. We formulate the optimal design problem of variable-length packet transmission policy into a semi-Markov decision process (MDP) problem, which minimizes the long-term average cost function of the WNCS. For the variable-length policy, we derived the stability condition of the WNCS, i.e., the necessary and sufficient condition on the existence of an optimal policy that can stabilize the WNCS (i.e., make the cost function-bounded) in terms of transmission reliabilities with different packet lengths and the control system parameter. The analysis is not trivial, since neither the optimal policy nor its long-term average cost function of the WNCS has a closed-form formula. We prove the necessity by constructing a tractable virtual policy that can achieve a better performance than the optimal policy. The sufficiency is proved by analyzing fixed-length policies that achieve a worse performance than the optimal variable-length policy. We investigate the fixed-length packet transmission policies of the WNCS with different packet lengths, and derive a closed-form stability condition in terms of packet length, transmission reliability and control system parameter. Such result has not been obtained before in the literature, and will provide an important design guideline for the fixed-length policy in WNCS.

Chapter 2

Real-Time Remote Estimation with Hybrid ARQ in Wireless Networked Control

2.1 Introduction

Real-time remote estimation is critical for networked control applications such as industrial automation, smart grid, vehicle platooning, drone swarming, immersive virtual reality (VR) and the tactile Internet [61]. For such real-time applications, high quality remote estimation of the states of dynamic processes over unreliable links is a major challenge. The sensor's sampling policy, the estimation scheme at a remote receiver, and the communication protocol for state-information delivery between the sensor and the receiver should be designed jointly.

To enable the optimal design of wireless remote estimation, the performance metric for the remote estimation system needs to be selected properly. For some applications, the model of the dynamic process being monitored is unknown and the receiver is not able to estimate the current state of the process based on the previously received states, i.e., a state-monitoring-only scenario [62]. In this scenario,

the performance metric is the age-of-information (AoI), which reflects how old the freshest received sensor measurement is, since the moment that measurement was generated at the sensor [62]. However, in practice, most of the dynamic processes are time-correlated, and the state-changing rules can be known by the receiver to some extent. Therefore, the receiver can estimate the current state of the process based on the previously received measurements and the model of the dynamic process (see e.g., [63, 64]), especially when the transmission of the packet that carries the current sensor measurement fails or is delayed. In this sense, the estimation mean squared error (MSE) is the appropriate performance metric.

A HARQ protocol is considered in the design of wireless transmission in remote estimation, where a retransmission has a more successful detection probability. There exists a fundamental tradeoff between the reliability and freshness of the sensor's measurement transmission. When a failed transmission occurs, the sensor can either retransmit the previous measurement, such that the receiver can obtain a more reliable old measurement, or transmit a new, but less reliable measurement. In this chapter, we propose a HARQ-based real-time remote estimation framework and optimally design the sensor's transmission policy to minimize the estimation MSE.

2.2 System Model

We consider a basic system setting where a *smart sensor*[1] periodically samples, pre-estimates and sends its local estimation of a dynamic process to a remote receiver through a wireless link with packet dropouts, as illustrated in Fig. 2.1.

[1]Unlike a traditional sensor that sends raw measurement data to the receiver, a smart sensor with a better computing capacity is able to pre-process the measurement data before transmission which can improve the quality of remote estimation [46, 47, 65, 66].

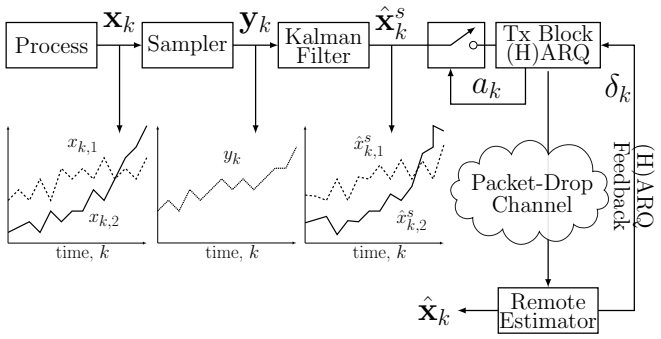


Figure 2.1: Proposed remote estimation system with HARQ, where $\mathbf{x}_k \triangleq [x_{k,1}, x_{k,2}]^T$ is the two-dimensional state vector of the dynamic process, and $\hat{\mathbf{x}}^s \triangleq \left[\hat{x}_{k,1}^s, \hat{x}_{k,1}^s\right]^T$.

2.2.1 Dynamic Process Modeling

We consider a general discrete LTI model for the dynamic process as (see e.g., [63, 65, 66])

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{v}_k, \end{aligned} \tag{2.2.1}$$

where the discrete time steps are determined by the sensor's sampling period T_s, $\mathbf{x}_k \in \mathbb{R}^n$ is the process state vector, $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the state transition matrix, $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement vector of the smart sensor attached to the process, $\mathbf{C} \in \mathbb{R}^{m\times n}$ is the measurement matrix, and $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^m$ are the process and measurement noise vectors, respectively. We assume $\mathbf{w}_k$ and $\mathbf{v}_k$ are independent and are identically distributed (i.i.d.) zero-mean Gaussian processes with corresponding covariance matrices $\mathbf{Q}_w$ and $\mathbf{Q}_v$, respectively. The initial state $\mathbf{x}_0$ is a zero-mean Gaussian with covariance matrix $\mathbf{\Sigma}_0$. To avoid trivial problems, we assume that $\rho^2(\mathbf{A}) > 1$, where $\rho^2(\mathbf{A})$ is the maximum squared eigenvalue of $\mathbf{A}$ [49].

2.2.2 State Estimation at the Smart Sensor

Since the sensor's measurements are noisy, a smart sensor with sufficient computation and storage capacity is required to estimate the state of the process, $\mathbf{x}_k$, using a Kalman filter [65, 66], which gives the minimum estimation MSE based on the current and previous raw measurements:

$$\mathbf{x}^s_{k|k-1} = \mathbf{A}\mathbf{x}^s_{k-1|k-1} \tag{2.2.2a}$$

$$\mathbf{P}^s_{k|k-1} = \mathbf{A}\mathbf{P}^s_{k-1|k-1}\mathbf{A}^T + \mathbf{Q}_w \tag{2.2.2b}$$

$$\mathbf{K}_k = \mathbf{P}^s_{k|k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}^s_{k|k-1}\mathbf{C}^T + \mathbf{Q}_v)^{-1} \tag{2.2.2c}$$

$$\mathbf{x}^s_{k|k} = \mathbf{x}^s_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}\mathbf{x}^s_{k|k-1}) \tag{2.2.2d}$$

$$\mathbf{P}^s_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}^s_{k|k-1} \tag{2.2.2e}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix, $\mathbf{x}^s_{k|k-1}$ is the priori state estimation, $\mathbf{x}^s_{k|k}$ is the posteriori state estimation at time k, $\mathbf{K}_k$ is the Kalman gain, $\mathbf{P}^s_{k|k-1}$ and $\mathbf{P}^s_{k|k}$ represent the priori and posteriori error covariance at time k, respectively. The first two equations present the prediction steps while the last three equations present the updating steps [67]. Note that $\mathbf{x}^s_{k|k}$ is the output of the Kalman filter at time k, i.e., the pre-filtered measurement of $\mathbf{y}_k$, with the estimation error covariance $\mathbf{P}^s_{k|k}$.

As the performance of remote estimation system depends on both the qualities of the local state estimation and the imperfect communication link, the quality of remote estimation is worse than that of the local estimation. Thus, the remote estimation system cannot be stabilized without stabilizing the local estimation system. In this work, we focus on the effect of communication protocols on the stability and quality of the remote estimation system. Thus, we assume that the local estimation of the Kalman filter is stable, i.e., the local estimation error covariance converges to

a bounded matrix [46, 47, 63, 65, 66].

Assumption 2.1. The LTI system (2.2.1) has the properties that $(\mathbf{A}, \mathbf{C})$ is observable and $(\mathbf{A}, \sqrt{\mathbf{Q}_w})$ is reachable, where $\sqrt{\mathbf{Q}_w}$ is the square root of the positive definite matrix $\mathbf{Q}_w$.[2] Thus, the local Kalman filter of the system in (2.2.2) is stable, i.e., the error covariance matrix $\mathbf{P}^s_{k|k}$ converges to a finite matrix $\bar{\mathbf{P}}_0$ when k is sufficiently large [68].

Although the convergence rate of the Kalman filter cannot be derived directly, our simulation results show that the local Kalman filter usually converges to the steady state in less then 10 steps. Therefore, in the rest of the chapter, we assume that the local Kalman filter operates in the steady state [46, 47, 63, 65, 66], i.e., $\mathbf{P}^s_{k|k} = \bar{\mathbf{P}}_0$. To simplify the notation, we use $\hat{\mathbf{x}}^s_k$ to denote the sensor's estimation, $\mathbf{x}^s_{k|k}$.

2.2.3 Wireless Channel

We consider both a static channel and a finite-state time-homogeneous Markov fading channel. In general, the channel input output equation is given as

$$z = \tilde{h}\vartheta + \omega, \tag{2.2.3}$$

where ϑ and z are the input and the output symbols of the channel, $\tilde{h}$ is the channel coefficient and ω is the additive white Gaussian noise (AWGN) [69].

For the static channel, the channel power gain h_k does not change with time, i.e., $h_k = h > 0, \forall k$. For the Markov channel, the channel power gain h_k remains constant during the kth time slot and changes slot by slot, where $h_k > 0, \forall k$. We assume that the Markov channel has B states, i.e, $\mathcal{U} \triangleq \{u_1, ..., u_B\}$, and $h_k \in \mathcal{U}$. The probability

[2] $(\mathbf{A}, \mathbf{C})$ is observable iff the matrix concatenation $\left[\mathbf{C}^T, \mathbf{A}^T\mathbf{C}^T, \cdots, (\mathbf{A}^n)^T\mathbf{C}^T\right]$ is of full rank; $(\mathbf{A}, \sqrt{\mathbf{Q}_w})$ is reachable iff the matrix concatenation $\left[\sqrt{\mathbf{Q}_w}, \mathbf{A}\sqrt{\mathbf{Q}_w}, \cdots, \mathbf{A}^n\sqrt{\mathbf{Q}_w}\right]$ is of full rank [68].

of transition from state i to state j is $p_{i,j}$, and the matrix of channel state transition probability is given as

$$\mathbf{\Pi} \triangleq \begin{bmatrix} p_{1,1} & \cdots & p_{B,1} \\ \vdots & \ddots & \vdots \\ p_{1,B} & \cdots & p_{B,B} \end{bmatrix}. \tag{2.2.4}$$

We assume that the channel state information (CSI) is available at both the sensor and the receiver [70–72]. In general, the receiver-side CSI is obtained by letting the sensor send a known sequence at the beginning of each time slot (called a training sequence) and estimating the channel using symbols detected earlier [69]. Then, the CSI needs to be quantized and sent to the sensor over a limited-rate feedback channel [73]. In practice, the CSI estimation and feedback may not be perfect, and their qualities depend on the length of the training sequence and the rate of the feedback channel. In this work, we focus on the maximum achievable performance of the remote estimation system when the sensor and the receiver have the perfect CSI. The impact of imperfect CSI and delays induced by CSI estimation and feedback will be considered in our future work.

2.2.4 HARQ-Based Communication

The sensor's estimation is quantized into $(L \times R)$ bits and then coded into a packet with L symbols, where the symbol duration is T_s' and R is the coding rate. We assume that the packet length is equal to the sampling period, i.e., $LT_s' = T_s$. In other words, the sensors perform the next sampling once the current measurement-carrying packet has been delivered to the receiver. Thus, there exists a unit packet-transmission delay between the sensor and the receiver. For example, the sensor's measurement at the beginning of time slot k, $\mathbf{y}_k$, is filtered and sent to the receiver before time slot $(k+1)$.

It is assumed that the sensor and the receiver are perfectly synchronized.

The acknowledgment/negative-acknowledgment (ACK/NACK) message is fed back from the receiver to the sensor perfectly without any delay when the packet detection succeeds/fails. If an ACK is received by the sensor, it will send a new (pre-filtered) measurement in the next time slot. If a NACK is received, the sensor can decide whether to retransmit the unsuccessfully transmitted measurement using its ARQ protocol or to send a new measurement. We introduce the binary variable $\gamma_k \in \{1, 0\}$ to indicate the successful and failed packet detection in time slot k, respectively. Note that the sensor discards the old measurement once the decision of a new transmission is made.

For the standard ARQ protocol, the receiver discards the failed packets, and the sensor simply resends the previously failed packet if a retransmission is required. Thus, the successful packet detection probability at each time is independent of the current number of retransmissions.

For a HARQ protocol, the receiver buffers the incorrectly received packets and the detection of the retransmitted packet will utilize all the buffered related packets. In the CC-HARQ case, the sensor resends the previously failed packet if a retransmission is required, and the receiver optimally combines (i.e., the maximal ratio combining method) all the previously received replicas of the packet of the same message and make a detection. In the IR-HARQ case, each retransmitted packet is an incremental redundancy of the same message, and the receiver treats the sequence of all the buffered replicas as a long codeword to detect the transmitted massage.

Given the channel power gains, the probability that the message cannot be detected within r transmission attempts started from time slot $k-(r-1)$ is given as [74]

and [75]

$$\begin{aligned}&\mathbb{P}\left[\gamma_k = 0 | r \text{ transmission attampts}\right]\\ &\approx \begin{cases} Q\left(\frac{L^{\frac{1}{2}}(\log_2(1+\sum_{i=0}^{r-1} h_{k-i}\mathsf{SNR})+\frac{\log_2 L}{L}-R)}{\sqrt{1-\frac{1}{(1+\sum_{i=0}^{r-1} h_{k-i}\mathsf{SNR})^2}}\log_2 e}\right), \text{CC-HARQ} \\ Q\left(\frac{L^{\frac{1}{2}}(\sum_{i=0}^{r-1}\log_2(1+h_{k-i}\mathsf{SNR})+\frac{\log_2 rL}{L}-R)}{\sqrt{\sum_{i=0}^{r-1}(1-\frac{1}{(1+h_{k-i}\mathsf{SNR})^2})}\log_2 e}\right), \text{IR-HARQ} \end{cases}\end{aligned} \tag{2.2.5}$$

where SNR is the signal-to-noise ratio at the receiver with unit channel power gain, and the approximation in (2.2.5) is based on the results of the finite-blocklength information theory for AWGN channel; see e.g. [74] and Eq. (9) in [75].

2.2.5 State Estimation at the Receiver

Although the sensor sends a packet to the receiver in every time slot, the receiver is not able to always detect the packet successfully due to the noise. Therefore, the receiver can only estimate the current process state from the most recently received sensor's estimation when a failed packet detection occurs.

We assume that the latest sensor's estimation that is available at the receiver at the beginning of time slot k, i.e., $\hat{\mathbf{x}}^s_{t_k}$, was generated at the beginning of time slot t_k. Therefore, the receiver-side AoI at the beginning of time slot k can be defined as [62]

$$q_k \triangleq k - t_k, \forall k, \tag{2.2.6}$$

and $q_k \geq 1$.

As the latest available sensor's estimation was generated q_k-step earlier, the receiver needs to estimate the current state based on the dynamic process model (2.2.1). The receiver's MSE optimal estimator at the beginning of time slot k is given as [63]

$$\hat{\mathbf{x}}_k = \mathbf{A}^{k-t_k}\hat{\mathbf{x}}^s_{t_k} = \mathbf{A}^{q_k}\hat{\mathbf{x}}^s_{t_k}, \tag{2.2.7}$$

and the corresponding estimation error covariance is [46, 47, 63, 65, 66]

$$\mathbf{P}_k \triangleq \mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\right] \tag{2.2.8}$$

$$= f^{q_k}(\bar{\mathbf{P}}_0), \tag{2.2.9}$$

where (2.2.9) is obtained by substituting (2.2.7) and (2.2.1) into (2.2.8), $f(\mathbf{X}) \triangleq \mathbf{A}\mathbf{X}\mathbf{A}^T + \mathbf{Q}_w$, $f^{n+1}(\cdot) \triangleq f(f^n(\cdot))$ when $n \geq 1$, and $f^1(\cdot) \triangleq f(\cdot)$. Note that $\mathbf{P}_k$ takes value from a countable infinity set [49], i.e., $\mathbf{P}_k \in \{f(\bar{\mathbf{P}}_0), f^2(\bar{\mathbf{P}}_0), \cdots\}$.

The receiver's estimation MSE at the beginning of time slot k is $\mathrm{Tr}\left(\mathbf{P}_k\right)$. Note that the operator $\mathrm{Tr}\left(f^n(\cdot)\right)$ is monotonically increasing with respect to (w.r.t.) n, i.e., $\mathrm{Tr}\left(f^{n_1}(\bar{\mathbf{P}}_0)\right) \leq \mathrm{Tr}\left(f^{n_2}(\bar{\mathbf{P}}_0)\right)$ if $\rho^2(\mathbf{A}) > 1$ and $1 \leq n_1 \leq n_2$ (see Lemma 3.1 in [49]).

Remark 2.1. *From* (2.2.9)*, the estimation MSE is a non-linear function of the AoI, and thus, q_k can also be treated as the estimation quality indicator.*

2.2.6 Performance Metric

The long-term average MSE of the dynamic process is defined as

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{P}_k\right)\right], \tag{2.2.10}$$

where $\limsup_{K\to\infty}$ is the limit superior operator.

2.3 Optimal Transmission Control: Analysis and Problem Formulation

For the standard ARQ, as the chance of the successful detection of a new transmission and that of a retransmission are the same, *the optimal policy is to always*

transmit the current sensor's estimation, i.e., a non-retransmission policy [57]. For a HARQ protocol, the probability of successful packet detection in time slot k depends on the number of consecutive transmission attempts of the original message and the experienced channel conditions. Since a new transmission is less reliable than a retransmission, there exists an inherent trade-off between retransmitting a previously failed local state estimation with a higher success probability, and sending the current state estimation with a lower success probability. Therefore, when a packet detection error occurs, the sensor needs to optimally make a decision on whether to retransmit it or to start a new transmission.

2.3.1 Transmission-Control Policy

Let $a_k \in \{0, 1\}$ be the sensor's decision variable at time k as illustrated in Fig. 2.2.1. If $a_k = 0$, the sensor sends the new measurement to the receiver in time slot k; otherwise, it retransmits the unsuccessfully transmitted measurement. It is clear that $a_k = 0$ if the the packet transmitted in time slot $(k-1)$ was successful.

Let r_k denote the number of consecutive transmission attempts before time slot k. As r_k only depends on the sensor's transmission-control policy, it has the updating rule as

$$r_k = \begin{cases} 1, & \text{if } a_{k-1} = 0 \\ r_{k-1} + 1, & \text{otherwise,} \end{cases} \tag{2.3.1}$$

where $r_k \geq 1, \forall k$. From the definition of the estimation quality indicator (2.2.6), the updating rule of q_k is given as

$$q_k = \begin{cases} r_k, & \gamma_{k-1} = 1 \\ q_{k-1} + 1, & \gamma_{k-1} = 0, \end{cases} \tag{2.3.2}$$

where $q_k \geq 1, \forall k$. As the estimation quality indicator depends on the current number

of transmission attempts and also the control policy, plugging (2.3.1) into (2.3.2), we further have

$$q_k = \begin{cases} 1, & a_{k-1} = 0, \gamma_{k-1} = 1 \\ r_{k-1} + 1, & a_{k-1} = 1, \gamma_{k-1} = 1 \\ q_{k-1} + 1, & \gamma_{k-1} = 0. \end{cases} \tag{2.3.3}$$

2.3.2 Packet Error Probability with Online Transmission Control

If the sensor decides to transmit a new measurement in time slot k, i.e., $a_k = 0$, the packet error probability in time slot k is obtained directly from (2.2.5) as

$$\mathbb{P}\left[\gamma_k = 0 | a_k = 0\right] = \mathbb{P}\left[\gamma_k = 0 | r = 1\right]. \tag{2.3.4}$$

If a transmission is failed and a retransmission decision has been made, i.e., $a_k = 1$, the packet error probability based on (2.2.5) can be obtained as

$$\begin{aligned} \mathbb{P}\left[\gamma_k = 0 | a_k = 1\right] &= \mathbb{P}\left[\gamma_k = 0 | \gamma_{k-1}, ..., \gamma_{k-r_k} = 0\right] \\ &= \frac{\mathbb{P}\left[\gamma_k = 0 | r = r_k + 1\right]}{\mathbb{P}\left[\gamma_{k-1} = 0 | r = r_k\right]}. \end{aligned} \tag{2.3.5}$$

In the Markov channel scenario, we assume that the packet error probability in (2.3.5) is a function of the current channel power gain h_k, and the state indicator $\mathbf{\Omega}_k$ is a function of the previously experienced r_k channel states, which does not rely on their order [3]. To be specific, we define $\mathbf{\Omega}_k \triangleq [r_k(1), r_k(2), ..., r_k(B)]$, where $r_k(i)$ is the occurrence number of the channel state with channel power gain u_i during time slots $k - r_k$ to $k - 1$. Thus, $\sum_{j=1}^{B} r_k(j) = r_k$. In other words, $\mathbf{\Omega}_k$ is a sorted counter of the relevant historical channel states, and $\mathbf{\Omega}_k \in \mathbb{N}_0^B \backslash \mathbf{0}$. By introducing the current

[3] This assumption is in line with the approximation in (2.2.5).

channel state index $\Xi_k \in \{1, \cdots, B\}$, i.e., $h_k = u_{\Xi_k} \in \mathcal{U}$, $\mathbf{\Omega}_k$ is updated as

$$\mathbf{\Omega}_k = \begin{cases} \mathbf{1}_{\Xi_{k-1}}, & \text{if } a_{k-1} = 0 \\ \mathbf{\Omega}_{k-1} + \mathbf{1}_{\Xi_{k-1}}, & \text{if } a_{k-1} = 1 \end{cases} \tag{2.3.6}$$

where $\mathbf{1}_i \triangleq [\underbrace{0, \cdots, 0, 1}_{i}, 0, \cdots, 0]$.

Therefore, using (2.2.5), the detection error probabilities of a new transmission and $(r_k + 1)$ transmission attempts can be rewritten as functions of Ξ_k and $\{\mathbf{\Omega}_k, \Xi_k\}$, respectively, as

$$\begin{aligned} &\mathbb{P}\left[\gamma_k = 0 | r = 1\right] \\ &\approx \begin{cases} Q\left(\frac{L^{\frac{1}{2}}(\log_2(1+\left(u_{\Xi_k}\right)\mathsf{SNR})+\frac{\log_2 L}{L}-R)}{\sqrt{1-\frac{1}{(1+\left(u_{\Xi_k}\right)\mathsf{SNR})^2}}\log_2 e}\right), \text{CC-HARQ} \\ Q\left(\frac{L^{\frac{1}{2}}(\log_2(1+u_{\Xi_k}\mathsf{SNR})+\frac{\log_2 rL}{L}-R)}{\sqrt{(1-\frac{1}{(1+u_{\Xi_k}\mathsf{SNR})^2})\log_2 e}}\right), \text{IR-HARQ} \end{cases} \end{aligned} \tag{2.3.7}$$

and

$$\begin{aligned} &\mathbb{P}\left[\gamma_k = 0 | r = r_k + 1\right] \\ &\approx \begin{cases} Q\left(\frac{L^{\frac{1}{2}}(\log_2(1+\left(\sum_{j=1}^{B} r_k(j)u_j+u_{\Xi_k}\right)\mathsf{SNR})+\frac{\log_2 L}{L}-R)}{\sqrt{1-\frac{1}{(1+\left(\sum_{j=1}^{B} r_k(j)u_j+u_{\Xi_k}\right)\mathsf{SNR})^2}}\log_2 e}\right), \text{CC} \\ Q\left(\frac{L^{\frac{1}{2}}(\sum_{j=1}^{B} r_k(j)\log_2(1+u_j\mathsf{SNR})+\log_2(1+u_{\Xi_k}\mathsf{SNR})+\frac{\log_2 rL}{L}-R)}{\sqrt{\sum_{j=1}^{B} r_k(j)(1-\frac{1}{(1+u_j\mathsf{SNR})^2})+(1-\frac{1}{(1+u_{\Xi_k}\mathsf{SNR})^2})\log_2 e}}\right), \text{IR} \end{cases} \end{aligned} \tag{2.3.8}$$

Thus, by taking (2.3.7) and (2.3.8) into (2.3.4) and (2.3.5), the packet error probability can be uniformly written as

$$\mathbb{P}\left[\gamma_k = 0\right] = \begin{cases} \mathbb{P}\left[\gamma_k = 0 | r = 1\right] \triangleq \tilde{g}(\mathbf{0}, \Xi_k), & a_k = 0, \\ \dfrac{\mathbb{P}\left[\gamma_k = 0 | r = r_k + 1\right]}{\mathbb{P}\left[\gamma_{k-1} = 0 | r = r_k\right]} \triangleq \tilde{g}(\mathbf{\Omega}_k, \Xi_k), & a_k = 1. \end{cases} \tag{2.3.9}$$

where $\mathbf{0} \triangleq \underbrace{[0, 0, ..., 0]}_{B}$, and $\mathbb{P}\left[\gamma_{k-1} = 0 | r = r_k\right]$ can be obtained by (2.3.7) and (2.3.8), respectively, for the cases $r_k = 1$ and $r_k = r_{k-1} + 1 > 1$.

In the static channel scenario, i.e., a special case of the Markov channel scenario, as the channel power gains are identical to each other the packet error probability in

(2.3.9) can be rewritten as a function of the current number of transmission attempts based on (2.3.7) and (2.3.8) as

$$\mathbb{P}\left[\gamma_k = 0\right] = \begin{cases} \mathbb{P}\left[\gamma_k = 0 | r = 1\right] \triangleq g(1), & a_k = 0, \\ \dfrac{\mathbb{P}\left[\gamma_k = 0 | r = r_k + 1\right]}{\mathbb{P}\left[\gamma_{k-1} = 0 | r = r_k\right]} \triangleq g(r_k + 1), & a_k = 1. \end{cases} \tag{2.3.10}$$

As the packet error probability of a retransmission is smaller than a new transmission under the same channel conditions, we have the following inequalities for the Markov and static channel scenarios, respectively:

$$\Lambda_i' \triangleq \tilde{g}(\mathbf{0}, i) > \tilde{g}(\mathbf{\Omega}_k, i), \forall k, \tag{2.3.11}$$

and

$$\Lambda_0' \triangleq g(1) > g(r_k + 1), \forall k, \tag{2.3.12}$$

where Λ_i' is the packet error probability of a new transmission with the channel power gain u_i in the Markov channel scenario, and Λ_0' is the packet error probability of a new transmission in the static channel scenario.

For the Markov channel, the largest packet error rate of a retransmission with channel power gain u_i is defined as

$$\Lambda_i \triangleq \max_{\mathbf{\Omega_k} \in \mathbb{N}_0^B \backslash \mathbf{0}} \tilde{g}(\mathbf{\Omega_k}, i), i = 1, 2, ..., B. \tag{2.3.13}$$

For the static channel, the largest packet error rate of a retransmission is defined as

$$\Lambda_0 \triangleq \max_{r>1} g(r). \tag{2.3.14}$$

2.3.3 Problem Formulation

The sensor's transmission control policy is defined as the sequence $\{a_1, a_2, ..., a_k, \cdots\}$, where a_k is the control action in time slot k. In what follows we optimize the sensor's

policy such that the long-term estimation error is minimized, i.e.,

$$\min_{a_1,a_2,\ldots,a_k,\cdots} \limsup_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{P}_k\right)\right]. \tag{2.3.15}$$

It is possible that the long-term estimation error may never be bounded no matter how we choose the policy, if the channel quality is always bad or the dynamic process (2.2.1) changes rapidly. Therefore, *it is also important to investigate the condition in terms of the transmission reliability and the dynamic process parameters under which the remote estimation system can be stabilized, i.e., the long-term estimation MSE can be bounded.*

To shed light on the stability condition and the optimal policy structure, we first consider the simplified case, i.e., the static channel scenario, in Sec. 2.4. The insights are leveraged to investigate the general Markov channel scenario in Sec. 2.5.

2.4 Optimal Policy: Static Channel

In this section, we investigate the optimal transmission control policy in static channel.

2.4.1 MDP Formulation

From (2.2.9), (2.3.1) and (2.3.3), the estimation MSE $\mathrm{Tr}\left(\mathbf{P}_k\right)$ and states r_k and q_k only depend on the previous action and states, i.e., a_{k-1}, r_{k-1} and q_{k-1}. Therefore, the online decision problem (2.3.15) can be formulated as a discrete time Markov decision process (MDP) as follows.

1) The state space is defined as $\mathbb{S} \triangleq \{(r,q) : r \leq q,\ (r,q) \in \mathbb{N}\times\mathbb{N}\}$, where the number of transmission attempts, r, should be no larger than the estimation quality

indicator (i.e., the AoI), q, from the definition. The state of the MDP at time k is $s_k \triangleq (r_k, q_k) \in \mathbb{S}$.

2) The action space is defined as $\mathbb{A} \triangleq \{0, 1\}$. A policy is a mapping from states to actions, i.e., $\pi : \mathbb{S} \to \mathbb{A}$. Recall that the action at time k, $a_k \triangleq \pi(s_k) \in \mathbb{A}$ indicates a new transmission ($a_k = 0$) or a retransmission ($a_k = 1$).

3) The state transition function $P(s'|s, a)$ characterizes the probability that the state transits from s at time $(k-1)$ to s' at time k, with action a at time $k-1$. As the transition is time-homogeneous, we can drop the time index k here. Let $s = (r, q)$ and $s' = (r', q')$ denote the current and next state, respectively. Based on the packet error probability (2.3.10) and the state updating rules (2.3.1) and (2.3.3), we have the state transition function as

$$P(s'|s,a) = \begin{cases} 1 - g(1), & \text{if } a = 0, s' = (1,1) \\ g(1), & \text{if } a = 0, s' = (1, q+1) \\ 1 - g(r+1), & \text{if } a = 1, s' = (r+1, r+1) \\ g(r+1), & \text{if } a = 1, s' = (r+1, q+1) \\ 0, & \text{otherwise.} \end{cases} \tag{2.4.1}$$

4) The one-stage (instantaneous) cost, i.e., the estimation MSE based on (2.2.9), is a function of the current state of q:

$$c(s,a) \triangleq \mathrm{Tr}\left(f^q(\bar{\mathbf{P}}_0)\right), \tag{2.4.2}$$

which is independent of the action.

Therefore, the problem (2.3.15) is equivalent to solving the classical *average cost optimization* problem of the MDP. Assuming the existence of a stationary and deterministic optimal policy, we can effectively solve the MDP problem using standard

methods such as the relative value iteration algorithm [76, Chapter 8].

2.4.2 Optimal Policy: Condition of Existence

Since the cost function grows exponentially with the state q, it is possible that the long-term average cost with a HARQ-based transmission control policy, π, in the state space $\mathbb{S}$ cannot be bounded, i.e., the remote estimation system is unstable. We give the following sufficient condition of the existence of an optimal policy that has a bounded long-term estimation MSE.

Theorem 2.1. For the static channel, there exists a stationary and deterministic optimal policy π^* of problem (2.3.15), if the following condition holds:

$$\Lambda_0 \rho^2(\mathbf{A}) < 1, \tag{2.4.3}$$

where Λ_0 is the largest packet error probability of a retransmission defined in (2.3.14).

Proof. See Appendix A. □

Remark 2.2. *From Theorem 2.1, it is clear that the optimal policy exists if we have a good channel condition and a good HARQ scheme that guarantees high retransmission reliability (i.e., a small $g(r)$ and hence a small Λ_0), or if the dynamic process does not change quickly, which is easy to estimate (i.e., a small $\rho^2(\mathbf{A})$).*

2.4.3 Optimal Policy: The Structure

We show that the optimal policy has a switching structure as follows.

Theorem 2.2. The optimal policy π^* of problem (2.3.15) is a switching-type policy, i.e., (i) if $\pi^*(r, q) = 0$, then $\pi^*(r + z, q) = 0$; (ii) if $\pi^*(r, q) = 1$, then $\pi^*(r, q + z) = 1$, where z is any positive integer.

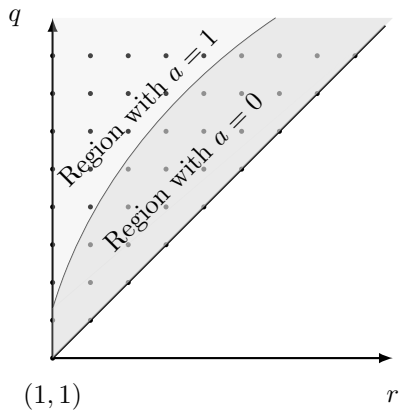


Figure 2.2: The switching structure of the optimal policy in the state space $\mathbb{S}$.

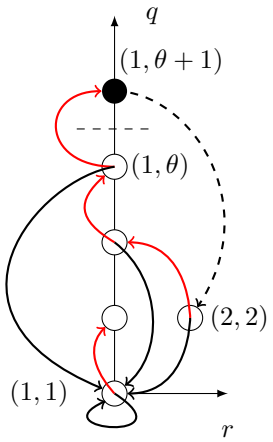


Figure 2.3: Illustration of the state space and state transitions in the high SNR scenario, where $\theta = 4$.

Proof. See Appendix A. □

In other words, for the optimal policy, the two-dimensional state space $\mathbb{S}$ is divided into two regions by a curve, and the decision actions of the states within each region are the same, as illustrated in Fig. 2.2.

Remark 2.3. *Note that the switching structure can help save storage space for online implementation, since the smart sensor only needs to store switching-boundary states rather than the actions on the entire state space. At each time, the sensor simply needs to compare the current state with the boundary states to make the optimal decision.*

2.4.4 Optimal Policy: A Special Case (Ultra-Low Retransmission Packet Error Probabilities)

We consider the high SNR scenario where retransmissions have ultra-low packet error probabilities that are much smaller than a new transmission. Therefore, in the

high SNR scenario, we assume that a retransmission is always successful while a new transmission is not, and the optimal policy always exists from Theorem 2.1.

Due to the successful retransmissions, it can be noted from (2.4.1) that the states in $\mathbb{S}$ with $r > 1$ and $r \neq q$ are transient states. Also, since a successful retransmission must be followed by a new transmission, the states in $\mathbb{S}$ with $r = q$ and $r > 1$ are transient, and the state $(2, 2)$ has the action of new transmission. Furthermore, due to the switching structure of the optimal policy in Theorem 2.2, we set a policy-switching threshold $\theta \in \mathbb{N}$ for the states with $r = 1$, where the states $q > \theta, r = 1$ choose the action of retransmission while the states with $q \leq \theta, r = 1$ choose the action of new transmission. Then, it is easy to see that the states with $r = 1$ and $q > \theta + 1$ are transient states. Finally, the countably infinite state space $\mathbb{S}$ is reduced to a finite state space $\mathbb{S}' = \{(2, 2), (1, q), \forall q \in \{1, ..., \theta + 1\}$ as illustrated in Fig. 2.3. Only the state $(1, \theta + 1)$ has the action $a = 1$, and the other states have the action $a = 0$. Note that the next state will be $(2, 2)$ if the current state is $(1, \theta + 1)$.

Therefore, θ is the key design parameter to be optimally designed. The policy optimization problem in the state space $\mathbb{S}$ is transformed to the one-dimensional problem. By calculating the stationary distribution of the states in $\mathbb{S}'$ with a given θ, the average cost of the Markov chain (i.e., the average MSE) denoted by ζ can be obtained, and we have the following result.

Proposition 2.1. In the high SNR scenario, the minimum long-term average MSE of the static channel is given as

$$\zeta^{\star}=\begin{cases}\frac{(1-\Lambda_0')\mathrm{Tr}f(\bar{\mathbf{P}}_0)+(2\Lambda_0'-(\Lambda_0')^2)\mathrm{Tr}f^2(\bar{\mathbf{P}}_0)+(\Lambda_0')^2\mathrm{Tr}f^3(\bar{\mathbf{P}}_0)}{1+\Lambda_0'}, & \theta^{\star}=1\\ \frac{(1-\Lambda_0')\left(\sum_{i=1}^{\theta^{\star}+1}\mathrm{Tr}\left(f^i(\bar{\mathbf{P}}_0)\right)(\Lambda_0')^{i-1}-\mathrm{Tr}\left(f(\bar{\mathbf{P}}_0)\right)(\Lambda_0')^{\theta^{\star}-1}\right)}{1-(\Lambda_0')^{\theta^{\star}-1}+(\Lambda_0')^{\theta^{\star}}+(\Lambda_0')^{\theta^{\star}+1}}, & \theta^{\star}>1\end{cases} \tag{2.4.4}$$

where $\theta^\star$ is the optimal policy-switching threshold.

In Proposition 2.1, the optimal θ can be numerically obtained by linear search methods, yielding the minimum estimation MSE. The optimal policy structure in the high SNR scenario shows that if the current estimation quality is bad, i.e., a large q, a retransmission is required; otherwise, a new transmission is a better choice. This is mainly because although a retransmission is more reliable than a new transmission, a successful new transmission leads to a better estimation quality. When the current estimation quality is pretty good, it is reasonable to take the risk of a new transmission to achieve a better estimation, since the estimation quality will not be too bad even if it is failed. However, when the current estimation quality is bad, it is not wise to take the risk since the estimation quality will be even worse if it is failed.

2.4.5 Suboptimal Policy

The optimal policy of the MDP problem in Sec. 2.4.1 does not have a closed-form expression for low-complexity computation. Besides, since the MDP problem has infinitely many states, it has to be approximated by a truncated MDP problem with finite states for numerical evaluation and solved offline. Therefore, we propose a easy-to-compute suboptimal policy, which is a myopic policy that makes decisions simply to maximize the expected instantaneous cost.

Based on (2.4.1) and (2.4.2), the expected next step cost $c'((r,q),a)$ given the current state (r,q) and action a can be derived as

$$\begin{aligned} &c'((r,q),a) \\ &= \begin{cases} g(1)\mathrm{Tr}\left(f^{q+1}(\bar{\mathbf{P}}_0)\right) + (1-g(1))\mathrm{Tr}\left(f(\bar{\mathbf{P}}_0)\right), & a=0, \\ g(r+1)\mathrm{Tr}\left(f^{q+1}(\bar{\mathbf{P}}_0)\right) + (1-g(r+1))\mathrm{Tr}\left(f^{r+1}(\bar{\mathbf{P}}_0)\right), & a=1. \end{cases} \end{aligned} \tag{2.4.5}$$

Then, we have

$$\begin{aligned} & c'((r,q),1) - c'((r,q),0) \\ &= (g(r+1)-g(1))\mathrm{Tr}\left(f^{q+1}(\bar{\mathbf{P}}_0)\right) + (1-g(r+1))\mathrm{Tr}\left(f^{r+1}(\bar{\mathbf{P}}_0)\right) \\ &- (1-g(1))\mathrm{Tr}\left(f(\bar{\mathbf{P}}_0)\right). \end{aligned} \tag{2.4.6}$$

Using the inequality (2.3.12), $c'((r,q),1) - c'((r,q),0) \geq 0$ if and only if (r,q) satisfies

$$\begin{aligned} & \mathrm{Tr}\left(f^{q+1}(\bar{\mathbf{P}}_0)\right) \\ & \leq \frac{(1-g(r+1))\mathrm{Tr}\left(f^{r+1}(\bar{\mathbf{P}}_0)\right) - (1-g(1))\mathrm{Tr}\left(f(\bar{\mathbf{P}}_0)\right)}{g(1)-g(r+1)}. \end{aligned} \tag{2.4.7}$$

Thus, we have the following result.

Proposition 2.2. The myopic policy of problem (2.3.15) is

$$a = \begin{cases} 0 & \text{if the condition (2.4.7) is satisfied,} \\ 1 & \text{otherwise.} \end{cases} \tag{2.4.8}$$

It can be proved that the suboptimal policy in Proposition 2.2 is also a switching-type policy. *Due to its simplicity, the suboptimal policy–which, unlike the optimal policy, does not need any iteration for policy calculation–can be applied as an online decision algorithm.* In Sec. 2.6, we will show that the performance of the suboptimal policy is close to the optimal one for practical system parameters. The detailed computation-complexity analysis will be given in Sec. 2.6 as well.

2.5 Optimal Policy: Markov Channel

In this section, we investigate the sensor's optimal transmission control policy in the Markov channel.

2.5.1 MDP Formulation

We also formulate the problem as a MDP.

1) The state space is defined as $\mathbb{S} \triangleq \{(\mathbf{\Omega}, q, \Xi) : \mathbf{\Omega} \in \mathbb{N}_0^B \backslash \mathbf{0}, q \in \mathbb{N}, \Xi \in \{1, 2, \cdots, B\}\}$.

2) The action space is defined as $\mathbb{A} \triangleq \{0, 1\}$.

3) Let $s = (\mathbf{\Omega}, q, \Xi)$ and $s' = (\mathbf{\Omega}', q', \Xi')$ denote the current and next state, respectively. The transition probability can be written as

$$P(s'|s,a) = \begin{cases} p_{\Xi,\Xi'}(1-\tilde{g}(\mathbf{0},\Xi)) & \text{if } a=0, s'=(\mathbf{1}_\Xi, 1, \Xi'), \\ p_{\Xi,\Xi'}\tilde{g}(\mathbf{0},\Xi) & \text{if } a=0, s'=(\mathbf{1}_\Xi, q+1, \Xi'), \\ p_{\Xi,\Xi'}\left(1-\tilde{g}(\mathbf{\Omega},\Xi)\right) & \text{if } a=1, s'=(\mathbf{\Omega}+\mathbf{1}_\Xi, \|\mathbf{\Omega}\|_1+1, \Xi'), \\ p_{\Xi,\Xi'}\tilde{g}(\mathbf{\Omega},\Xi) & \text{if } a=1, s'=(\mathbf{\Omega}+\mathbf{1}_\Xi, q+1, \Xi'), \\ 0 & \text{otherwise.} \end{cases} \tag{2.5.1}$$

where $\|\mathbf{\Omega}\|_1 \triangleq \sum_{i=1}^B r_i$.

4) The one-stage cost is given in (2.4.2).

2.5.2 Optimal Policy: Condition of Existence

Inspired by the static channel scenario, we derive the following condition under which the long term average MSE can be bounded.

Theorem 2.3. For a Markov channel, there exists a stationary and deterministic optimal policy π^* of problem (2.3.15), if the following condition holds:

$$\rho\left(\mathbf{\Pi\Lambda}\right)\rho^2(\mathbf{A}) < 1, \tag{2.5.2}$$

where $\mathbf{\Pi}$ is defined in (2.2.4), and

$$\mathbf{\Lambda} \triangleq \operatorname{diag}\left\{\Lambda_1, ..., \Lambda_B\right\}, \tag{2.5.3}$$

and Λ_i is the largest packet error probability of a retransmission when the channel power gain is u_i defined in (2.3.13).[4]

Proof. See Appendix A. □

Remark 2.4. *It is interesting to see that when retransmissions have very high reliability, i.e., $\Lambda_i \to 0$ $\forall i = 1, \cdots, B$, the eigenvalues of the matrix $\mathbf{\Pi\Lambda}$ approach zero, and thus the left-hand side of* (2.5.2) *is much less than one and the remote estimation system can be stabilized.*

The stability regions of a two-state Markov channel in terms of Λ_1 and Λ_2 with different $\rho^2(\mathbf{A})$ are illustrated in Fig. 2.4, where $\mathbf{\Pi} = \begin{bmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{bmatrix}$. We see that a larger $\rho^2(\mathbf{A})$ results in a smaller stability region.

2.5.3 Optimal Policy: The Structure

The optimal policy in the Markov channel also has a switching structure in the state space.

Theorem 2.4. (i) if $\pi^*(\mathbf{\Omega}, q, \Xi) = 0$, then $\pi^*(\mathbf{\Omega} + z\mathbf{1}_i, q, \Xi) = 0, \forall 1 \leq i \leq B$;

(ii) if $\pi^*(\mathbf{\Omega}, q, \Xi) = 1$, then $\pi^*(\mathbf{\Omega}, q + z, \Xi) = 1$, where z is any positive integer.

Proof. The proof is similar to that of Theorem 2.2 and is omitted here. □

[4]Note that Theorem 2.3 only works for a Markov channel with finite state and cannot be extended to Markov channels with infinite countable states. The analysis of the latter case involves multiplication and eigenvalue of infinite matrices, which need a general theory concerning operator on Hilbert spaces in functional analysis and is beyond the scope of our work.

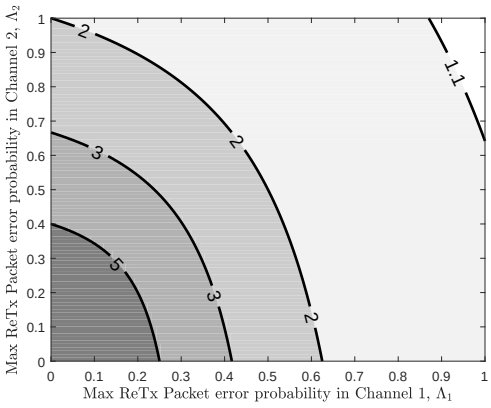


Figure 2.4: The stability regions in terms of Λ_1 and Λ_2 with $\rho^2(\mathbf{A}) = 1.1,\ 2,\ 3$ and 5, respectively.

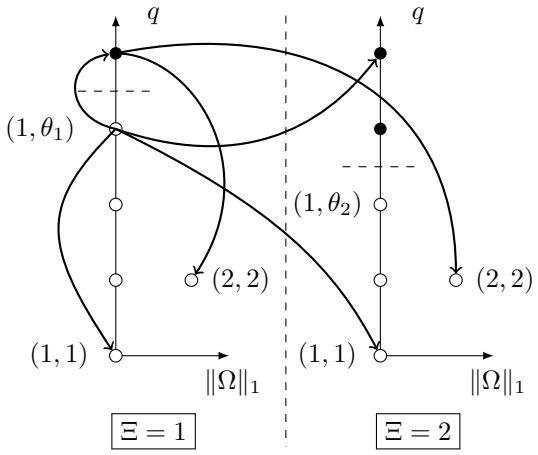


Figure 2.5: Illustration of the state space and state transitions of the optimal transmission control policy in the high SNR scenario, where the channel is a 2-state Markov channel, $\theta_1 = 4$ and $\theta_2 = 3$.

2.5.4 Optimal Policy: A Special Case

For the high SNR scenario, we assume that a retransmission is always successful. Thus, the state transition probability (2.5.1) does not depend on all the individual elements of the historical channel-state vector $\mathbf{\Omega}$, and we can simply combine the states in $\mathbb{S}$ by $\|\mathbf{\Omega}\|_1$ as $s = (\|\mathbf{\Omega}\|_1, q, \Xi)$ to reduce the state space.

Similar to the static channel scenario, the state space of the optimal policy can be further reduced as $\mathbb{S}' = \{(2,2,\Xi), (1,q,\Xi), \forall q = 1,2,...,\theta_{\max}+1, \Xi = 1,2,...,B\}$, and the optimal policy for states $(1,q,\Xi), \forall q \in \{\theta_\Xi + 1, \cdots, \theta_{\max}+1\}$ is $a = 1$, where $\theta_{\max} \triangleq \max \theta_\Xi$, and the other states have the action $a = 0$, as illustrated in Fig. 2.5. Different from the static channel scenario, the optimal policy for the B-state Markov channel has a set of parameters, i.e., $\{\theta_1, \cdots, \theta_B\}$, to be optimally designed.

We can reorder the three dimensional states as a $B \times (\theta_{\max}+2)$ state (column) vector, $\mathbf{b}$, and the states $(2,2,\Xi)$ and $(1,q,\Xi)$ are the $(1+(\Xi-1)(\theta_{\max}+2))$th and $(1+q+(\Xi-1)(\theta_{\max}+2))$th elements of $\mathbf{b}$, respectively. Using the state transition probability (2.5.1) and the transition rule of the special case as illustrated in Fig. 2.5, the matrix of the state transition probability can be written as

$$\mathbf{M} = \left[\begin{array}{c|c|c|c} \mathbf{p}_1 \otimes \mathbf{M}_1 & \mathbf{p}_2 \otimes \mathbf{M}_2 & \cdots & \mathbf{p}_B \otimes \mathbf{M}_B \end{array} \right], \tag{2.5.4}$$

where the $\otimes$ is the Kronecker product operator, $\mathbf{p}_i$ is the ith column of $\mathbf{\Pi}$ defined in (2.2.4), and $\mathbf{M}_i$ is the

$$\mathbf{M}_i = \left[\begin{array}{c|c} \mathbf{E}_i & \mathbf{F}_i \\ \hline [0]_{(\theta_{\max}-\theta_i)\times(\theta_i+2)} & [0]_{(\theta_{\max}-\theta_i)\times(\theta_{\max}-\theta_i)} \end{array} \right], \tag{2.5.5}$$

$$\mathbf{E}_i = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1-\Lambda_i' & 1-\Lambda_i' & 1-\Lambda_i' & \cdots & 1-\Lambda_i' & 0 \\ 0 & \Lambda_i' & 0 & 0 & \cdots & 0 \\ \Lambda_i' & 0 & \Lambda_i' & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \Lambda_i' & 0 \end{bmatrix}_{(\theta_i+2)\times(\theta_i+2)}$$

$$\mathbf{F}_i = \begin{bmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{(\theta_i+2)\times(\theta_{\max}-\theta_i)} . \tag{2.5.6}$$

Based on the stochastic matrix (2.5.4), we can calculate the steady state distribution with a given set of policy-switching parameters. By numerically optimizing $\{\theta_1, \cdots, \theta_B\}$, we have the following result.

Proposition 2.3. In the high SNR scenario, the minimum long-term average MSE

of the Markov channel is given as

$$\zeta^\star = \frac{\mathbf{c}^T\mathbf{e}}{\|\mathbf{e}\|_1}, \tag{2.5.7}$$

where

$$\mathbf{c} = \underbrace{[1,\cdots,1]^T}_{B} \otimes [\mathrm{Tr}f^2(\bar{\mathbf{P}}_0), \mathrm{Tr}f(\bar{\mathbf{P}}_0), \mathrm{Tr}f^2(\bar{\mathbf{P}}_0), \cdots, \mathrm{Tr}f^{\theta^\star_{\max}+1}(\bar{\mathbf{P}}_0)]^T,$$

and $\mathbf{e}$ is a null-space vector of $(\mathbf{M}^\star - \mathbf{I})$ with non-negative values, and here $\mathbf{I}$ is the $B(\theta^\star_{\max}+2)$ by $B(\theta^\star_{\max}+2)$ identity matrix.

2.6 Numerical Results

2.6.1 Delay-Optimal Policy: A Benchmark

We also consider a delay-optimal policy based on the HARQ protocol in [77], as the benchmark of the proposed optimal policy. We use the average AoI to measure the delay of the system. Therefore, similar to the MSE-optimization problem (2.3.15), the delay optimization problem is formulated as

$$\min_{\pi} \limsup_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\left[q_k\right]. \tag{2.6.1}$$

This problem can also be converted to a MDP problem with the same state space, action space and state transition function as presented in Sec. 2.4.1. The one-stage cost in terms of delay is

$$c((r,q),a) = q. \tag{2.6.2}$$

Comparing (2.6.2) with (2.4.2), we see that the cost function of the delay-optimal policy is a linear function of q, while it grows exponentially fast with q in the optimal policy. Thus, these two policies should be very different and their performance will

be compared in the following section.

2.6.2 Simulation and Policy Comparison

In the remainder of the section, we present the numerical results of the optimal policies for static and Markov channels in Sec. 2.4 and Sec. 2.5, respectively, and their performance. Also, we numerically compare these optimal policies with the benchmark policy in Sec. 2.6.1. Unless otherwise stated, we consider CC-HARQ, and we set $\mathsf{SNR} = 10$ dB, $L = 100$, $R = 4$, $\mathbf{A} = \begin{bmatrix} 2.4 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $\mathbf{Q}_w = I$, $\mathbf{Q}_v = 1$, and thus $\rho^2(\mathbf{A}) = 1.8385^2$, $\bar{\mathbf{P}}_0 = \begin{bmatrix} 2.5548 & -1.6233 \\ -1.6233 & 1.6179 \end{bmatrix}$.

The packet error probabilities for a new transmission and a retransmission of the CC/IR-HARQ protocol are based on taking the approximation (2.2.5) into (2.3.4) and (2.3.5), respectively. We use the relative value iteration algorithm in a MDP toolbox in MATLAB [78] to solve the MDP problems in Sections 2.4, 2.5 and 2.6.1.

Static Channel

Policy Comparison. To solve the MDP problem with an infinite state space, the unbounded state space $\mathbb{S}$ is truncated as $\{(r, q) : 1 \leq r \leq q \leq 20\}$ to enable the evaluation. We set the channel power gain $h = 2$. Using Theorem 2.1, we can verify that the CC/IR-HARQ based optimal policy exists. Fig. 2.6 shows different policies within the truncated state space. In Fig. 2.6(a), we see that in line with Theorem 2.2, the optimal policy of CC-HARQ protocol is a switching-type one, where the actions of the states close to the states with $r = q$ are equal to zero; i.e., new transmissions

are required. Also, we see that the myopic policy plotted in Fig. 2.6(d) is a good approximation of the optimal one within the truncated state space. However, the delay-optimal policy plotted in Fig. 2.6(c) is very different from the previous ones, where more states have the action of new transmission. Therefore, HARQ-based retransmissions are more important to reduce the estimation MSE than the delay. Fig. 2.6(b) presents the optimal policy of the IR-HARQ protocol, which is identical with that of CC-HARQ in Fig. 2.6(a). This is because when the channel power gain is high, e.g., $h = 2$, both IR- and CC-HARQ can provide sufficiently high retransmission reliability and the transmission control policy are the same.

Moreover, in Figs. 2.7(a)-(b), we calculate the optimal policies of different state transition matrices $\mathbf{A}$. We see that the optimal policy with a larger $\rho(\mathbf{A})$ has more states in the state space $\mathbb{S}$ with the action of new transmission. This is consistent with the intuition that if the system changes strongly per step, it would be preferable to use fresh information, whereas if the change is mild, more retransmissions can be allowed.

Inspired by Theorem 2.2, we investigate other suboptimal policies with switching structures within the state space. To be specific, we consider a type of suboptimal policy named α policy. For an α policy, the state space $\mathbb{S}$ is divided by a straight line passing through the origin with the slope α, where $\alpha \geq 1$. Thus, the actions of an α policy in the state space are given as

$$a = \begin{cases} 1, \text{ if } q > \alpha r, \text{ and } (q, r) \in \mathbb{S}, \\ 0, \text{ if } r \leq q \leq \alpha r, \text{ and } (q, r) \in \mathbb{S}, \end{cases}$$

Figs. 2.7(c)-(d) show α policies with different parameters, where $\mathbf{A} = [2.4, 0.2; 0.2, 0.8]$. It is clear that the policy with $\alpha = 1$ is similar to the myopic policy in Fig. 2.6(d).

Performance Comparison. Based on the above numerically obtained policies and

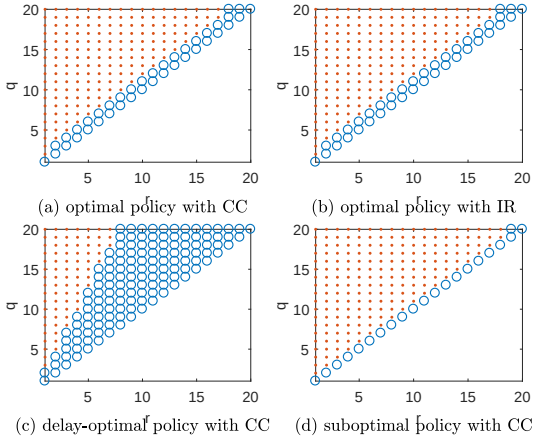


Figure 2.6: Different transmission control policies, where 'o' and '·' denote $a = 0$ and $a = 1$, respectively.

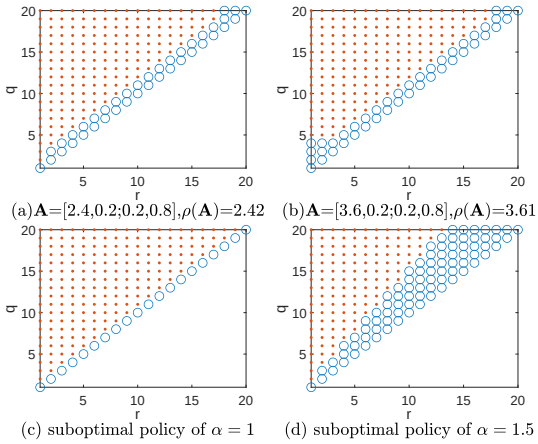


Figure 2.7: Optimal policies with different $\mathbf{A}$ and suboptimal polices with different α.

the policy with the standard ARQ, i.e., the one without retransmission (see Sec. 2.3), we further evaluate their performances in terms of the long-term average MSE using (2.2.10). We run the remote estimation process with 10^4 time slots and set the initial value of $\mathbf{P}_k$ as $\mathbf{P}_0 = f(\bar{\mathbf{P}}_0) = \begin{bmatrix} 14.2218 & -1.6966 \\ -1.6966 & 1.6179 \end{bmatrix}$. Also, we set $\mathrm{Tr}(\mathbf{P}_0) = 15.8$ as the *performance baseline*, as $\mathrm{Tr}(\mathbf{P}_0) \leq \mathrm{Tr}(\mathbf{P}_k)$, $\forall k$.

Fig. 2.8 plots the average MSE versus the simulation time K, using different transmission control policies. Our simulation shows that the conventional non-retransmission policy has an *unbounded average MSE*, which is not shown in the figure due to the ultra fast growth rate. However, we show that the average MSEs of different HARQ-based policies quickly converge to bounded steady state values. Therefore, the proposed HARQ-based policy can significantly improve the estimation quality against the conventional policy. Also, we see that the performance of the myopic policy is very close to the optimal one. Given the performance baseline, the optimal policy gives a 10% MSE reduction of the delay-optimal policy, which demonstrates the superior

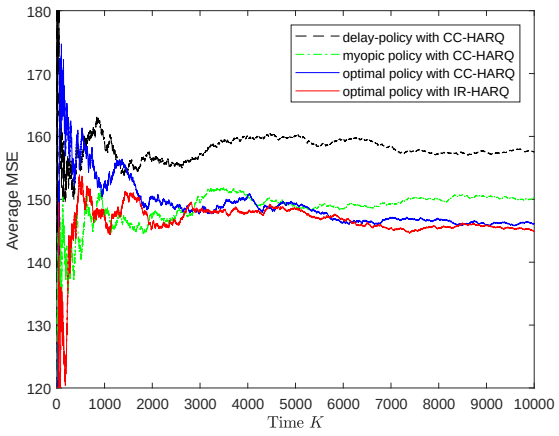


Figure 2.8: Average MSE with different policies of the static channel.

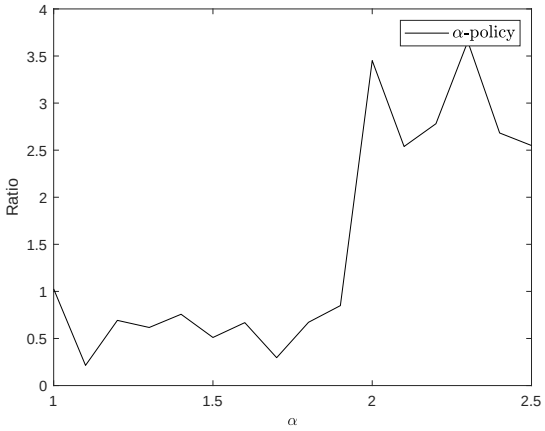


Figure 2.9: Comparison of α policies and the myopic policy.

performance of the proposed optimal policy.

In Fig. 2.9, we compare α policies with the myopic policy with CC-HARQ. Specifically, to measure the relative performance gain provided by the suboptimal policies, we use the ratio of the average MSE gap between the suboptimal and the optimal policies to the gap between the myopic and the optimal policies. Thus, a ratio less than one indicates a better performance achieved by the suboptimal policy. From Fig. 2.9, we see that some α policies with $\alpha \in (1, 1.8)$ are able to achieve better performances than the myopic policy, and the policies with $\alpha \geq 2$ induce much larger average MSEs than the myopic one. Therefore, the α policy with a small parameter can be a better choice than the myopic policy.

Markov Channel

We consider a two state Markov channel with the channel power gains $u_1 = 2$ and $u_2 = 1$. The matrix of channel state transition probability $\mathbf{\Pi} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$. Using Theorem 2.3, we can verify that the CC/IR-HARQ based optimal policy exists. To

Table 2.1: Computation complexity of the policies in the static and Markov channels.

	Computation complexity [79]					The number of convergence steps, K		
	M	N	Optimal	Myopic	Delay-optimal	Optimal	Myopic	Delay-optimal
Static channel	2	210	$\mathcal{O}(MN^2K)$	$\mathcal{O}(MN)$	$\mathcal{O}(MN^2K)$	58	1	41
Markov channel	2	328	$\mathcal{O}(MN^2K)$	$\mathcal{O}(MN)$	$\mathcal{O}(MN^2K)$	90	1	30

solve the MDP problem in the Markov channel scenario with an infinite state space, the unbounded state space $\mathbb{S}$ is truncated as $\{(r_1, r_2, q, \Xi) : 0 \leq r_1 \leq 4, 0 \leq r_2 \leq 4, 1 \leq (r_1 + r_2) \leq q \leq 10, \Xi = 1, 2\}$ to enable the evaluation, where $(r_1, r_2) \triangleq \mathbf{\Omega}$ is the state vector of the historical channel states. Figs. 2.10 and 2.11 show the optimal transmission control policy under channel 1 ($h = 2$) and 2 ($h = 1$), respectively. We can see the switching structure of the optimal policy. Also, we see that new transmission occurs more often in the good channel than in the bad channel.

We also calculate the myopic policy and delay-optimal policy of the Markov channel, and the computation complexity of these policies together with the ones of the static channel are listed in Table 2.1. Note that the myopic policy of the Markov channel case is similar to that of the static channel case in Sec. 2.4.5, which makes simply the decision to maximize the expected instantaneous cost. The details are omitted due to space limitations. We see that the numbers of convergence steps for calculating these policies are less than 100, and the optimal policy has a larger number of convergence steps than the delay-optimal policy and hence a higher computation complexity.

Performance Comparison. We evaluate the non-retransmission policy, the CC/IR-HARQ based optimal policy, the CC-HARQ based myopic policy and the delay-optimal policy in terms of the long-term average MSE using (2.2.10). We run the

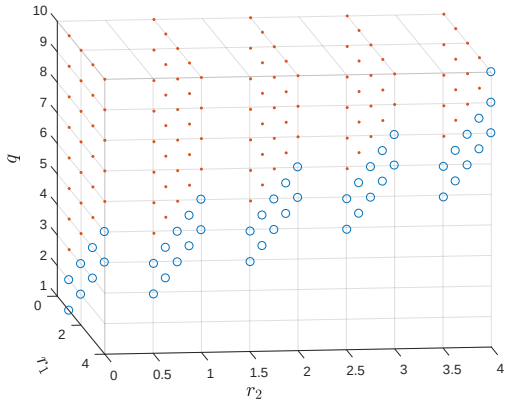


Figure 2.10: Optimal transmission control policy at channel 1, where 'o' and '·' denote $a = 0$ and $a = 1$, respectively.

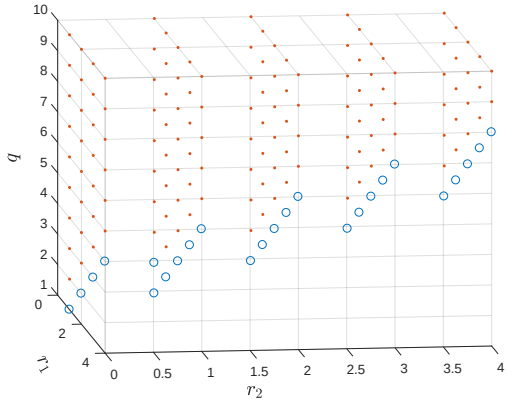


Figure 2.11: Optimal transmission control policy at channel 2, where 'o' and '·' denote $a = 0$ and $a = 1$, respectively.

remote estimation process with 10^4 time slots and set the initial value of $\mathbf{P}_k$ as $\mathbf{P}_0$.

Fig. 2.12 plots the average MSE versus the simulation time K, using different transmission control policies. Our simulation shows that the non-retransmission policy has an *unbounded average MSE*. We show that the average MSEs of different policies quickly converge to bounded steady state values. Therefore, the proposed HARQ-based policy can also significantly improve the estimation quality against the conventional policy in the Markov channel scenario. We see that the performance of the myopic policy is very close to the optimal one. Given the performance baseline, the optimal policy reduces MSE by 33% for the delay-optimal policy. Also, unlike the static channel scenario, we see that the IR-HARQ based optimal policy significantly reduces the average MSE by 87% compared to the CC-HARQ based optimal policy. This is because IR-HARQ can provide much better retransmission reliability than CC-HARQ, especially when the channel quality is bad.

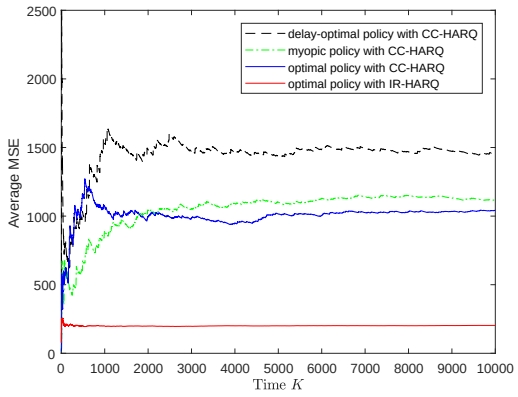


Figure 2.12: Average MSE with different policies of the Markov channel

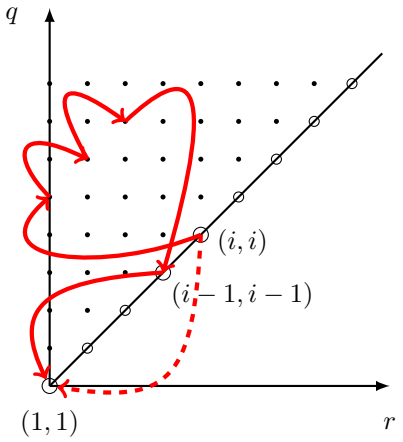


Figure 2.13: Two of the possible first passage paths from (i,i) to $(1,1)$, i.e., $(i,i) \to (1,1)$ and $(i,i) \to (1,i+1) \to (2,i+2) \to (3,i+3) \to (i-1,i-1) \to (1,1)$, where $i = 5$.

2.7 Conclusions

We have proposed and optimized a HARQ-based remote estimation protocol for real-time applications. We derived a sufficient condition of the existence of a stationary and deterministic optimal policy that stabilizes the remote estimation system and minimizes the MSE. Also, we proved that the optimal policy has a switching structure, and accordingly derived a low-complexity suboptimal policy. Our results have shown that the optimal policy can significantly reduce the estimation MSE for some practical settings.

Chapter 3

Optimal Downlink–Uplink Scheduling of Wireless Networked Control for Industrial IoT

3.1 Introduction

Driven by recent development of mission-critical Industrial Internet of Things (IIoT) applications [80–82] and significant advances in wireless communications, networking, computing, sensing and control [83–86], wireless networked control systems (WNCSs) have recently emerged as a promising technology to enable reliable and remote control of industrial control systems. They have a wide range of applications in factory automation, process automation, smart grid, tactile Internet and intelligent transportation systems [87–91]. Essentially, a WNCS is a spatially distributed control system consisting of a plant with dynamic states, a set of sensors, a remote controller, and a set of actuators.

A WNCS has two types of wireless transmissions, i.e., the sensor's measurement transmission to the controller and the controller's command transmission to the actuator. The packets carrying plant-state information and control commands can

be lost, delayed or corrupted during their transmissions. Most existing research in WNCS adopted a *separate design approach*, i.e., either focusing on remote plant-state estimation or remote plant-state control through wireless channels. In [63] and [92], the optimal policies of remote plant-state estimation with a single and multiple sensors' measurements were proposed, respectively. Some advanced remote plant-state control methods were investigated to overcome the effects of transmission delay [93] and detection errors [94, 95].

The fundamental *co-design* problem of a WNCS in terms of the optimal remote estimation and control were tackled in [12]. Specifically, the controller was ideally assumed to work in a *full-duplex* (FD) mode that can simultaneously receive the sensor's packet and transmit its control packet by default. The scheduling of the sensor's and controller's transmissions has rarely been considered in the area of WNCSs, although transmission scheduling is actually an important issue for practical wireless communication systems [96–98]. Moreover, although an FD system can improve the spectrum efficiency, it faces challenges such as balancing between the performance of self-interference cancellation, device cost and power consumption, and may not be feasible in practical systems [59].

In this chapter, we focus on the design of a WNCS using a practical HD controller and optimze the uplink-downlink transmission scheduling of WNCS. The remainder of the chapter is organized as follows: In Section 3.2, we introduce a WNCS with an HD controller. In Section 3.3, we analyze the estimation-error covariance and the plant-state covariance of the WNCS and formulate an uplink-downlink transmission-scheduling problem. In Sections 3.4 and 3.5, we analyze and solve the transmission-scheduling problem for one-step and multi-step controllable WNCSs, respectively, in

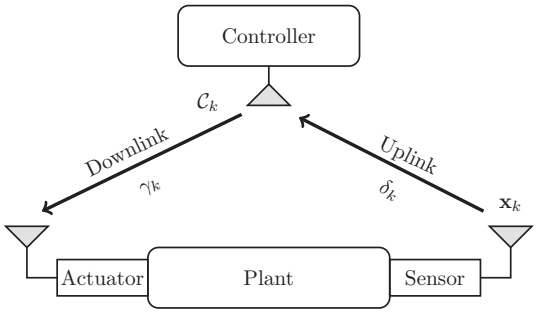


Figure 3.1: The system architecture.

the static channel scenario. In Section 3.6, we extend the design to the fading channel scenario. Section 3.7 numerically evaluates the performance of WNCSs with different transmission-scheduling policies. Finally, Section 3.8 concludes the work.

3.2 System Model

We consider a discrete-time WNCS consisting of a dynamic plant with multiple states, a wireless sensor, an actuator, and a remote controller, as illustrated in Fig. 3.1. In general, the sensor measures the states of the plant and sends the measurements to the remote controller through a wireless uplink (i.e., *sensor-controller*) channel. The controller generates control commands based on the sensor's feedback and sends the commands to the actuator through a wireless downlink (i.e., *controller-actuator*) channel. The actuator controls the plant using the received control commands.

3.2.1 Dynamic Plant

The plant is a linear time invariant (LTI) discrete-time system modeled as [36, 94, 95]

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k, \forall k \tag{3.2.1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the plant-state vector at time k, $\mathbf{u}_k \in \mathbb{R}^m$ is the control input applied by the actuator, and $\mathbf{w}_k \in \mathbb{R}^n$ is the plant disturbance independent of $\mathbf{x}_k$, and is a discrete-time zero-mean Gaussian white noise process with the covariance matrix $\mathbf{R} \in \mathbb{R}^{n\times n}$. $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$ are the system-transition matrix and the control-input matrix, respectively, which are constant. The discrete time step of the system (3.2.1) is T_0, i.e., the plant states keep constant during a time slot of T_0 and change slot-by-slot.

We assume that the plant is an unstable system [63, 94], i.e., the *spectral radius* of $\mathbf{A}$, $\rho(\mathbf{A})$, is larger than 1. In other words, the plant-state vector $\mathbf{x}_k$ grows unbounded without the control input, i.e., $\mathbf{u}_k = \mathbf{0}, \forall k$.

We consider the long-term *average (quadratic) cost* of the dynamic plant defined as (see e.g. [12, 94])

$$J = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[\mathbf{x}_k^\top \mathbf{Q} \mathbf{x}_k\right] = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathrm{Tr}\left(\mathbf{Q}\mathbf{P}_k\right), \tag{3.2.2}$$

where $\mathbf{Q}$ is a symmetric positive semidefinite matrix, and $\mathbf{P}_k$ is the plant-state covariance, defined as

$$\mathbf{P}_k \triangleq \mathbb{E}\left[\mathbf{x}_k \mathbf{x}_k^\top\right]. \tag{3.2.3}$$

Definition 3.1 (Closed-loop Stability [12, 94])**.** The plant (3.2.1) is stabilized by the sequence $\{\mathbf{u}_k\}$ if the average cost function (3.2.2) is bounded.

3.2.2 HD Operation of the Controller

We assume that the controller is an HD device, and thus it can either receive the sensor's measurement or transmit its control command to the actuator at a time. Let $a_k \in \{1, 2\}$ be the controller's transmission-scheduling variable in time slot k. The

sensor's or the controller's transmission is scheduled in time slot k if $a_k = 1$ or 2, respectively.

The sensor measures the plant states at the beginning of each time slot. The measurement is assumed to be perfect [36, 94, 95]. We use δ_k to indicate the success of the sensor's transmission in time slot k. Thus, $\delta_k = 1$ if the sensor is scheduled to send a packet carrying its measurement to the controller in time slot k (i.e., $a_k = 1$) and the transmission is successful, and $\delta_k = 0$ otherwise.

The controller generates a control-command-carrying packet at the beginning of each time slot. Similarly, we use γ_k to indicate the success of the controller's transmission in time slot k. Thus, $\gamma_k = 1$ if the controller is scheduled to send the control packet to the actuator in time slot k (i.e., $a_k = 2$) and the transmission is successful, and $\gamma_k = 0$ otherwise. We also assume that the controller has a perfect feedback from the actuator indicating successful packet detection [12]. *Thus, the controller knows whether its control command will be applied or not.* We assume that the packets in both the sensor-to-controller and controller-to-actuator channels have the same packet length and that this is less than T_0 [12, 63].

3.2.3 Wireless Channel

We consider both the static channel and the fading channel scenarios of the WNCS. The static channel scenario is for IIoT applications with low mobilities, e.g., process control of chemical and oil-refinery plants, while the fading channel scenario is for high mobility applications, e.g., motion control of automated guided vehicles in warehouses.

For the static channel scenario, we assume that the *packet-error probabilities* of

the uplink (sensor-controller) and downlink (controller-actuator) channels are p_s and p_c, respectively, which do not change with time, where $p_s, p_c \in (0, 1)$.

For the fading channel scenario, we adopt a practical finite-state Markov channel model, which captures the inherent property of practical fading channels for which the channel states change with memories [99]. It is assumed that the uplink channel and the downlink channel have B_s and B_c states, respectively, and the packet loss probability of the ith channel state of the uplink channel and the jth channel state of the downlink channel are ω_i and ξ_j, respectively. The matrices of the channel state transition probabilities of the uplink and downlink channels are given as

$$\mathbf{D}_s \triangleq \begin{bmatrix} d^s_{1,1} & \cdots & d^s_{B_s,1} \\ \vdots & \ddots & \vdots \\ d^s_{1,B_s} & \cdots & d^s_{B_s,B_s} \end{bmatrix}, \tag{3.2.4}$$

and

$$\mathbf{D}_c \triangleq \begin{bmatrix} d^c_{1,1} & \cdots & d^c_{B_c,1} \\ \vdots & \ddots & \vdots \\ d^c_{1,B_c} & \cdots & d^c_{B_c,B_c} \end{bmatrix}, \tag{3.2.5}$$

respectively. The packet-error probabilities of the uplink and downlink channels at time k are $p_{s,k}$ and $p_{c,k}$, respectively, where $p_{s,k} \in \{\omega_1, \cdots, \omega_{B_s}\}$ and $p_{c,k} \in \{\xi_1, \cdots, \xi_{B_c}\}$.

3.2.4 Optimal Plant-State Estimation

At the beginning of time slot $(k + 1)$, before generating a proper control command, the controller needs to estimate the current states of the plant, $\mathbf{x}_{k+1}$, using the previously received sensor's measurement and also the implemented control input based on the dynamic plant model (3.2.1). The optimal plant-state estimator is given

as [63]

$$\hat{\mathbf{x}}_{k+1} = \begin{cases} \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, & a_k = 1, \delta_k = 1, \\ \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k, & \text{otherwise.} \end{cases} \tag{3.2.6}$$

3.2.5 v-Step Predictive Plant-State Control

As the transmission between the controller and the actuator is unreliable, the actuator may not successfully receive the controller's packet containing the current control command. To provide robustness against packet failures, we consider a predictive control approach [100]. In general, the controller sends both the current command and the predicted future commands to the actuator at each time. If the current command-carrying packet is lost, the actuator will apply the previously received command that was predicted for the current time slot. The details of the predictive control method are given below.

The controller adopts a conventional linear predictive control law [94], which generates a sequence of v control commands including one current command and $(v-1)$ predicted future commands in each time slot k as

$$\mathcal{C}_k = \big[\mathbf{K}\hat{\mathbf{x}}_k, \underbrace{\mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})\hat{\mathbf{x}}_k, \cdots, \mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})^{v-1}\hat{\mathbf{x}}_k}_{(v-1) \text{ predicted control commands}}\big], \tag{3.2.7}$$

where the constant v is the length of predictive control, and the constant $\mathbf{K} \in \mathbb{R}^{m \times n}$ is the controller gain, which satisfies the condition that[1]

$$\rho(\mathbf{A}+\mathbf{B}\mathbf{K}) < 1. \tag{3.2.8}$$

Such predicitve control law comes from the intuiative combination of linear feedback law ($\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_k$) and predictive estimation process ($\hat{\mathbf{x}}_{k+1} = (\mathbf{A}+\mathbf{B}\mathbf{K})\hat{\mathbf{x}}_k$).

If time slot k is scheduled for the controller's transmission, the controller sends

[1] If (3.2.8) is not satisfied, the plant (3.2.1) can never be stabilized even if the uplink and downlink transmissions are always perfect see e.g., [12, 50].

a packet containing v control commands $\mathcal{C}_k$ to the actuator. Note that in most communication protocols, the minimum packet length is longer than the time duration required for transmitting a single control command [100], and thus it is wise to send multiple commands in the one packet without increasing the packet length.

The actuator maintains a command buffer of length v, $\mathcal{U}_k \triangleq \left[\mathbf{u}_k^0, \mathbf{u}_k^1, \cdots, \mathbf{u}_k^{v-1}\right]$. If the current controller's packet is successfully received, the actuator resets the buffer with the received command sequence; otherwise, the buffer shifts one step forward, i.e.,

$$\mathcal{U}_k = \begin{cases} \mathcal{C}_k, & a_k = 2, \gamma_k = 1, \\ \left[\mathbf{u}_{k-1}^1, \mathbf{u}_{k-1}^2, \cdots, \mathbf{u}_{k-1}^{v-1}, \mathbf{0}\right], & \text{otherwise.} \end{cases} \tag{3.2.9}$$

The actuator always applies the first command in the buffer to the plant. Thus, the actuator's control input in time slot k is

$$\mathbf{u}_k \triangleq \mathbf{u}_k^0. \tag{3.2.10}$$

To indicate the number of passed time slots from the last successfully received control packet, we define the *control-quality indicator* of the plant in time slot k as

$$\eta_k = \begin{cases} 1, & a_k = 2, \gamma_k = 1, \\ \eta_{k-1} + 1, & \text{otherwise.} \end{cases} \tag{3.2.11}$$

Specifically, η_{k-1} is the number of the time slots passed from the most recent controller's successful transmission to the current time slot k.

From (3.2.7), (3.2.9), (3.2.10) and (3.2.11), the control input can be rewritten as

$$\mathbf{u}_k = \begin{cases} \mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})^{\eta_k - 1}\hat{\mathbf{x}}_{k+1-\eta_k}, & \text{if } \eta_k \leq v, \\ \mathbf{0}, & \text{if } \eta_k > v. \end{cases} \tag{3.2.12}$$

To better explain the intuition behind the predictive control method (3.2.7), (3.2.9) and (3.2.10), we give an example below.

Example 3.1. *Assume that a sequence of the controller's commands is successfully*

received in time slot k and the actuator will not receive any further commands in the following $v-1$ time slots. Consider an ideal case that the estimation is accurate in time slot k, i.e., $\hat{\mathbf{x}}_k = \mathbf{x}_k$, and the plant disturbance $\mathbf{w}_k = \mathbf{0}, \forall k$. Taking (3.2.12) *into* (3.2.1)*, the plant-state vector at $(k+j), \forall j \leq v$ can be derived as*

$$\mathbf{x}_{k+j} = (\mathbf{A} + \mathbf{BK})^j \mathbf{x}_k. \tag{3.2.13}$$

Therefore, if the controller gain $\mathbf{K}$ is chosen properly and makes the spectral radius of $(\mathbf{A} + \mathbf{BK})$ less than 1, each state in $\mathbf{x}_k$ can approach zero gradually within the v steps even without receiving any new control packets.

In this work, we mainly focus on two types of plants applying the predictive control method as follows.

Case 1: The controller gain $\mathbf{K}$ satisfies the condition that

$$\mathbf{A} + \mathbf{BK} = \mathbf{0}. \tag{3.2.14}$$

This case is named as the *one-step controllable* case [17], since once a control packet is received successfully, the plant-state vector $\mathbf{x}_k$ can be driven to zero in one step in the above mentioned ideal setting, i.e., $\mathbf{x}_{k+1} = \mathbf{0}\mathbf{x}_k = \mathbf{0}$ in (3.2.13). By taking (3.2.14) into (3.2.7), the $(v-1)$ predicted commands are all $\mathbf{0}$, thus the controller only needs to send the current control command to the actuator *without any prediction*, and the length of $\mathcal{U}$ and $\mathcal{C}$, v, is equal to 1.

Case 2: The controller gain $\mathbf{K}$ satisfies the condition that [17]

$$(\mathbf{A} + \mathbf{BK})^v = \mathbf{0}, v > 1. \tag{3.2.15}$$

This case is named as the *v-step controllable* case [17], since the plant state $\mathbf{x}_k$ can be driven to zero in v steps after a successful reception of a control packet in the ideal setting[2], i.e., $\mathbf{x}_{k+v} = \mathbf{0}$ in (3.2.13).

[2]Note that the ideal setting here is only for the explanation of the term of "one-step controllable",

The other cases not satisfying the conditions (3.2.14) or (3.2.15) will also be discussed in the following section.

3.3 Analysis of the Downlink-Uplink Scheduling

As the controller estimates the current plant states and utilizes this estimation to control future ones, we analyze the estimation-error covariance and the plant-state covariance below.

3.3.1 Estimation-Error Covariance

Using (3.2.1) and (3.2.6), the estimation error in time slot $(k+1)$ is obtained as

$$\mathbf{e}_{k+1} \triangleq \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1} = \begin{cases} \mathbf{w}_k, & a_k = 1, \delta_k = 1, \\ \mathbf{A}\mathbf{e}_k + \mathbf{w}_k, & \text{otherwise.} \end{cases} \tag{3.3.1}$$

Thus, we have the updating rule of the estimation-error covariance, $\mathbf{U}_k \triangleq \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^\top]$, as

$$\mathbf{U}_{k+1} \triangleq \mathbb{E}[\mathbf{e}_{k+1}\mathbf{e}_{k+1}^\top] = \begin{cases} \mathbf{R} & a_k = 1, \delta_k = 1, \\ \mathbf{A}\mathbf{U}_k\mathbf{A}^\top + \mathbf{R} & \text{otherwise.} \end{cases} \tag{3.3.2}$$

We define the *estimation-quality indicator* of the plant in time slot k, τ_k, as the number of passed time slots from the last successfully received sensor packet. Then, the state-updating rule of τ_k is obtained as

$$\tau_{k+1} = \begin{cases} 1, & a_k = 1, \delta_k = 1, \\ \tau_k + 1, & \text{otherwise.} \end{cases} \tag{3.3.3}$$

Once a successful sensor transmission occurs (e.g., there exists k' such that $\mathbf{U}_{k'} = \mathbf{R}$), from (3.3.2) and (3.3.3), it can be shown that the estimation-error covariance

while we only consider practical settings in the rest of the chapter.

$\mathbf{U}_k, \forall k \geq k'$ is simply a function of the estimation-quality indicator τ_k, i.e.,

$$\mathbf{U}_k = \mathbf{F}(\tau_k), \tag{3.3.4}$$

where the function $\mathbf{F}(\cdot)$ is defined as

$$\mathbf{F}(\tau) \triangleq \sum_{i=1}^{\tau} \mathbf{A}^{i-1}\mathbf{R}(\mathbf{A}^{\top})^{i-1}, \tau \in \mathbb{N}. \tag{3.3.5}$$

As we focus on the long-term performance of the system, without loss of generality, we assume that $\mathbf{U}_k \in \{\mathbf{F}(1), \mathbf{F}(2), \mathbf{F}(3), \cdots\}$ for all k. From (3.3.3) and (3.3.4), the updating rule of $\mathbf{U}_k$ is obtained as

$$\mathbf{U}_{k+1} = \mathbf{F}(\tau_{k+1}) = \begin{cases} \mathbf{F}(1) & a_k = 1, \delta_k = 1, \\ \mathbf{F}(\tau_k + 1) & \text{otherwise.} \end{cases} \tag{3.3.6}$$

3.3.2 Plant-State Covariance of One-Step Controllable Case

Taking (3.2.11) and (3.2.14) into (3.2.12), the control input of the one-step controllable case can be simplified as

$$\mathbf{u}_k = \begin{cases} \mathbf{K}\hat{\mathbf{x}}_k, & a_k = 2, \gamma_k = 1, \\ \mathbf{0}, & \text{otherwise.} \end{cases} \tag{3.3.7}$$

Substituting (3.3.7) into (3.2.1) and using (3.2.14), the plant-state vector can be rewritten as

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{K}\hat{\mathbf{x}}_k + \mathbf{w}_k = \mathbf{A}\mathbf{e}_k + \mathbf{w}_k, & a_k = 2, \gamma_k = 1, \\ \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, & \text{otherwise.} \end{cases} \tag{3.3.8}$$

Thus, the plant-state covariance, $\mathbf{P}_k$, has the updating rule as

$$\mathbf{P}_{k+1} \triangleq \mathbb{E}[\mathbf{x}_{k+1}\mathbf{x}_{k+1}^{\top}] = \begin{cases} \mathbf{A}\mathbf{U}_k\mathbf{A}^{\top} + \mathbf{R} & a_k = 2, \gamma_k = 1, \\ \mathbf{A}\mathbf{P}_k\mathbf{A}^{\top} + \mathbf{R} & \text{otherwise.} \end{cases} \tag{3.3.9}$$

From (3.3.5), (3.3.6) and (3.3.9), we see that the plant-state covariance $\mathbf{P}_k$ will only take value from the countable infinity set $\{\mathbf{F}(2), \mathbf{F}(3), \cdots\}$ after a successful controller's transmission. Again, as we focus on the long-term performance of the system,

we assume that $\mathbf{P}_k \in \{\mathbf{F}(2), \mathbf{F}(3), \cdots\}$ for all k, without loss of generality.

By introducing the variable $\phi_k \in \{2, 3, \cdots\}$, the plant-state covariance in time slot k can be written as

$$\mathbf{P}_k = \mathbf{F}(\phi_k), \tag{3.3.10}$$

where ϕ_k is the *state-quality indicator* of the plant in time slot k. Note that the state covariance only depends on the state parameter ϕ_k.

From (3.3.9) and (3.3.4), the updating rules of $\mathbf{P}_k$ and ϕ_k in (3.3.10) are given by, respectively, as

$$\mathbf{P}_{k+1} = \mathbf{F}(\phi_{k+1}) = \begin{cases} \mathbf{F}(\tau_k + 1) & a_k = 2, \gamma_k = 1, \\ \mathbf{F}(\phi_k + 1) & \text{otherwise}, \end{cases} \tag{3.3.11}$$

$$\phi_{k+1} = \begin{cases} \tau_k + 1, & a_k = 2, \gamma_k = 1, \\ \phi_k + 1, & \text{otherwise}. \end{cases} \tag{3.3.12}$$

From (3.3.3) and (3.3.12), it is easy to prove that $\phi_k \geq \tau_k, \forall k$.

3.3.3 Plant-State Covariance of v-Step Controllable Case

Taking (3.2.12) into (3.2.1), the plant-state vector is rewritten as

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{K}(\mathbf{A} + \mathbf{B}\mathbf{K})^{\eta_k - 1}\hat{\mathbf{x}}_{k+1-\eta_k} + \mathbf{w}_k, & \text{if } \eta_k \leq v, \\ \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, & \text{if } \eta_k > v. \end{cases} \tag{3.3.13}$$

Using the property (3.2.15), we have the state-updating rule as

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{K}(\mathbf{A} + \mathbf{B}\mathbf{K})^{\eta_{k-1} - 1}\hat{\mathbf{x}}_{k-\eta_{k-1}} + \mathbf{w}_{k-1}. \tag{3.3.14}$$

Unlike the one-step controllable case in (3.3.8), where the current state vector relies on the previous-step estimation, current plant state vector in the v-step controllable case depends on the state estimation η_{k-1} steps ago.

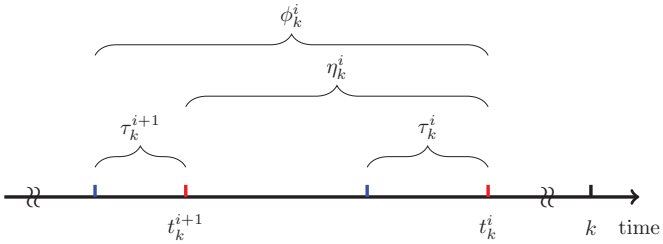


Figure 3.2: Illustration of the state parameters, where red vertical bars denote successful controller transmissions and blue vertical bars denote the most recent successful sensor transmissions prior to the successful controller transmissions.

Inspired by the one-step controllable case (3.3.10), we aim at deriving the plant-state covariance in terms of a set of state parameters. First, we define a sequence of variables, t_k^i, $i = 1, \cdots, v$, where t_k^i is the time-slot index of the ith latest successful controller's transmission prior to the current time slot k, as illustrated in Fig. 3.2. Then, we define the following state parameters

$$\tau_k^i \triangleq \begin{cases} \tau_k, & i = 0, \\ \tau_{t_k^i}, & i = 1, 2, \cdots v, \end{cases} \tag{3.3.15}$$

$$\eta_k^i \triangleq \begin{cases} \eta_{k-1} = k - t_k^1, & i = 0, \\ t_k^i - t_k^{i+1}, & i = 1, 2, \cdots v-1. \end{cases} \tag{3.3.16}$$

Specifically, η_k^i measures the delay between two consecutive controller's successful transmissions; τ_k^i is the estimation-quality indicator of time slot t_k^i. Last, we define the state parameters ϕ_k^i as

$$\phi_k^i \triangleq \eta_k^i + \tau_k^{i+1}, i = 0, \cdots, v-1. \tag{3.3.17}$$

Using the state-transition rules of η_k and τ_k in (3.2.11) and (3.3.3), and the definitions (3.3.15), (3.3.16) and (3.3.17), the state-transition rules of τ_k^i, η_k^i and ϕ_k^i can

be obtained, respectively, as

$$\tau_{k+1}^{i} = \begin{cases} 1, & i = 0, a_k = 1, \delta_k = 1, \\ \tau_k^0 + 1, & i = 0, \text{ otherwise}, \\ \tau_k^{i-1}, & i = 1, \cdots, v-1, a_k = 2, \gamma_k = 1, \\ \tau_k^i, & i = 1, \cdots, v-1, \text{ otherwise}, \end{cases} \tag{3.3.18}$$

$$\eta_{k+1}^{i} = \begin{cases} 1, & i = 0, a_k = 2, \gamma_k = 1, \\ \eta_k^0 + 1, & i = 0, \text{ otherwise}, \\ \eta_k^{i-1}, & i = 1, \cdots, v-1, a_k = 2, \gamma_k = 1, \\ \eta_k^i, & i = 1, \cdots, v-1, \text{ otherwise}, \end{cases} \tag{3.3.19}$$

$$\phi_{k+1}^{i} = \begin{cases} \tau_k^0 + 1, & i = 0, a_k = 2, \gamma_k = 1, \\ \phi_k^i + 1, & i = 0, \text{ otherwise}, \\ \phi_k^{i-1}, & i = 1, \cdots, v-1, a_k = 2, \gamma_k = 1, \\ \phi_k^i, & i = 1, \cdots, v-1, \text{ otherwise}. \end{cases} \tag{3.3.20}$$

Then, we can derive the plant-state covariance in a closed form in terms of the state parameters as follows.

Proposition 3.1. The plant-state covariance $\mathbf{P}_k$ in time slot k is

$$\begin{aligned} \mathbf{P}_k = \mathbf{F}(\phi_k^0) + \sum_{i=0}^{v-2} \mathbf{G} \Bigg(& \sum_{j=0}^{i} \phi_k^j - \sum_{j=0}^{i} \tau_k^{j+1}, \\ & \mathbb{1}(\phi_k^{i+1} > \tau_k^{i+1}) \left(\mathbf{F}(\phi_k^{i+1}) - \mathbf{F}(\tau_k^{i+1}) \right) \Bigg), \end{aligned} \tag{3.3.21}$$

where the summation operator has the property that $\sum_{i=a}^{b}(\cdot) = 0$ if $a > b$, $\mathbf{F}(\cdot)$ is defined in (3.3.5), and

$$\mathbf{G}(x, \mathbf{Y}) \triangleq (\mathbf{A} + \mathbf{BK})^x \mathbf{Y} ((\mathbf{A} + \mathbf{BK})^x)^\top. \tag{3.3.22}$$

Proof. See Appendix A. □

Remark 3.1. *Proposition 3.1 states that the state covariance* $\mathbf{P}_k$ *of a* v*-step controllable plant is determined by* $(2v-1)$ *state parameters, i.e.,* τ_k^i*,* $i = 1, \cdots, v-1$ *and* ϕ_k^i*,* $i = 0, \cdots, v-1$*.*

Remark 3.2. *In practice, it is possible that the plant* (3.2.1) *is* $\bar{v}$*-step controllable, i.e.,* $(\mathbf{A}+\mathbf{BK})^{\bar{v}} = \mathbf{0}$*, where* $\bar{v} > v$*; it is also possible that when the controller gain* $\mathbf{K}$ *is predetermined and fixed, one cannot find* $\bar{v} \in \mathbb{N}$ *such that* $(\mathbf{A}+\mathbf{BK})^{\bar{v}} = \mathbf{0}$*. Moreover, the plant may not be finite-step controllable, i.e., one cannot find a set of* $\mathbf{K}$ *and* $\bar{v} \in \mathbb{N}$ *such that* $(\mathbf{A}+\mathbf{BK})^{\bar{v}} = \mathbf{0}$*. In these cases, where conditions* (3.2.14) *and* (3.2.15) *are not satisfied, we can show that the covariance* $\mathbf{P}_k$ *has incountably infinite values and cannot be expressed by a finite number of state parameters as in Proposition 3.1. Furthermore, the process* $\{\mathbf{P}_k\}$ *is not stationary, making the long-term average cost function* (3.2.2) *difficult to evaluate. However, when* v *is sufficiently large,* $(\mathbf{A}+\mathbf{BK})^v$ *approaches* $\mathbf{0}$ *as* $\rho(\mathbf{A}+\mathbf{BK}) < 1$*. Thus, the plant-state vector in* (B.1.4) *of the proof of Proposition 3.1, obtained by letting* $(\mathbf{A}+\mathbf{BK})^v = \mathbf{0}$*, is still a good approximation of* $\mathbf{x}_k$ *for these cases, and hence Proposition 3.1 can be treated as a countable-state-space approximation of the plant-state covariance.*

3.3.4 Problem Formulation

The uplink-downlink transmission-scheduling policy is defined as the sequence $\{a_1, a_2, \cdots, a_k, \cdots\}$, where a_k is the transmission-scheduling action in time slot k. In the following, we optimize the transmission-scheduling policy for both the one-step and multi-step controllable plants such that the average cost of the plant in (3.2.2) is

minimized[3], i.e.,

$$\min_{a_1,a_2,\cdots,a_k,\cdots} \quad J = \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathrm{Tr}\left(\mathbf{Q}\mathbf{P}_k\right). \tag{3.3.23}$$

3.4 One-Step Controllable Case: Optimal Transmission-Scheduling Policy

We first investigate the optimal transmission scheduling policy for the one-step controllable case, as it will also shed some light onto the optimal policy design of general multi-step controllable cases. Note that in this section and the following, we focus on the static channel scenario, and the design method of the optimal scheduling policies can be extended to the fading channel scenario, which will be discussed in Section 3.6.

3.4.1 MDP Formulation

From (3.3.11), (3.3.3) and (3.3.12), the next state cost $\mathbf{P}_{k+1}$, and the next states τ_{k+1} and ϕ_{k+1} only depend on the current transmission-scheduling action a_k and the current states τ_k and ϕ_k. Therefore, we can reformulate the problem (3.3.23) into a Markov Decision Process (MDP) as follows.

1) The state space is defined as $\mathbb{S} \triangleq \{(\tau,\phi) : \phi \geq \tau, \phi \neq \tau+1, \tau \in \mathbb{N}, \phi \in \{2,3,\cdots\}\}$ as illustrated in Fig. 3.3. Note that the states with $\phi = \tau+1$ are transient states (which can be verified using (3.3.3) and (3.3.12)) and are not included in $\mathbb{S}$, since we only focus on the long-term performance of the system. The state of the

[3]In this work, we only focus on the design of the scheduling policy $\{a_k\}$, when the controller gain $\mathbf{K}$ and the length of predictive control v are given and fixed. In our future work, the controller gain, the length of predictive control and the scheduling sequence will be jointly optimized.

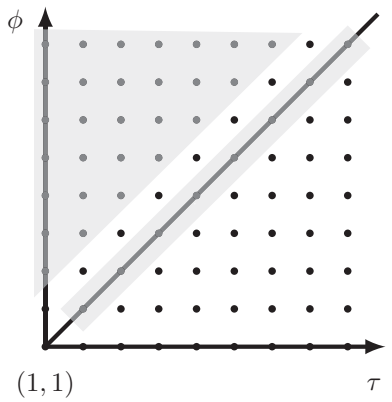


Figure 3.3: The state space $\mathbb{S}$ (shaded dots) of the MDP.

MDP at time k is $\mathbf{s}_k \triangleq (\tau_k, \phi_k) \in \mathbb{S}$.

2) The action space of the MDP is defined as $\mathbb{A} \triangleq \{1, 2\}$. The action at time k, $a_k \triangleq \pi(\mathbf{s}_k) \in \mathbb{A}$, indicates the sensor's transmission ($a_k = 1$) or the controller's transmission ($a_k = 2$) in time slot k.

3) The state-transition probability $P(\mathbf{s}'|\mathbf{s}, a)$ is the probability that the state $\mathbf{s}$ at time $(k-1)$ transits to $\mathbf{s}'$ at time k with action a at time $(k-1)$. We drop the time index k here since the transition is time-homogeneous. Let $\mathbf{s} = (\tau, \phi)$ and $\mathbf{s}' = (\tau', \phi')$ denote the current and next state, respectively. From (3.3.3) and (3.3.12), the state-transition probability can be obtained as

$$P(\mathbf{s}'|\mathbf{s}, a) = \begin{cases} p_s, & \text{if } a = 1, \mathbf{s}' = (\tau + 1, \phi + 1) \\ 1 - p_s, & \text{if } a = 1, \mathbf{s}' = (1, \phi + 1) \\ p_c, & \text{if } a = 2, \mathbf{s}' = (\tau + 1, \phi + 1) \\ 1 - p_c, & \text{if } a = 2, \mathbf{s}' = (\tau + 1, \tau + 1) \\ 0, & \text{otherwise.} \end{cases} \tag{3.4.1}$$

4) The one-stage cost of the MDP, i.e., the one-step quadratic-cost of the plant in (3.2.2), is a function of the current state ϕ as

$$c(\mathbf{s}) = c(\phi) \triangleq \mathrm{Tr}\left(\mathbf{QP}\right) = \mathrm{Tr}\left(\mathbf{QF}(\phi)\right), \tag{3.4.2}$$

which is independent of the state τ and the action a. The function $c(\cdot)$ has the following property:

Lemma 3.1. The one-stage cost function $c(\phi)$ in (3.4.2) is a strictly monotonically increasing function of ϕ, where $\phi \in \{2, 3, \cdots\}$.

Proof. Since $\mathbf{R}$ is a positive definite matrix, $\mathbf{MRM}^\top$ is positive definite for any n-by-n non-zero matrix $\mathbf{M}$. Also, we have $\mathbf{A}^i \neq \mathbf{0}, \forall i \in \mathbb{N}$, as it is assumed that $\rho(\mathbf{A}) > 1$ in Section 3.2.1. Due to the fact that the product of positive-definite matrices has positive trace and $\mathbf{Q}$ is positive definite, $\mathrm{Tr}\left(\mathbf{QA}^i\mathbf{R}(\mathbf{A}^i)^\top\right)$ is positive, $\forall i \in \mathbb{N}$. From the definition of $\mathbf{F}(\cdot)$ in (3.3.5), we have

$$\begin{aligned} c(\phi+z) - c(\phi) &= \mathrm{Tr}(\mathbf{QF}(\phi+z)) - \mathrm{Tr}(\mathbf{QF}(\phi)) \\ &= \sum_{i=\phi+1}^{\phi+z} \mathrm{Tr}\left(\mathbf{QA}^i\mathbf{R}(\mathbf{A}^i)^\top\right) > 0, \forall z \in \mathbb{N}. \end{aligned} \tag{3.4.3}$$

This completes the proof. □

Therefore, the problem (3.3.23) is equivalent to finding the optimal policy $\pi(\mathbf{s}s), \forall \mathbf{s} \in \mathbb{S}$ by solving the classical *average cost minimization problem* of the MDP [76]. If a stationary and deterministic optimal policy of the MDP exists, we can effectively find the optimal policy by using standard methods such as the relative value iteration algorithm see e.g., [76, Chapter 8].

3.4.2 Existence of the Optimal Scheduling Policy

If the uplink and downlink channels have high packet-error probabilities, the average cost in (3.3.23) may never be bounded no matter what policy we choose. Therefore, *we need to study the condition in terms of the transmission reliability of the*

uplink and downlink channels under which the dynamic plant can be stabilized, i.e., the average cost can be bounded. We derive the following result.

Theorem 3.1. In the static channel scenario, there exists a stationary and deterministic optimal transmission-scheduling policy that can stabilize the one-step controllable plant (3.2.1) iff

$$\max\{p_s, p_c\} < \frac{1}{\rho^2(\mathbf{A})}, \tag{3.4.4}$$

where we recall that $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$.

Proof. The necessity of the condition can be easily proved as (3.4.4) is the necessary and sufficient condition that an ideal FD controller with the uplink-downlink packet-error probabilities $\{p_s, p_c\}$ can stabilize the remote plant [12]. Intuitively, if (3.4.4) does not hold, an FD controller cannot stabilize the plant and thus an HD controller cannot either, no matter what transmission-scheduling policy it applies.

The sufficiency part of the proof is conducted by proving the existence of a stationary and deterministic policy π' that can stabilize the plant if (3.4.4) is satisfied, where

$$\pi'(\mathbf{s}) = \pi'(\tau, \phi) = \begin{cases} 1, \ \tau = \phi, (\tau, \phi) \in \mathbb{S} \\ 2, \ \text{otherwise.} \end{cases} \tag{3.4.5}$$

The details of the proof are given in Appendix B. □

Remark 3.3. *Theorem 3.1 states that the optimal policy, which stabilizes the plant, exists if both the channel conditions of the uplink and downlink channels are good (i.e., small p_s and p_c) and the dynamic process does not change rapidly (i.e., a small $\rho^2(\mathbf{A})$). Also, it is interesting to see that the HD controller has exactly the same condition as the FD controller [12] to stabilize the plant. However, since the HD operation*

naturally introduces longer delays in both transmissions of the sensor measurement and the control command than the FD operation, the bounded average cost of the HD controller should be higher than the FD one, which will be illustrated in Section 3.7.

Assuming that the condition (3.4.4) is satisfied, we have the following property of the optimal policy.

Proposition 3.2. The stationary and deterministic optimal policy of the problem (3.3.23), $\pi^*(\tau, \phi)$, is a switching-type policy in terms of τ and ϕ, i.e., (i) if $\pi^*(\tau, \phi) = 1$, then $\pi^*(\tau+z, \phi) = 1$, $\forall z \in \mathbb{N}$ and $(\tau+z, \phi) \in \mathbb{S}$; (ii) if $\pi^*(\tau, \phi) = 2$, then $\pi^*(\tau, \phi+z) = 2$, $\forall z \in \mathbb{N}$ and $(\tau, \phi + z) \in \mathbb{S}$.

Proof. The proof follows the same procedure as Theorem 2.2 in chapter 2 and is omitted here. □

Therefore, for the optimal policy, the state space is divided into two parts by a curve, and the scheduling actions of the states in each part are the same, which will be illustrated in Section 3.7. Such a switching structure helps save storage space for on-line transmission scheduling, as the controller only needs to store the states of the switching boundary instead of the entire state space.

3.4.3 Suboptimal Policy

In practice, to solve the MDP problem in Section 3.4.1 with an infinite number of states, one needs to approximate it by a truncated MDP problem with finite states for offline numerical evaluation. The computing complexity of the problem is $\mathcal{O}(AB^2C)$ [79], where A and B are the cardinalities of the action space and the state

space, respectively, and C is the number of convergence steps for solving the problem. To reduce the computation complexity, we propose a myopic policy $\psi(\mathbf{s}), \forall \mathbf{s} \in \mathbb{S}$, which simply makes the online decision to optimize the expected next stage cost.

From (3.4.1) and (3.4.2), the expected next stage cost $\mathbb{E}[c(\phi')|\mathbf{s}, a = \psi(\mathbf{s})]$, where $\mathbf{s} = (\tau, \phi)$, is derived as

$$\begin{aligned} &\mathbb{E}[c(\phi')|\mathbf{s}, \psi(\mathbf{s}) = 1] = c(\phi + 1), \\ &\mathbb{E}[c(\phi')|\mathbf{s}, \psi(\mathbf{s}) = 2] = p_c c(\phi + 1) + (1 - p_c)c(\tau + 1). \end{aligned} \tag{3.4.6}$$

1) For states $\{\mathbf{s}|(\tau, \phi) \in \mathbb{S}, \phi > \tau\}$, from (3.4.6), the action $\psi(\mathbf{s}) = 2$ *results in a smaller next stage cost than* $\psi(\mathbf{s}) = 1$.

2) For states $\{\mathbf{s}|(\tau, \phi) \in \mathbb{S}, \phi = \tau\}$, from (3.4.6), since the two actions lead to the same next stage cost, i.e.,

$$\mathbb{E}[c(\phi')|\mathbf{s}, \psi(\mathbf{s}) = 1] = \mathbb{E}[c(\phi')|\mathbf{s}, \psi(\mathbf{s}) = 2] = c(\phi + 1), \tag{3.4.7}$$

we need to compare the second stage cost led by the actions. If $\psi(\mathbf{s}) = 1$, $\mathbf{s}' \in \{(1, \phi + 1), (\phi + 1, \phi + 1)\}$. If $\mathbf{s}' = (1, \phi + 1)$, since $\phi + 1 > 1$, the next stage myopic action is $\psi(1, \phi + 1) = 2$ as discussed earlier, and the second stage state $\mathbf{s}'' \in \{(2, \phi + 2), (2, 2)\}$. If $\mathbf{s}' = (\phi + 1, \phi + 1)$, from (3.4.7), the expected second stage cost is $c(\phi + 2)$ for both $\psi(\mathbf{s}') = 1$ and 2. Based on this analysis and (3.4.1), we have the expected second stage cost with $\phi(\mathbf{s}) = 1$ as

$$\begin{aligned} \mathbb{E}[c(\phi'')|\mathbf{s}, \psi(\mathbf{s}) = 1] = (1 - p_s) \left(p_c c(\phi + 2) + (1 - p_c)c(2)\right) \\ + p_s c(\phi + 2). \end{aligned} \tag{3.4.8}$$

Similarly, we can obtain the expected second stage cost with $\phi(\mathbf{s}) = 2$ as

$$\mathbb{E}[c(\phi'')|\mathbf{s}, \psi(\mathbf{s}) = 2] = c(\phi + 2). \tag{3.4.9}$$

Since $p_c, p_s < 1$ and $c(2) < c(\phi + 2)$ from Lemma 3.1, $\psi(\mathbf{s}) = 1$ *results in a smaller cost than* $\psi(\mathbf{s}) = 2$. From the above analysis, the myopic policy $\psi(\mathbf{s})$ is equal to $\pi'(\mathbf{s})$

in (3.4.5), $\forall \mathbf{s} \in \mathbb{S}$.

Proposition 3.3. The myopic policy of problem (3.3.23) is π' in (3.4.5).

Remark 3.4. *From the myopic policy* (3.4.5) *and the state-updating rules* (3.3.3) *and* (3.3.12)*, we see that the policy π' is actually a persistent scheduling policy, which consecutively schedules the uplink transmission until one is successful, and then consecutively schedules the downlink transmission until one is successful, and so on.*

From the property of the persistent scheduling policy, we can easily obtain the result below.

Corollary 3.1. *For the persistent uplink-downlink scheduling policy π' in Proposition 3.3, the chances of scheduling the sensor and the controller transmissions are $\frac{1-p_c}{(1-p_c)+(1-p_s)}$ and $\frac{1-p_s}{(1-p_c)+(1-p_s)}$, respectively.*

3.4.4 Naive Policy: A Benchmark

We consider a naive uplink-downlink scheduling policy of the HD controller as a benchmark of the proposed optimal scheduling policy. The naive policy simply schedules the sensor and the controller transmissions alternatively, i.e., $\{\cdots$, sensing, control, sensing, control, $\cdots\}$, without taking into account the state-estimation quality of the controller nor the state-quality of the plant. Such a naive policy is also noted as the *round-robin scheduling policy*.

Theorem 3.2. In the static channel scenario, the alternative scheduling policy can stabilize the one-step controllable plant (3.2.1) iff

$$\max\{p_s, p_c\} < \frac{1}{\left(\rho^2(\mathbf{A})\right)^2}. \tag{3.4.10}$$

Proof. See Appendix C. □

Remark 3.5. *Compared with Theorem 3.1, to stabilize the same plant the naive policy may require smaller packet-error probabilities of the uplink and downlink channels than the proposed optimal scheduling policy. This also implies that the optimal policy can result in a notably smaller average cost of the plant than the naive policy, which will be illustrated in Section 3.7.*

3.5 v-Step Controllable Case: Optimal Transmission-Scheduling Policy

3.5.1 MDP Formulation

Based on Proposition 3.1, the average cost minimization problem (3.3.23) can be formulated as an MDP similar to the one-step controllable case in Section 3.4 as:

1) The state space is defined as $\mathbb{S} \triangleq \{(\tau_k^0, \tau_k^1, \cdots, \tau_k^{v-1}, \phi_k^0, \phi_k^1, \cdots, \phi_k^{v-1}) : \phi_k^i \geq \tau_k^i, \phi_k^i \neq \tau_k^i + 1, \tau_k^i \in \mathbb{N}, \phi_k^i \in \{2, 3, \cdots\}, \forall i = 0, \cdots, v-1\}$.

2) The action space of the MDP is exactly the same as that of the one-step controllable plant in Section 3.4.1.

3) Let $P(\mathbf{s}'|\mathbf{s}, a)$ denote the state-transition probability, where $\mathbf{s} = (\tau^0, \cdots, \tau^{v-1}, \phi^0, \cdots, \phi^{v-1})$

and $\mathbf{s}' = ((\tau^0)', \cdots, (\tau^{v-1})', (\phi^0)', \cdots, (\phi^{v-1})')$ are the current and next state, respectively, after dropping the time indexes. From (3.3.18) and (3.3.20), the state-transition probability is obtained as

$$P(\mathbf{s}'|\mathbf{s}, a) = \begin{cases} p_s, \text{if } a=1, \mathbf{s}'=(\tau^0+1, \tau^1, \cdots, \tau^{v-1}, \phi^0+1, \phi^1, \cdots, \phi^{v-1}) \\ 1-p_s, \text{if } a=1, \mathbf{s}'=(1, \tau^1, \cdots, \tau^{v-1}, \phi^0+1, \phi^1, \cdots, \phi^{v-1}) \\ p_c, \text{if } a=2, \mathbf{s}'=(\tau^0+1, \tau^1, \cdots, \tau^{v-1}, \phi^0+1, \phi^1, \cdots, \phi^{v-1}) \\ 1-p_c, \text{if } a=2, \mathbf{s}'=(\tau^0+1, \tau^0, \cdots, \tau^{v-2}, \tau^0+1, \phi^0, \cdots, \phi^{v-2}) \\ 0, \text{otherwise.} \end{cases} \tag{3.5.1}$$

4) The one-stage cost of the MDP is a function of the current state $\mathbf{s}$, and is obtained from (3.2.2) and Proposition 3.1 as

$$\begin{aligned} c(\mathbf{s}) &= c(\tau^1, \cdots, \tau^{v-1}, \phi^0, \cdots, \phi^{v-1}) \\ &= \mathrm{Tr}\left(\mathbf{Q}\left[\mathbf{F}(\phi^0) + \sum_{i=0}^{v-2} \mathbf{G}\left(\sum_{j=0}^{i} \phi^j - \sum_{j=0}^{i} \tau^{j+1}, \right.\right.\right. \\ &\qquad \left.\left.\left. \mathbb{1}(\phi^{i+1} > \tau^{i+1})\left(\mathbf{F}(\phi^{i+1}) - \mathbf{F}(\tau^{i+1})\right)\right)\right]\right). \end{aligned} \tag{3.5.2}$$

Remark 3.6. *Different from the one-step controllable case, where the one-stage cost function is a monotonically increasing function of the state parameter ϕ, the cost function in* (3.5.2) *is more complex and does not have this property. Thus, the switching structure of the optimal policy does not hold in general for the v-step controllable case.*

3.5.2 Existence of the Optimal Scheduling Policy

Theorem 3.3. In the static channel scenario, there exists a stationary and deterministic optimal transmission-scheduling policy that can stabilize the v-step controllable plant (3.2.1) using the predictive control method (3.2.7), (3.2.9) and (3.2.10) iff (3.4.4) holds.

Proof. See Appendix D. □

Remark 3.7. *The stability condition of a v-step controllable plant is exactly the same as that of the one-step controllable plant in Theorem 3.1. Thus, whether a plant can be stabilized by an HD controller simply depends on the spectral radius of the plant parameter* $\mathbf{A}$ *and the uplink and downlink transmission reliabilities.*

Remark 3.8. *Although the stability conditions of a one-step and a v-step plants are the same, to find the optimal uplink-downlink scheduling policy, the state space and the computation complexity of the MDP problem increase linearly and exponentially with v [79], respectively. However, in the following section, we will show that the persistent scheduling policy in Proposition 3.3 (which can be treated as a policy that makes decision simply relying on two state parameters, i.e., ϕ^0 and τ^0, instead of the entire $2v$ state parameters) can provide a remarkable performance close to the optimal one.*

3.6 Extension to Fading Channels

In this section, we investigate the optimal transmission-scheduling policy for the general v-step controllable case in the fading channel scenario, where $v \geq 1$.

3.6.1 MDP Formulation

Compared with the static channel scenario, the transmission scheduling of the WNCS in the fading channel scenario should take into account the channel states of both the uplink and downlink channels, and hence expand the dimension of the state space. Also, the state-transition probabilities of the MDP problem should also rely on

the transition probabilities of channel states. Therefore, the detailed MDP problem for solving the average cost minimization problem (3.3.23) can be formulated as:

1) The state space is defined as $\mathbb{S} \triangleq \{(\tau_k^0, \tau_k^1, \cdots, \tau_k^{v-1}, \phi_k^0, \phi_k^1, \cdots, \phi_k^{v-1}, h_{s,k}, h_{c,k}) : \phi_k^i \geq \tau_k^i, \phi_k^i \neq \tau_k^i+1, \tau_k^i \in \mathbb{N}, \phi_k^i \in \{2, 3, \cdots\}, h_{s,k} \in \{1, \cdots, B_s\}, h_{c,k} \in \{1, \cdots, B_c\}, \forall i = 0, \cdots, v-1\}$, where $h_{s,k}$ and $h_{c,k}$ are channel-state indexes of the uplink and downlink channels at time k, respectively.

2) The action space is the same as that of the static channel scenario in Section 3.5.1.

3) As the state transition is time-homogeneous, we drop the time index k here. Let $h \triangleq (h_s, h_c)$ and $h' \triangleq (h'_s, h'_c)$ denote the current and the next uplink-downlink channel states, respectively. As the uplink and downlink channel are action-invariant and independent of each other, the overall channel state transition probability can be directly obtained from (3.2.4) and (3.2.5) as

$$P(h'|h) = d^s_{h_s,h'_s} d^c_{h_c,h'_c}. \tag{3.6.1}$$

Let $\mathbf{s} \triangleq (\tau^0, \cdots, \tau^{v-1}, \phi^0, \cdots, \phi^{v-1}, h)$ and $\mathbf{s}' \triangleq ((\tau^0)', \cdots, (\tau^{v-1})', (\phi^0)', \cdots, (\phi^{v-1})', h')$

denote the current and the next states of the WNCS, respectively. The state-transition probability $P(s'|s,a)$ can be obtained as

$$P(\mathbf{s}'|\mathbf{s},a) = \begin{cases} P(h'|h)\omega_{h_s}, \text{if } a=1 \text{ and} \\ \qquad \mathbf{s}' = (\tau^0+1,\cdots,\tau^{v-1},\phi^0+1,\cdots,\phi^{v-1},h'), \\ P(h'|h)(1-\omega_{h_s}), \text{if } a=1 \text{ and} \\ \qquad \mathbf{s}' = (1,\cdots,\tau^{v-1},\phi^0+1,\cdots,\phi^{v-1},h'), \\ P(h'|h)\xi_{h_c}, \text{if } a=2 \text{ and} \\ \qquad \mathbf{s}' = (\tau^0+1,\cdots,\tau^{v-1},\phi^0+1,\cdots,\phi^{v-1},h'), \\ P(h'|h)(1-\xi_{h_c}), \text{if } a=2 \text{ and} \\ \qquad \mathbf{s}' = (\tau^0+1,\cdots,\tau^{v-2},\tau^0+1,\cdots,\phi^{v-2},h'), \\ 0, \text{ otherwise.} \end{cases} \tag{3.6.2}$$

4) The one-stage cost of the MDP is the same as (3.5.2).

Such an MDP problem with $(2v+2)$ state dimensions and a small action space can be solved by standard MDP algorithms similar to that of the static channel scenario discussed earlier.

3.6.2 Existence of the Optimal Scheduling Policy

In the fading channel scenario, since each state of the Markov chain induced by a scheduling policy has $(2v+2)$ dimensions, it is difficult to analyze the average cost of the Markov chain and determine whether it is bounded or not. Therefore, it is hard to give a necessary and sufficient condition in terms of the properties of

the Markov channels and the plant, under which the MDP problem has a scheduling-policy solution leading to a bounded minimum average cost. However, inspired by the result of the static channel scenario in Section 3.5.2, we can directly give a necessary condition and a sufficient one by considering the best and the worst Markov channel conditions of the uplink and downlink channels as below.

Theorem 3.4. In the fading channel scenario, a necessary condition and a sufficient condition of existing a stationary and deterministic optimal transmission-scheduling policy that can stabilize the general v-step controllable plant (3.2.1) using the predictive control method (3.2.7), (3.2.9) and (3.2.10) are given by

$$\max\left\{\underline{p_s}, \underline{p_c}\right\} < \frac{1}{\rho^2(\mathbf{A})}, \tag{3.6.3}$$

and

$$\max\left\{\overline{p_s}, \overline{p_c}\right\} < \frac{1}{\rho^2(\mathbf{A})}, \tag{3.6.4}$$

respectively, where $\underline{p_s} \triangleq \min\{\omega_1, \cdots, \omega_{B_s}\}$, $\overline{p_s} \triangleq \max\{\omega_1, \cdots, \omega_{B_s}\}$, $\underline{p_c} \triangleq \min\{\xi_1, \cdots, \xi_{B_c}\}$, $\overline{p_c} \triangleq \max\{\xi_1, \cdots, \xi_{B_c}\}$.

In general, Theorem 3.4 says that the plant can be stabilized by a transmission scheduling policy as long as the worst achievable channel conditions of the uplink and downlink Markov channels are good enough, and it cannot be stabilized by any scheduling policy if the best achievable channel conditions of the uplink and downlink Markov channels are poor.

In the following section, we will numerically evaluate the performance of the plant using the optimal transmission scheduling policy, where the sufficient condition of the existence of an optimal policy (3.6.4) is satisfied.

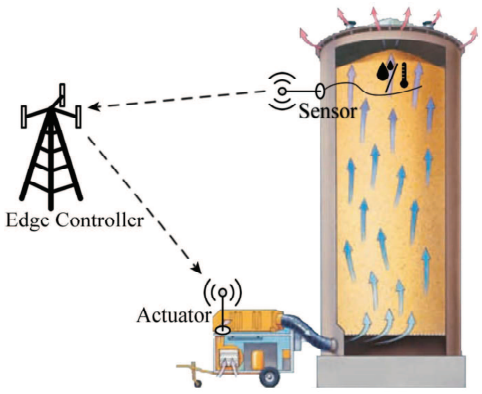


Figure 3.4: Temperature and humidity control in grain conservation.

3.7 Numerical Results

The uplink-downlink scheduling policies that we developed can be applied to a large range of real IIoT applications, including temperature control of hot rolling process in the iron and steel industry, flight path control of delivery drones, voltage control in smart grids, and lighting control in smart homes/buildings. Specifically, in this section, we apply the uplink-downlink scheduling policies to a real application of smart farms as illustrated in Fig. 3.4. The system contains a grain container, a sensor measuring the temperature (°C) and the humidity (%) of the grain pile, an actuator which has a high pressure fan and/or an air cooler, and an edge controller, which receives the sensor's measurements, and then computes and sends the command to the actuator. Given the present values of temperature and humidity, the state vector $\mathbf{x}_k$ in (3.2.1) contains two parameters, i.e., the current temperature and humidity offsets. Note that since grain absorbs water from the air and generates heat naturally, the temperature and the humidity levels of the grain pile will automatically increase without proper control, leading to severe insect and mold development [101]. In general, by using a high pressure fan for ventilation, both the temperature and the

humidity can be controlled in a proper range. If an air cooler is also available, the temperature can be controlled to the preset value faster. Thus, if the actuator has a high pressure fan only, given a preset fan speed, its control input $\mathbf{u}_k$ in (3.2.1) has only one parameter, which is the relative fan speed (measured by the flow volume [m^3/h]). If the actuator has both high pressure fan and air cooler, the control input $\mathbf{u}_k$ has two parameters including the fan speed and the cooler temperature (°C). The former and the latter cases will be studied in Section 3.7.1, and Sections 3.7.2 and 3.7.3, respectively.

The discrete time step T_0 in this example is set to be one second [50]. Unless otherwise stated, we assume the system parameters as $\mathbf{A} = \begin{bmatrix} 1.1 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$, $\mathbf{Q} = \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and thus $\rho^2(\mathbf{A}) = 1.44$. Since the controller, the sensor and the actuator have very low mobility in this example, we focus on the static channel scenario and set the packet error probabilities of the uplink and downlink channels as $p_s = 0.1$ and $p_c = 0.1$, respectively, and also study the fading channel scenario in Section 3.7.1.

In the following, we present the numerical results of the optimal policies and the optimal average costs of the plant in Sections 3.4 and 3.5 for one-step and v-step controllable cases, respectively. Also, we numerically compare the performance of the optimal scheduling policy with the persistent scheduling policy in Section 3.4.3, the benchmark (naive) policy in Section 3.4.4, and also the ideal FD policy in [12], i.e., where the controller works in the FD mode and have the same packet-error probabilities of the uplink-downlink channels as in the HD mode.

Note that to calculate the optimal policies in Sections 3.4.1, 3.5.1 and 3.6.1 by solving the MDP problems, the infinite state space $\mathbb{S}$ is first truncated by limiting

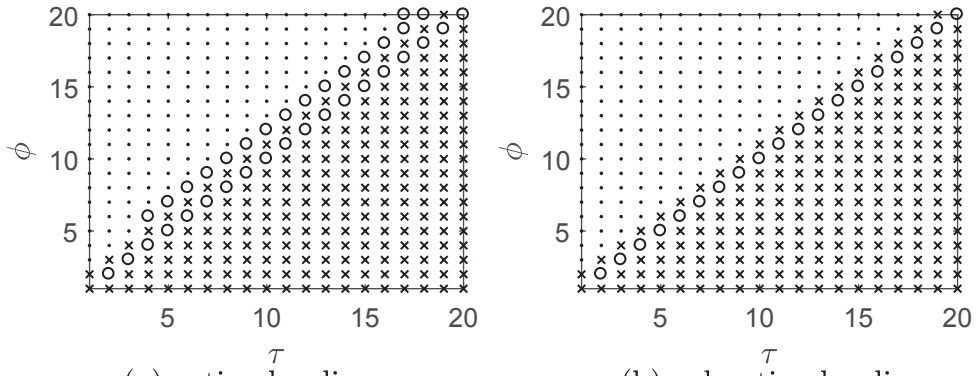
the range of the state parameters as $1 \leq \tau^i, \phi^i \leq 20, \forall i = 0, \cdots, v-1$, to enable the evaluation. For example, if we consider a two-step controllable case, i.e., $v = 2$, there will be $20^{2\times v} = 160,000$ states in the static channel scenario, and there will be many more in the fading channel scenario. To solve finite-state MDP problems, in general, there are two classical methods: policy iteration and value iteration. The policy iteration method converges faster in solving small-scale MDP problems, but is more computationally burdensome than the value iteration method when the state space is large [76]. Since our problems have large state spaces, we adopt the classical value iteration method for solving MDP problems by using a well recognized Matlab MDP toolbox [78].

3.7.1 One-Step Controllable Case

In this case, we assume that $\mathbf{B} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and $\mathbf{K} = \mathbf{A}$ satisfying $\mathbf{A} + \mathbf{BK} = \mathbf{0}$.

Optimal and suboptimal policies. Fig. 3.5 shows the optimal policy and the persistent (suboptimal) policy in Proposition 3.3 within the truncated state space. We see that although the optimal policy has more states choosing to schedule the sensor's transmission than the persistent policy, these two policies look similar. Also, we see that the optimal policy is a switching-type one, in line with Proposition 3.2.

Performance comparison. We further evaluate the performances of the optimal scheduling policy, the persistent policy, the naive policy and the FD policy in terms of the K-step average cost of the plant using $\frac{1}{K}\sum_{k=0}^{K-1} \mathbf{x}_k^\top \mathbf{Q}\mathbf{x}_k$. We run 10^4-step simulations with the initial value of the plant-state vector $\mathbf{x}_0 = [1, -1]^\top$. The initial state for the optimal and persistent policies is $(\tau_0, \phi_0) = (2, 2)$. The initial scheduling of the naive policy is the sensor's transmission.



(a) optimal policy (b) suboptimal policy

Figure 3.5: The uplink-downlink scheduling policies, where where 'o' and '.' denote $a = 1$ and $a = 2$, respectively, and 'x' denotes a state that does not belong to $\mathbb{S}$.

Fig. 3.6 shows the average cost versus the simulation time, using different policies. We see that the average costs induced by different policies converge to the steady state values when $K > 3000$. Given the baseline of the FD (non-scheduling) policy, the optimal scheduling policy gives a significant 60% average cost reduction of the naive policy. Also, we see that the persistent policy provides a performance close to the optimal one. We note that there is a noticeable performance gap between the optimal scheduling policy of the HD controller and the FD policy of the FD controller, since the HD operation introduces extra delays in uplink-downlink transmissions and deteriorates the performance of the control system.

Performance versus transmission reliabilities. In Fig. 3.7, we show a contour plot of the average cost of the plant with different uplink-downlink packet-error probabilities (p_s, p_c) within the rectangular region that can stabilize the plant, i.e., $p_s, p_c < 1/\rho^2(\mathbf{A}) = 0.7$ based on Theorem 3.1. The average cost is calculated by running a 10^6-step simulation and then taking the average, and it does not have a steady-state value if (p_s, p_c) lies outside the rectangular region. We see that the average cost increases quickly when p_s or p_c approaches the boundary $1/\rho^2(\mathbf{A})$. Also, it is interesting to see that in order to guarantee a certain average cost, e.g., $J = 8$, the

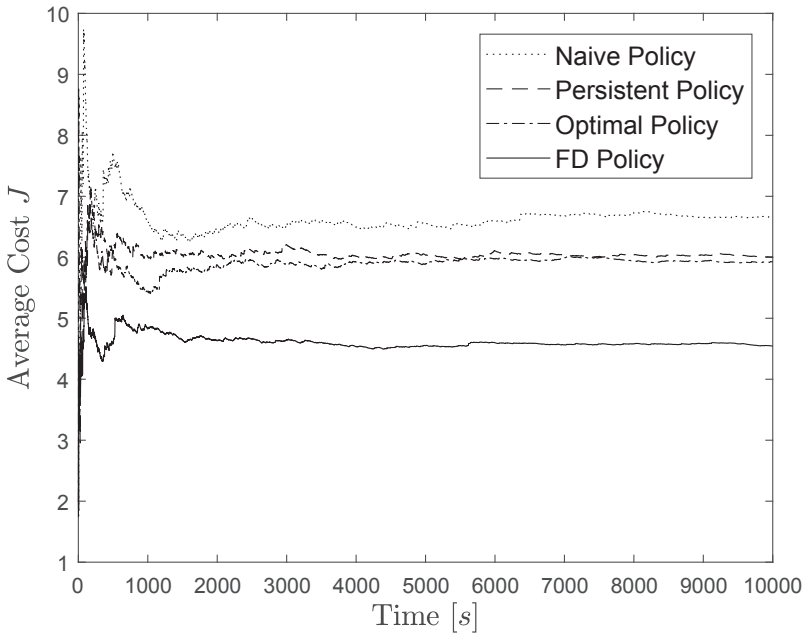


Figure 3.6: One-step controllable case: average cost versus time.

required p_s is less than p_c in general, which implies that the transmission reliability of the sensor-controller channel is more important than that of the controller-actuator channel.

Fading channel scenario. Assume that both the uplink and downlink channels have two Markov channel states with the packet error probabilities 0.1 (i.e., the good channel state) and 0.4 (i.e., the bad channel state), respectively, i.e., $\omega_1 = \xi_1 = 0.1$ and $\omega_2 = \xi_2 = 0.4$. Figs. 3.8 and 3.9 show the average cost versus the simulation time with different channel state transition probabilities. In Fig. 3.8, we set the matrices of the channel state transition probabilities of the uplink and downlink channels as $\mathbf{D}_s = \mathbf{D}_c = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$. Taking the uplink channel as an example, the transition probabilities from the bad and the good channel state are the same, and thus the Markov channel does not have any memory [100]. Since the uplink and downlink channels have the same Markovian property, both the uplink and downlink channels

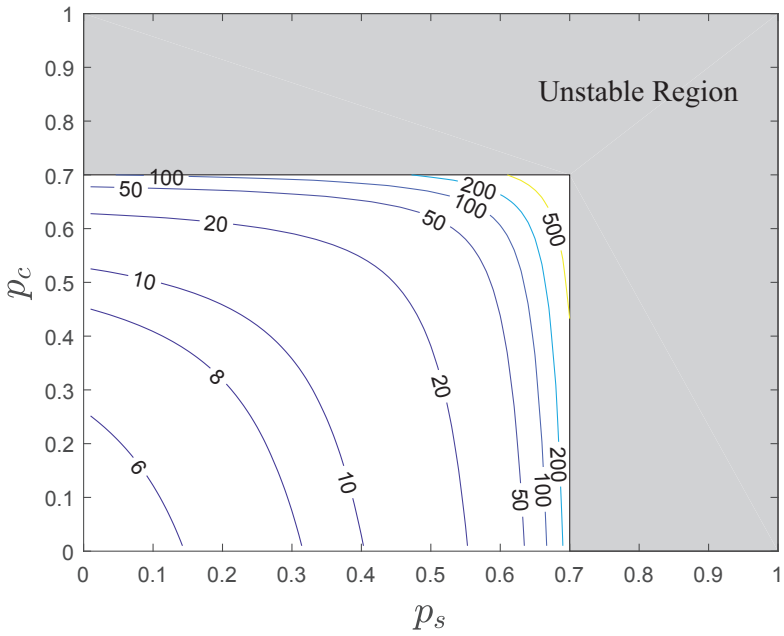


Figure 3.7: One-step controllable case: average cost versus packet-error probabilities, i.e., p_s and p_c.

are memoryless. In Fig. 3.9, we set $\mathbf{D}_s = \mathbf{D}_c = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$, where the probability of remaining in any given state is higher than jumping to the other state. In this case, both the uplink and downlink channels have persistent memories.

In Figs. 3.8 and 3.9, we see that the persistent policy always provides a low average cost, which is close to that of the optimal policy and is much smaller than that of the naive policy. It is interesting to see that the average cost achieved by the optimal policy under the memoryless Markov channels in Fig. 3.8 is smaller than that of the Markov channels with memories in Fig. 3.9. This is because in the Markov channel with memory, if the current channel state is bad, it is more likely to have a bad channel state again in the next time slot, which can lead to consecutive packet losses and deteriorate the control performance of the WNCS.

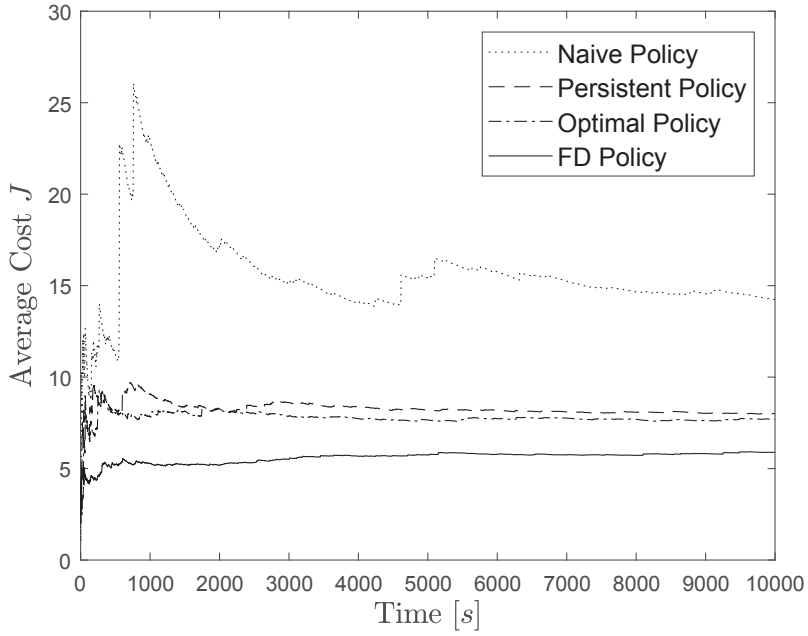


Figure 3.8: Markov channel scenario without memory: average cost versus time.

3.7.2 Two-Step Controllable Case

In this case, we assume that $\mathbf{B} = -[1,1]^{\top}$, and $\mathbf{K} = [2.9, -1]$ satisfying $(\mathbf{A}+\mathbf{BK})^2 = \mathbf{0}$. For fair comparison, all the policies considered in this subsection adopt the same predictive control method in (3.2.7), (3.2.9) and (3.2.10) with $v = 2$.

In Fig. 3.10, we plot the average cost function versus the packet-error probability of the downlink channel with different uplink-downlink transmission-scheduling policies, where the uplink packet-error probability $p_s = 0.1$. We see that the persistent policy can still provide a good performance close to the optimal policy. Given the FD policy as a benchmark, it is clear that the optimal scheduling policy provides at least a 66% reduction of the average cost than the naive policy when $p_c \geq 0.1$.

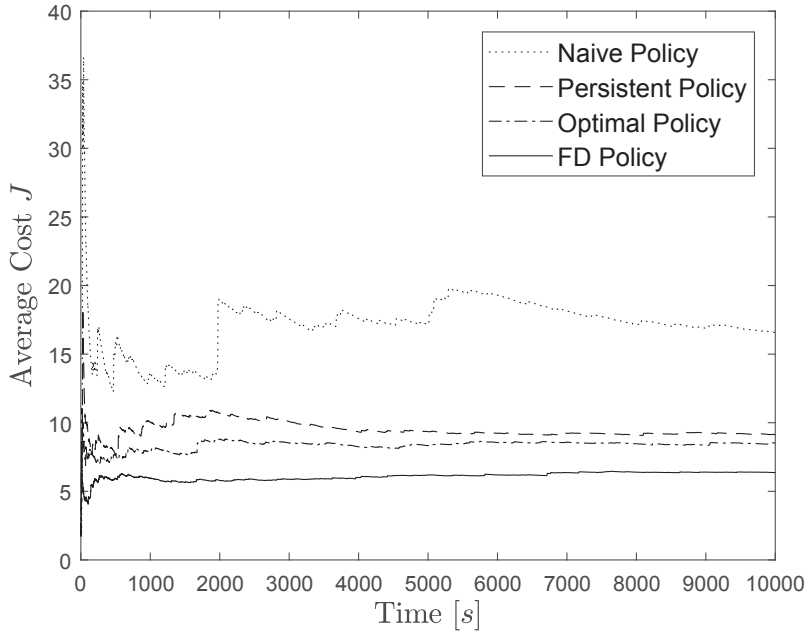


Figure 3.9: Markov channel scenario with memory: average cost versus time.

3.7.3 Non-Finite-Step-Controllable Case

We now look at the non-finite-step-controllable case as discussed in Remark 3.2, where $\mathbf{B} = -[1,1]^\top$, and $\mathbf{K} = [0.7, 0.4]$. It can be verified that $\rho(\mathbf{A}+\mathbf{BK}) = 0.72 < 1$ and $(\mathbf{A}+\mathbf{BK})^v \neq \mathbf{0}$ for a practical range of v, e.g. $v < 10$. We consider two predictive control protocols in (3.2.7) with $v = 2$ and $v = 3$, respectively, i.e., the controller sends two or three commands to the actuator each time. We have $(\mathbf{A}+\mathbf{BK})^1 = \begin{bmatrix} 0.4 & -0.2 \\ -0.5 & 0.4 \end{bmatrix}$, $(\mathbf{A}+\mathbf{BK})^2 = \begin{bmatrix} 0.26 & -0.16 \\ -0.4 & 0.26 \end{bmatrix}$ and $(\mathbf{A}+\mathbf{BK})^3 = \begin{bmatrix} 0.18 & -0.12 \\ -0.29 & 0.18 \end{bmatrix}$. It is clear that $(\mathbf{A}+\mathbf{BK})^v$ approaches $\mathbf{0}$ as v increases. By letting $(\mathbf{A}+\mathbf{BK})^v = \mathbf{0}$ in the analysis of plant-state vector in (B.1.4), where $v = 2$ or 3, the plant-state covariance matrix $\mathbf{P}_k$ is approximated by a function of $2v-1$ state parameters as in Proposition 3.1. Based on this approximation, we can formulate and solve the MDP problem in Section 3.5.1, resulting an *approximated optimal scheduling policy*.

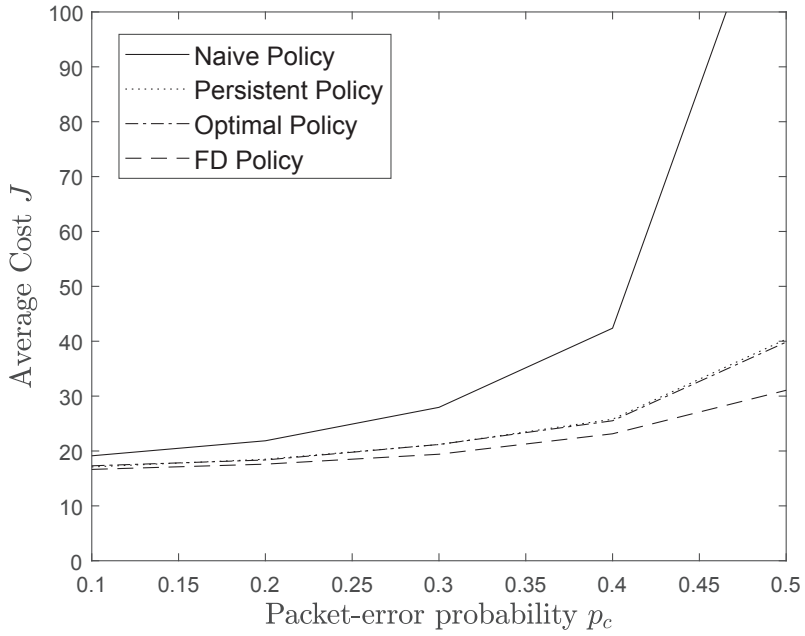


Figure 3.10: Two-step controllable case: average cost versus packet-error probability p_c.

In Fig. 3.11, we plot the average cost function versus the packet-error probability of the downlink channel with different downlink transmission-scheduling policies. We see that for both the cases $v = 2$ and 3, the performances of the approximated optimal and persistent uplink-downlink scheduling policies are quite close to the benchmark FD policy when $p_c < 0.2$, while the performance gap between the naive scheduling policy and the FD policy is large. This also implies that the approximated optimal policy is near optimal in this practical range of downlink transmission reliability, and the persistent scheduling policy is also effective, yet low-complexity in this case.

3.8 Conclusions

In this work, we have proposed an important uplink-downlink transmission scheduling problem of a WNCS with an HD controller for practical IIoT applications, which

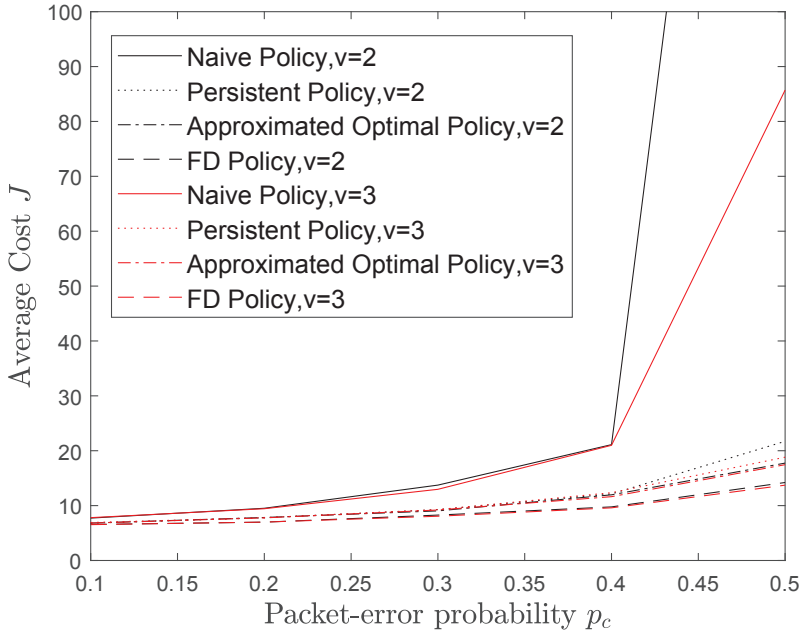


Figure 3.11: Non-finite-step-controllable case: average cost versus packet-error probability p_c.

has not been considered in the available literature. We have given a comprehensive analysis of the estimation-error covariance and the plant-state covariance of the HD-controller-based WNCS for both one-step and v-step controllable plants. Based on our analytical results, in both the static and fading channel scenarios, we have formulated the novel problem to optimize the transmission-scheduling policy depending on both the current estimation quality of the controller and the current cost function of the plant, so as to minimize the long-term average cost function. Moreover, for the static channel scenario, we have derived the necessary and sufficient condition of the existence of a stationary and deterministic optimal policy that results in a bounded average cost in terms of the transmission reliabilities of the uplink and downlink channels. For the fading channel scenario, we have derived a necessary condition and a sufficient condition in terms of the uplink and downlink channel qualities, under which the optimal transmission-scheduling policy exists. Our problem can be solved

effectively by the standard MDP algorithms if the optimal scheduling policy exists. Also, we have derived an easy-to-compute suboptimal policy, which provides a control performance close to the optimal policy and notably reduces the average cost of the plant compared to a naive alternative-scheduling policy.

Chapter 4

Wireless Feedback Control With Variable Packet Length for Industrial IoT

4.1 Introduction

The Industrial Internet of Things (IIoT) can be treated as an extension of consumer IoT in industrial applications. One of the most important applications for IIoT is industrial control [83], with user scenarios ranging from building and process automation to more mission-critical applications, such as factory automation and power system control [82]. Wireless networked control systems (WNCSs) are composed of spatially distributed controllers, sensors and actuators communicating through wireless channels, and physical processes to be controlled. Due to the enhanced flexibility and the reduced deployment and maintenance costs, WNCSs are becoming a fundamental infrastructure technology for mission-critical control applications [80]. In [12], the optimal control policy and the stability condition of a WNCS were investigated. In [50], the optimal transmission scheduling of multiple control systems over shared communication resources was studied. In [102], the uplink and downlink transmission

scheduling problem of a WNCS with a half-duplex controller was considered. In [94], an event-triggered WNCS was proposed to reduce the communication cost. In [103] and [104], WNCSs with low-power and high-performance multi-hop wireless networks were investigated, respectively.

In most of the existing works on WNCSs [12, 23, 50, 94, 102, 105–107], the status of the physical process was discretized by periodical sampling, and the transmission of the controller's packet was ideally assumed to be fixed and equal to the sampling period. From the theory of channel encoding, if a message is encoded into a longer codeword (with a longer packet length), it can be delivered to the receiver with a higher reliability, but it introduces a longer transmission delay on the other side. This introduces the fundamental tradeoff in transmission delay and reliability [74]. In a WNCS, the transmission of a short control-information-carrying packet results in frequent but unreliable control, while the transmission of a long control packet leads to a less timely but more reliable control. Packet-length design to achieve an optimal control-system performance has rarely been considered in the existing literature on WNCSs.

Moreover, a WNCS is a dynamic system, and the state of the physical process under control changes with time. Naturally, different status of the WNCS can require different levels of reliability and delay of control-packet transmission for achieving the optimal control performance. For some statuses, reliable transmission is more important, which needs a longer control packet; while for others a short-delay transmission is more important, which needs a shorter control packet. Variable-length packet transmission has been proposed and investigated in conventional communication systems [108, 109], which, however, has not been considered in WNCSs. In this

chapter, we will tackle the packet-length design problem in WNCSs.

4.2 System Model

Consider a wireless networked control system where the discrete-time dynamic physical process $\mathbf{x}_t \in \mathbb{R}^n$, $t \in \mathbb{N}$ is measured by the *controller*, which generates and sends control information to the remote *actuator* to control the process. $\mathbb{N}$ is the set of positive integers. The evolution of the dynamic process is modeled as a linear time-invariant system [12, 50, 94, 102]:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t, \tag{4.2.1}$$

where $\mathbf{u}_t \in \mathbb{R}^m$ is the actuator's control input, $\mathbf{w}_t \in \mathbb{R}^n$ is the process disturbance modelled as a zero-mean Gaussian white noise with the covariance $\mathbf{R} \in \mathbb{R}^{n\times n}$, and $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$ are the system transition matrix and the input matrix, respectively. The discrete time slot has a duration of T_0 s, which is also the sampling period of the process. For brevity, we only use the discrete time for analysis in the rest of the chapter.

4.2.1 Controller-Side Operation

To deliver the control information to the actuator, the controller converts its control signal into a packet by quantization and channel encoding (i.e., error-control coding). The communication channel for packet transmission is static for low-mobility industrial control applications [12, 94, 102]. We assume that the quantization noise is negligible due to the sufficiently high number of quantization levels, which is commonly considered in the literature [12, 23, 50, 94, 102, 105, 106]. Since a longer

channel-coding blocklength leads to a longer packet with a higher reliability [74], the packet error probability is a monotonically decreasing function $g(l)$ in terms of the packet length of l time slots, where $l \in \mathbb{N}$.

The transmission of control information introduces delay, so we adopt a predictive control method for delay compensation [94, 102]. To be specific, given the current time t and packet length l, since the control information is expected to be delivered at time $(t + l - 1)$ and there is no control input until then, the controller optimally predicts the process state $\mathbf{x}_{t+l-1}$ as [102]

$$\hat{\mathbf{x}}_{t+l-1|t} = \mathbf{A}^{l-1}\mathbf{x}_t. \tag{4.2.2}$$

By adopting a linear control law, the control signal that is generated at time t and to be applied at time $(t + l - 1)$ by the actuator is [94, 102]

$$\hat{\mathbf{u}}_{t+l-1|t} = \mathbf{K}\hat{\mathbf{x}}_{t+l-1|t} = \mathbf{K}\mathbf{A}^{l-1}\mathbf{x}_t, \tag{4.2.3}$$

where $\mathbf{K} \in \mathbb{R}^{m\times n}$ is the constant controller gain.

Assumption 4.1. The controller gain has the property that [102]

$$\mathbf{A} + \mathbf{B}\mathbf{K} = \mathbf{0}. \tag{4.2.4}$$

This assumes that the control system is one-step controllable[1], i.e., it can be verified by taking (4.2.3) into (4.2.1) so that the state vector $\mathbf{x}_t$ can be driven to zero in one time slot in the absence of the process disturbance with packet length $l = 1$.

[1]Multi-step controllable cases can also be handled by the following problem formulation and performance analysis framework, and the stability conditions in Theorems 4.1 and 4.2 remain the same.

4.2.2 Actuator-Side Operation

Let $\gamma_t = 1$ denote a successful packet detection at time t and $\tilde{l}_t$ denote the length of the successfully received packet. Thus, $\gamma_t = 0$ denotes the packet has not arrived at t or the detection of the arrived packet at t has failed.

Now, we introduce the age-of-information (AoI) at the actuator, d_t, which measures the time duration between the generation time of the most recently received control packet and the current time t [62, 105]. Then, it is easy to have the updating rule of d_t as

$$d_{t+1} = \begin{cases} \tilde{l}_t, & \gamma_t = 1, \\ d_t + 1, & \text{otherwise.} \end{cases} \tag{4.2.5}$$

The actuator adopts a zero-hold strategy: it remains the zero control input until a control packet is successfully detected [12, 50, 102]. Thus, the actuator's control input $\mathbf{u}_t$ in (4.2.1) is given as

$$\mathbf{u}_t = \begin{cases} \hat{\mathbf{u}}_{t|t-\tilde{l}_t+1}, & \gamma_t = 1 \\ \mathbf{0}, & \text{otherwise.} \end{cases} \tag{4.2.6}$$

Taking (4.2.6) into (4.2.1) and using the property (4.2.4), the state covariance matrix can be obtained as

$$\mathbf{P}_t \triangleq \mathbb{E}\left[\mathbf{x}_t \mathbf{x}_t^\top\right] = \mathbf{H}(d_t) \triangleq \sum_{i=0}^{d_t-1} \mathbf{A}^i \mathbf{R} (\mathbf{A}^i)^\top, \tag{4.2.7}$$

where $\mathbb{E}[\cdot]$ is the expectation operator, and $(\cdot)^\top$ is the operator of the matrix transpose. Therefore, the state covariance matrix $\mathbf{P}_t$ in (4.2.7) depends on the AoI status d_t.

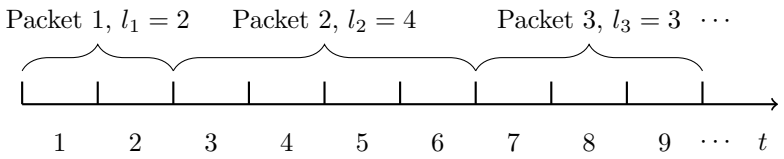


Figure 4.1: Variable-length packet transmission policy.

4.3 Control with Variable-Length Packets

4.3.1 Problem Formulation

The performance of the control system is measured by the quadratic average cost function as [12, 94, 102]

$$J = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\mathbf{x}_t^\top \mathbf{Q}\mathbf{x}_t\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathrm{Tr}\left(\mathbf{Q}\mathbf{P}_t\right), \tag{4.3.1}$$

where $\mathbf{Q}$ is a symmetric positive semidefinite weighting matrix, and $\mathrm{Tr}\left(\mathbf{Q}\mathbf{P}_t\right)$ is the cost function at time t.

We define the variable-length packet transmission policy for wireless control: the policy $\pi = \{l_1, l_2, .., l_k, ...\}$ is the sequence of the packet lengths during process control, where $l_k \in \mathbb{N}$ and k is the packet index, as illustrated in Fig. 4.1. Our problem is to find the optimal policy π^* that minimizes the infinite-horizon average cost, i.e.,

$$\pi^\star \triangleq \underset{\pi}{\mathrm{argmin}} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathrm{Tr}\left(\mathbf{Q}\mathbf{P}_t\right). \tag{4.3.2}$$

4.3.2 Semi-MDP Solution

From the definition of $\mathbf{P}_t$ in (4.2.7) and the updating rule of the AoI status (4.2.5), the problem (4.3.2) can be treated as an adaptive packet-length decision process with two properties: 1) if a decision of packet length l is made at the AoI state d, then each step's cost $\mathrm{Tr}(\mathbf{Q}\mathbf{P}_t)$ depends only on the present AoI state d until the

completion of packet transmission; 2) the sum cost incurred until completion of the packet transmission depends only on the present AoI state d and the packet length l.

Such an average-cost minimization problem is a typical semi-Markov decision process (semi MDP) [110], modelled as:

1) The state space is defined as $\mathbb{S} \triangleq \{d : d \in \mathbb{N}\}$. The state indicates the AoI at the beginning of a packet transmission. The state at the beginning of the kth packet is denoted as $d_k \in \mathbb{S}$.

2) The action space is defined as $\mathbb{A} \triangleq \{l : l \in \mathbb{N}\}$. The action $l_k \triangleq \pi(d_k)$ represents the length of the kth packet, where, with a slight abuse of notation, $\pi(d_k)$ is the stationary policy function in terms of the current state d_k.

3) The state-transition probability $P(d'|d, l)$ characterizes the probability that state transits from d at the beginning of the current packet to d' at the beginning of the next packet with the current action of l. As the transition is time-homogeneous, we drop the packet index k here. From the state-updating rule (4.2.5) and the packet error probability function $g(l)$, the state transition probability is:

$$P(d'|d, l) = \begin{cases} g(l) & \text{if } d' = d + l \\ 1 - g(l) & \text{if } d' = l. \end{cases} \tag{4.3.3}$$

4) The duration time $\delta(d, l)$ characterizes the expected time until the next transmission decision if the action l is chosen at the current state d. It is clear that the duration time is determined by the decided packet length in our scenario, i.e.,

$$\delta(d, l) = l. \tag{4.3.4}$$

5) The one-stage cost of the semi MDP problem is the sum cost during the current packet transmission, which is given as:

$$c(d, l) \triangleq \sum_{i=d}^{d+l-1} \mathrm{Tr}\left(\mathbf{QH}(i)\right). \tag{4.3.5}$$

From the semi-MDP formulation, the average cost J in (4.3.2) can be rewritten as:

$$J = \frac{\sum_{d\in\mathbb{S}} c(d, \pi(d))\phi_\pi(d)}{\sum_{d\in\mathbb{S}} \delta(d, \pi(d))\phi_\pi(d)}, \tag{4.3.6}$$

where $\phi_\pi(d)$ denotes the stationary probability of state $d \in \mathbb{S}$ under policy π.

Therefore, the optimal policy of problem (4.3.2), $\pi^\star(\cdot)$, can be obtained by solving the above semi-MDP problem with the target function (4.3.6). By using the classical data-transformation method, the semi-MDP problem can be transformed as an MDP problem, and thus can be solved effectively by the classical relative value iteration method [110].

4.3.3 Practical Implementation Issues of Variable-Length Policy

Since the control packet length of a WNCS changes with time, each packet header should include the information of the packet length. Thus, compared with a fixed-length policy, the variable-length policy requires a slightly higher communication overhead in practice. Moreover, when considering multiple WNCSs sharing the same wireless resources, it is not applicable to consider a time-division multiple access such as the medium-access control (MAC) protocol for multi-WNCS scheduling, as each WNCS requires a dynamic time slot length for transmissions. Therefore, different WNCSs need to be allocated to different frequency channels/sub-carriers when applying the variable-length policy. Commonly considered error-control codes, including polar codes and turbo codes, can be used for variable-length encoding.

4.4 Stability Condition of the Control System

If the packet transmissions are very unreliable with different packet lengths, the average cost J in (4.3.6) might be unbounded no matter what packet-transmission policy we choose and the semi-MDP problem discussed above might not have a feasible solution, i.e., the control system is *unstable*.

Now, we study the stability condition of the control system by investigating the fixed-length and variable-length packet transmission policies in the sequel, where the latter is a more general case of the former.

4.4.1 Fixed-Length Packet Transmission Policy

Theorem 4.1 (Fixed-length scenario)**.** Consider the control system described by (4.2.1)-(4.2.6). Let $(\mathbf{A}, \sqrt{\mathbf{R}})$ be controllable and let $(\mathbf{A}, \sqrt{\mathbf{Q}})$ be observable.[2] Assuming that the packet length is l and fixed during the process control, the dynamic process can be stabilized iff

$$g(l)\rho^{2l}(\mathbf{A}) < 1, \tag{4.4.1}$$

where $\rho(\mathbf{A})$ is the spectral radius of the matrix $\mathbf{A}$.

Proof. Consider the policy with fixed packet length $l_0 \in \mathbb{N}$, i.e., $\pi(d) = l_0, \forall d \in \mathbb{S}$. From (4.2.5), it is easy to see that the state space in Section 4.3.2 is degraded into $\mathbb{S} = \{l_0, 2l_0, 3l_0, \cdots\}$. Since the packet error probability is fixed, it can be proved that the process of the AoI, $\{d_k\}$, has a stationary distribution as

$$\phi_\pi\big(il_0\big) = (1 - g(l_0))g(l_0)^{i-1}, i = 1, 2, \cdots \tag{4.4.2}$$

[2] $\sqrt{\mathbf{R}}$ and $\sqrt{\mathbf{Q}}$ are the square roots of the positive definite matrices $\mathbf{R}$ and $\mathbf{Q}$, respectively. $(\mathbf{A}, \sqrt{\mathbf{R}})$ is controllable and $(\mathbf{A}, \sqrt{\mathbf{Q}})$ is observable if $\left[\sqrt{\mathbf{R}}, \mathbf{A}\sqrt{\mathbf{R}}, \cdots, \mathbf{A}^n\sqrt{\mathbf{R}}\right]$ and $\left[\sqrt{\mathbf{Q}}^\top, \mathbf{A}^\top\sqrt{\mathbf{Q}}^\top, \cdots, (\mathbf{A}^n)^\top\sqrt{\mathbf{Q}}^\top\right]$ are of full rank.

From (4.3.4) and (4.4.2), it is clear that the denominator of (4.3.6) is bounded. Thus, the average cost J is bounded iff the numerator of (4.3.6) is. Using the inequalities below [102, Lemma 1],

$$\mathrm{Tr}\left(\mathbf{QH}(d)\right) \leq c(d,l) \leq l\mathrm{Tr}\left(\mathbf{QH}(d+l)\right), \tag{4.4.3}$$

we can obtain

$$\sum_{i=1}^{\infty} c(il_0, l_0)\phi_\pi\big(il_0\big) < \frac{l_0(1-g(l_0))}{g^2(l_0)} \sum_{i=2}^{\infty} g^i(l_0)\mathrm{Tr}\left(\mathbf{QH}(il_0)\right), \tag{4.4.4}$$

$$\sum_{i=1}^{\infty} c(il_0, l_0)\phi_\pi\big(il_0\big) \geq (1-g(l_0)) \sum_{i=1}^{\infty} g^i(l_0)\mathrm{Tr}\left(\mathbf{QH}(il_0)\right). \tag{4.4.5}$$

From [12], if $q > 0$ and $(\mathbf{A}, \sqrt{\mathbf{R}})$ and $(\mathbf{A}, \sqrt{\mathbf{Q}})$ are controllable and observable, respectively, the following property holds:

$$\sum_{i=1}^{\infty} q^i \mathrm{Tr}\left(\mathbf{QH}(i)\right) < \infty \text{ iff } q\rho^2(\mathbf{A}) < 1. \tag{4.4.6}$$

Applying (4.4.6) to (4.4.4) and (4.4.5), it can be obtained that the average cost J is bounded iff $\left(g(l_0)\right)^{1/l_0} \rho^2(\mathbf{A}) < 1$, which completes the proof of Theorem 4.1. □

Remark 4.1. Theorem 4.1 says that the stability condition under the fixed-length policy depends on the packet error probability, the length of the packet and the control system parameter. The process (4.2.1) can be stabilized if the packet length l is properly chosen such that both the packet error probability $g(l)$ and the lth power of $\rho^2(\mathbf{A})$ are small.

4.4.2 Variable-Length Packet Transmission Policy

Theorem 4.2 (Variable-length scenario)**.** Consider the same system and conditions as defined in Theorem 4.1. There exists a stationary and deterministic variable-length

packet transmission policy that can stabilize the dynamic process iff

$$\min_{l\in\mathbb{N}} g(l)\rho^{2l}(\mathbf{A}) < 1. \tag{4.4.7}$$

Proof. The sufficiency is easy to prove based on Theorem 4.1, as the optimal variable-length policy results in an average cost no higher than that of a fixed-length policy.

We use a constructive method to prove the necessity. First, we consider a virtual updating rule below to replace (4.2.5)

$$d_{t+1} = \begin{cases} 1, & \gamma_t = 1 \\ d_t + 1, & \text{otherwise.} \end{cases} \tag{4.4.8}$$

It is clear that (4.4.8) is no larger than (4.2.5) for all $t \in \mathbb{N}$, and hence the average cost of the optimal packet-transmission policy by using the updating rule (4.4.8) is no higher than that of (4.2.5). Then, we will show that (4.4.7) holds if the average cost induced by a policy is bounded under the condition of (4.4.8).

Consider a general policy $\pi'(\cdot)$ that has $\pi'(1) = l'_1$. The state space can be rewritten as $\mathbb{S} = \{1, 1+l'_1, 1+l'_1+l'_2, 1+l'_1+l'_2+l'_3, \cdots\}$, where $l'_k = \pi'(1+\sum_{i=1}^{k-1} l'_i), \forall k \in \{2, 3, \cdots\}$. The average cost function in (4.3.6) with packet-transmission policy $\pi'(\cdot)$ can be rewritten as

$$J = \frac{\sum_{i=1}^{\infty} c'\left(\sum_{j=1}^{i} l'_j\right) \phi'_{\pi'}\left(\sum_{j=1}^{i} l'_j\right)}{\sum_{i=1}^{\infty} \left(\sum_{j=1}^{i} l'_j\right) \phi'_{\pi'}\left(\sum_{j=1}^{i} l'_j\right)}, \tag{4.4.9}$$

where

$$c'(l) = c(1,l) \geq \mathrm{Tr}\left(\mathbf{QH}(l)\right), \tag{4.4.10}$$

$$\phi'_{\pi'}\left(\sum_{j=1}^{i} l'_j\right) = \prod_{j=1}^{i-1} g(l'_j)\left(1 - g(l'_i)\right) \geq \left(1 - g(1)\right)\prod_{j=1}^{i} g(l'_j). \tag{4.4.11}$$

Since the function $c'(l)$ grows exponentially fast with l, J in (4.4.9) is bounded iff the numerator is. Then, by using the inequalities (4.4.10) and (4.4.11), the numerator of

(4.4.9) is lower bounded by

$$(1-g(1))\sum_{i=1}^{\infty}\prod_{j=1}^{i}g(l'_j)\mathrm{Tr}(\mathbf{QH}\left(\sum_{j=1}^{i}l'_j\right)). \tag{4.4.12}$$

From (4.4.6), it can be proved that $\mathrm{Tr}\left(\mathbf{QH}(i)\right)$ increases as fast as $\rho^{2i}(\mathbf{A})$ when $i\to\infty$. Thus, $\mathrm{Tr}\left(\mathbf{QH}\left(\sum_{j=1}^{i}l'_j\right)\right)$ can be approximated by $\eta\rho^{2\left(\sum_{j=1}^{i}l'_j\right)}(\mathbf{A})$ when i is large, where $\eta>0$. Thus, if (4.4.12) is bounded, $\min_{j\in\mathbb{N}}g(l'_j)\rho^{2l'_j}(\mathbf{A})<1$ holds, which completes the proof of Theorem 4.2. □

Remark 4.2. Theorem 4.2 shows that the stability condition of variable-length packet transmission policy is looser than that of a fixed-length policy in Theorem 4.1. The stability condition depends on the function of the packet error probability $g(l)$ and also the system parameter $\mathbf{A}$.

4.5 Numerical Results

In this section, we numerically evaluate the optimal variable-length packet transmission policy in the WNCS and compare it with the fixed-length packet transmission policies. In order to find the optimal policy, we need to solve the semi-MDP problem with finite state and action spaces. Thus, the infinite state space $\mathbb{S}$ is truncated as $\mathbb{S}=\{1,\cdots,N\}$. The action space is $\mathbb{A}=\{1,\cdots,M\}$. The function of the packet error probability in terms of packet length is approximated by an exponential function as $g(l)=0.8\times0.5^{l-1}$ [77, 105, 111]. Unless otherwise stated, we set $N=70$ and $M=5$ for solving the variable-length policy, and consider a scalar system [23, 112, 113], where $\mathbf{A}=1.2$, $\mathbf{B}=1$, $\mathbf{R}=1$, $\mathbf{Q}=1$ [112], and thus $\rho(\mathbf{A})=1.2$ and $\mathbf{K}=-1.2$.

Fig. 4.2 shows the optimal packet-transmission policy of the semi-MDP problem with different truncated state-space cardinality N. It is interesting to see that when

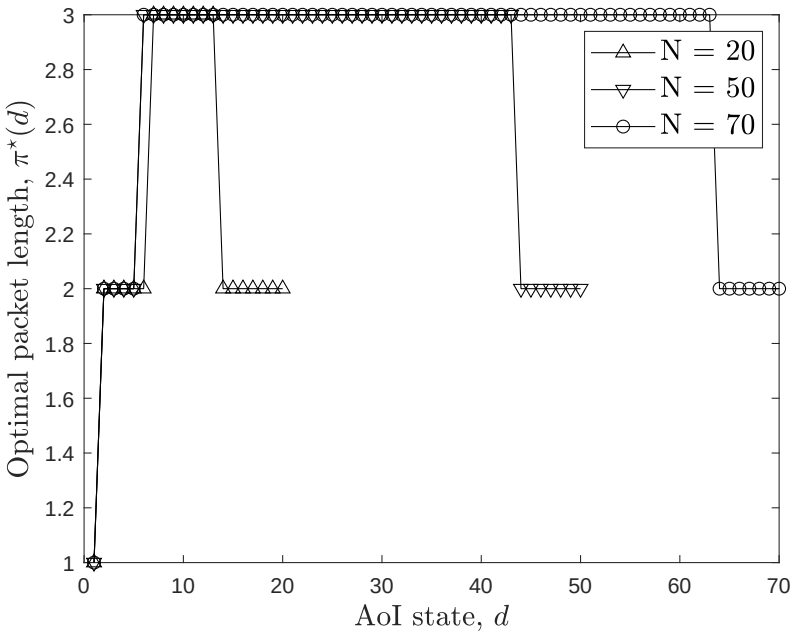


Figure 4.2: The optimal variable-length packet transmission policies within different truncated state spaces.

the AoI state is small ($d \leq 7$), the optimal packet lengths in different truncated state spaces are almost the same, and the optimal packet length increases with the increasing AoI state. Also, we see that when the state space is large (i.e., $N = 70$), which is closer to the ideal infinite state-space case, the optimal packet length tends to be invariant when $d > 7$. Thus, it is reasonable to infer that the optimal policy with the infinite state space has the property that $\pi^\star(d) = 3$ when $d > 7$. The structure of the optimal policy shows that when the current system AoI is pretty good, it is wise to take the risk of a transmission with a lower reliability to achieve good control quality, as the latter will not be too bad even if the transmission fails.

Fig. 4.3 plots the average costs of fixed-length packet transmission policies with different packet lengths and an optimal variable-length policy based on (4.3.1) with $T = 50,000$. From Theorem 4.1, it can be verified that when the fixed packet length $l = 1$, which is the conventional transmission policy in most of the existing work [12,

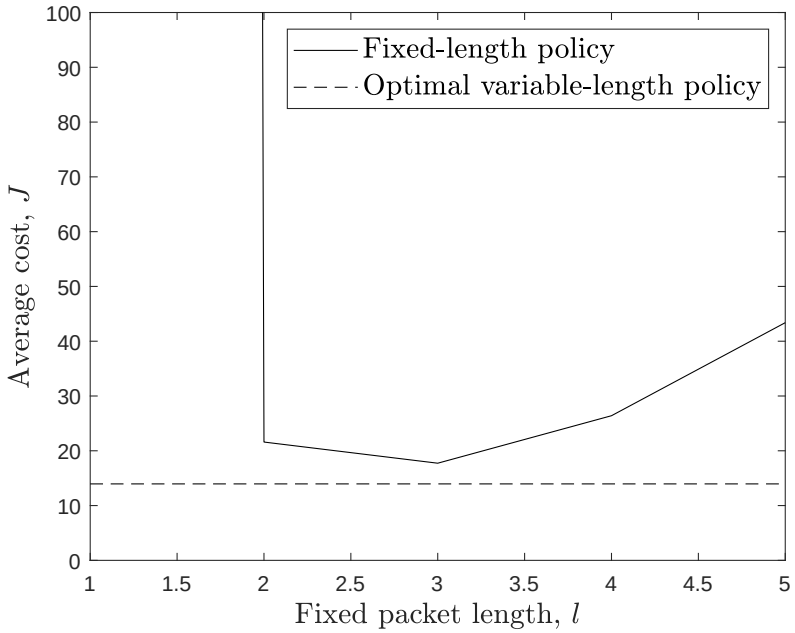


Figure 4.3: The average cost of fixed-length packet-transmission policy versus packet length, and the average costs of optimal variable-length policy.

23, 50, 94, 102, 105, 106], the control system is unstable. From Fig. 4.3, the system can be stabilized with longer transmission packets, and the average cost is minimized when the fixed packet length is 3. This optimal fixed-length policy is largely in agreement with the optimal variable-length policy illustrated in Fig. 4.2, where the optimal packet length is 3 for most of the states in the state space. Also, we see that the optimal variable-length policy gives a 22% average cost reduction of the optimal fixed-length policy, which shows the importance of adaptive packet-transmission design in WNCSs.

4.6 Conclusions

In this chapter, we have proposed and optimized the variable-length packet transmission policy. We have also derived the control-system stability conditions for both

the fixed-length and variable-length policies. Our numerical results have demonstrated the superiority of the proposed variable-length packet transmission method in wireless control systems.

Chapter 5

Conclusions and Future Work

In the thesis, we studied the transmission scheduling problems in terms of remote estimation, downlink-uplink schedule, and variable packet length control in wireless networked control for industrial IoT. These problems stem from the tradeoff between wireless transmission reliability and latency. This thesis answers a question about how to dynamically and efficiently balance the tradeoff, if it is unavoidable. Under different WNCS settings, the weighting of reliablity and latency would be various, which motivates our work. Below, we will summarise our main results and propose some possible extensions in future work.

5.1 Summary of Results

Firstly, we have proposed and optimized a HARQ-based remote estimation protocol for real-time applications where the sensor makes online decision on whether to send a new measurement or retransmit the previously failed one depending on both current estimation quality of the receiver and the current number of retransmissions of the sensor. We derive a switching-based structural property of the optimal policy

and give an easy-to-compute suboptimal policy. Our results have shown that the proposed optimal policy can significantly reduce the estimation MSE for some practical settings.

Secondly, we have proposed an important uplink-downlink transmission scheduling problem of a WNCS with an HD controller for practical IIoT applications, which has not been considered in the open literature. We have given a comprehensive analysis of the HD-controller-based WNCS for both one-step and v-step controllable plants and derived the necessary and sufficient condition of the existence of a stationary and deterministic optimal policy. We have also derived an easy-to-compute suboptimal policy, which provides a control performance close to the optimal policy and notably reduces the average cost of the plant compared to a naive alternative-scheduling policy.

Finally, we have proposed and optimized a variable-length packet transmission policy. The control-system stability conditions for both the fixed-length and variable-length policies have been derived. Our numerical results have demonstrated the superiority of the proposed variable-length packet transmission method in wireless control systems.

5.2 Future Work

We now propose some extensions of our problems and other potential works that could be investigated in the future.

In future work, the temporal correlations and differential encoding across transmissions of different estimates could be explored, rather than assuming each estimation is completely unrelated to any previous ones. Furthermore, a more practical

setting of remote estimation system should be considered where the exact model of channel-state transition is unknown to either the sensor or the receiver. To solve such a decision problem with unknown state-transition probabilities, we can seek a solution based on reinforcement learning (RL), which is a powerful tool to solve MDP problems especially when the state-transition model of the MDP is unknown. Besides, control input will also be included in our future work on HARQ-based remote estimation system.

In addition, the scenario where multiple plants are controlled by an HD controller for IIoT applications with a large number of devices has not been investigated and remains as an open challenge. It is also important to investigate the scheduling of different sensors' transmissions to the controller, and the controller's transmissions to different actuators, and consider different quality of service (QoS) requirements of different devices in the scheduling and how they affect the control. Moreover, for scheduling-policy design, it is more practical to take into account the transmission power constraints of the sensors and controller, which limit the reliability of wireless communication.

Moreover, a more general system with the combination of HARQ scheme, control policy and variable-length code design will be investigated. Due to the high complexity of the system model, machine learning based techniques would be prederable method to solve such problems.

Appendix A

Proofs for Chapter 2

A.1 Proof of Theorem 2.1

We adopt a constructive approach to prove the sufficient condition (2.4.3) on the existence of a stationary and deterministic optimal policy. Specifically, we use the technical corollary [114, Corollary 7.5.10] for the proof. Corollary 7.5.10 in [114] says that the optimal policy exists if (CAV*1) and (CAV*2) are both satisfied: (CAV*1) there exists a standard policy ψ such that the recurrent class R_ψ induced by ψ is equal to the whole state space $\mathbb{S}$; (CAV*2) given $U > 0$, the set $\mathbb{S}_U = \{s|c(s,a) \leq U \text{ for some } a\}$ is finite. Thus, in the following, we will prove that both (CAV*1) and (CAV*2) are satisfied if (2.4.3) holds. As a consequence, it proves that (2.4.3) is a sufficient condition of the existence of the optimal policy.

Condition (CAV*2) can be easily verified based on (2.4.2). In what follows, we verify (CAV*1) by first constructing a policy ψ and then proving that it is a standard policy.

The action of the policy ψ is given as

$$a = \begin{cases} 0, & \text{if } r = q \\ 1, & \text{otherwise.} \end{cases} \tag{A.1.1}$$

It is easy to prove that any state in $\mathbb{S}$ induced by ψ is a recurrent state. We then prove that ψ is a standard policy by verifying both the expected first passage cost and time from state $(r, q) \in \mathbb{S}$ to $(0, 0)$ are bounded [114]. Due to the space limitation, we only prove that any state with $r = q$ has bounded first passage cost and time. The other states can be proved similarly. For notational simplicity, the expected first passage cost of the state (i, i) is denoted as $d(i)$, and the one-stage cost (2.4.2) is rewritten as

$$c(q) \triangleq c((r, q), a) = \mathrm{Tr}\left(f^q(\bar{\mathbf{P}}_0)\right). \tag{A.1.2}$$

Based on (2.3.10), (A.1.1) and the law of total expectation of the first passage cost of all the possible first passage paths (as illustrated in Fig. 2.13), the expected first passage cost $d(i)$ can be obtained as

$$\begin{aligned} d(i) &= c(i) + (1 - g(1))c(1) + g(1)c(i+1) \\ &+ g(1)(1 - g(2))d(2) + g(1)g(2)c(i+2) \\ &+ g(1)g(2)(1 - g(3))d(3) + g(1)g(2)g(3)c(i+3) + \cdots \\ &= \nu(i) + (1 - g(1))c(1) + D, \forall i > 0, \end{aligned} \tag{A.1.3}$$

where $g(1) = \Lambda_0'$, $\nu(i) = c(i) + \sum_{j=1}^{\infty} \alpha_j c(i+j)$, $D = \sum_{j=2}^{\infty} \beta_j d(j)$, and $\alpha_j = \prod_{l=1}^{j} g(l)$ and $\beta_j = \prod_{l=1}^{j-1} g(l)(1 - g(j))$. Therefore, $d(i)$ is bounded if $\nu(i) < \infty$ and $D < \infty$. Using the inequality (2.3.12), we have

$$\alpha_j \leq \Lambda_0' \Lambda_0^{j-1}, \forall j \geq 1, \text{ and } \beta_j \leq \Lambda_0' \Lambda_0^{j-2}, \forall j \geq 2. \tag{A.1.4}$$

From [63], we have $\sum_{j=1}^{\infty} \Lambda_0^j c(j) < \infty$ iff $\Lambda_0 \rho^2(\mathbf{A}) < 1$. Thus, it is easy to prove that $\nu(i) < \infty$ if (2.4.3) holds.

From (A.1.3), D can be further derived after simplifications as

$$D = \frac{1}{1 - \sum_{i=2}^{\infty} \beta_i} \left(\sum_{i=2}^{\infty} \beta_i (1 - g(1)) c(1) + \sum_{i=2}^{\infty} \beta_i \nu(i) \right). \tag{A.1.5}$$

As $\sum_{i=2}^{\infty} \beta_i < g(1) < 1$, D is bounded as long as $\sum_{i=2}^{\infty} \beta_i \nu(i) < \infty$. Using the inequalities (A.1.4), after some simplifications, we have $\sum_{i=2}^{\infty} \beta_i \nu(i) \leq \eta \sum_{j=2}^{\infty} \Lambda_0^{j-1} c(j) + \eta^2 \sum_{j=3}^{\infty} (j-2) \Lambda_0^{j-1} c(j)$, where $\eta = \Lambda_0'/\Lambda_0$. It can be proved that $\sum_{j=3}^{\infty} (j-2) \Lambda_0^{j-1} c(j)$ is bounded if $\sum_{j=1}^{\infty} \Lambda_0^j c(j)$ is bounded. Again, using the result that $\sum_{j=1}^{\infty} \Lambda_0^j c(j) < \infty$ iff $\Lambda_0 \rho^2(\mathbf{A}) < 1$ in [63], $\sum_{i=1}^{\infty} \beta_i \nu(i) < \infty$ if $\Lambda_0 \rho^2(\mathbf{A}) < 1$, yielding the proof of the bounded expected first passage cost with condition (2.4.3). Similarly, we can verify that the expected first passage time is also bounded.

A.2 Proof of Theorem 2.2

The switching property is equivalent to the monotonicity of the optimal policy in r if q is fixed and in q if r is fixed. The monotonicity can be proved by verifying the following conditions (see Theorem 8.11.3 in [76]).

(1) $c(s, a)$ is nondecreasing in s for all $a \in \mathbb{A}$;

(2) $c(s, a)$ is a superadditive function on $\mathbb{S} \times \mathbb{A}$;

(3) $\tilde{P}(s'|s, a) = \sum_{i=s'}^{\infty} P(i|s, a)$ is nondecreasing in s for all $s' \in \mathbb{S}$ and $a \in \mathbb{A}$;

(4) $\tilde{P}(s'|s, a)$ is a superadditive function on $\mathbb{S} \times \mathbb{A}$ for all $s' \in \mathbb{S}$.

We first prove the monotonicity in r with q fixed. The state s is ordered by r, i.e., if $r^- \leq r^+$, we define $s^- \leq s^+$ with $s^- = (r^-, q)$ and $s^+ = (r^+, q)$. From the definition of one-stage cost, $c(s, a)$ is increasing in q. Therefore, condition (1) can be easily verified. For condition (2), the superadditive function is defined in (4.7.1) of [76]. A function $f(x, y)$ is superadditive for $x^- \leq x^+$ and $y^- \leq y^+$, if

$f(x^+, y^+) + f(x^-, y^-) \geq f(x^+, y^-) + f(x^-, y^+)$. Then, condition (2) can be easily verified as $c(s, a)$ is independent of a.

Given the current state $s = (r, q)$, from (2.4.1), the next possible states are $s_0 \triangleq (0, 0)$, $s_1 \triangleq (0, q+1)$, $s_2 \triangleq (r+1, r+1)$ and $s_3 \triangleq (r+1, q+1)$. Let $s' \triangleq \{(r', q') : q \in \mathbb{N}\}$. If $r' \leq r$, we define $s' \preceq s$ with $s = (r, q)$. Based on (2.4.1), $\tilde{P}(s'|s, a)$ with different actions are given as

$$\tilde{P}(s'|s, a=0) = \begin{cases} 1, & \text{if } s' \preceq s_0 \\ 0, & \text{otherwise} \end{cases}, \tag{A.2.1}$$

and

$$\tilde{P}(s'|s, a=1) = \begin{cases} 1, & \text{if } s' \preceq s_2 \\ 0, & \text{otherwise} \end{cases}. \tag{A.2.2}$$

Therefore, condition (3) can be easily verified.

For condition (4), let $s^+ = (r^+, q)$, $s^- = (r^-, q)$, $r^+ \geq r^-$ and $a^+ \geq a^-$ Then, we need to verify if $\tilde{P}(s'|s^+, a^+) + \tilde{P}(s'|s^-, a^-) \geq \tilde{P}(s'|s^+, a^-) + \tilde{P}(s'|s^-, a^+)$. Based on the definitions of $\tilde{P}(s'|s, a)$, s' and s_i, $i = 0, 1, 2, 3$, condition (4) can be verified straightforwardly. As all four conditions hold, the monotonicity of the optimal policy in r is proved. Similarly, the monotonicity of the optimal policy in q can be proved.

A.3 Proof of Theorem 2.3

The proof of the sufficient condition (2.5.2) follows the same steps of Theorem 2.1: 1) construction of a stationary policy in state space $\mathbb{S}$, 2) providing useful technical lemmas for problem transformation and 3) deriving a sufficient condition in terms of the packet error probability such that the long-term average cost of the stationary policy is bounded, completing the proof of existence condition of Theorem 2.3.

Step 1. Inspired by the proof of the static channel scenario, where the constructed stationary policy (A.1.1) is simply the policy that a retransmission is always required until the a successful transmission occurs, we consider a similar policy in the Markov channel scenario as

$$a = \psi(\mathbf{\Omega}, q, \Xi) = \begin{cases} 0, & \text{if } \|\mathbf{\Omega}\|_1 = q \\ 1, & \text{otherwise,} \end{cases} \tag{A.3.1}$$

We can prove that ψ is a stationary policy in the state space $\mathbb{S}$. In what follows, we prove that the long-term average cost induced by policy ψ is bounded if (2.5.2) holds, completing the proof of existence condition of Theorem 2.3.

As the state $s = (\mathbf{\Omega}, q, \Xi)$ has $B + 2$ dimensions, it is not possible to analyze the average cost directly and thus we have to reduce the dimension of the state space.

Step 2. Some useful lemmas.

Lemma A.1. Given the policy ψ and the packet loss function

$$\mathbb{P}\left[\gamma_k = 0\right] = \begin{cases} \tilde{g}(\mathbf{0}, \Xi_k), & a_k = 0 \\ \tilde{g}'(\mathbf{\Omega}_k, \Xi_k), & a_k = 1 \end{cases} \tag{A.3.2}$$

where $\tilde{g}(\mathbf{\Omega}_k, \Xi_k) < \tilde{g}'(\mathbf{\Omega}_k, \Xi_k) < 1$, $\forall k$, the average cost of the MDP with the packet loss function (2.3.9) is bounded, if that of (A.3.2) is.

The proof is straightforward due to the fact that a larger packet error probability results in a larger average cost, and thus is omitted here.

In the following, we derive a sufficient condition that can stabilize the average cost of the MDP with the packet loss function

$$\mathbb{P}\left[\gamma_k = 0\right] = \begin{cases} \Lambda_i' = \tilde{g}(\mathbf{0}, \Xi_k), & a_k = 0, \Xi_k = i \\ \Lambda_i \geq \tilde{g}(\mathbf{\Omega}_k, \Xi_k), & a_k = 1, \Xi_k = i. \end{cases} \tag{A.3.3}$$

Since the packet error function (A.3.3) does not depends on the individual elements of the state vector $\mathbf{\Omega}_k$, the state space $\mathbb{S}$ is reduced to a three-dimensional space

$\mathbb{S}' \triangleq \{(r, q, \Xi) : r \leq q,\ (r, q) \in \mathbb{N} \times \mathbb{N}, \Xi \in \{1, 2, ..., B\}\}$, where $r \triangleq \|\mathbf{\Omega}\|_1$.

Again using [114, Corollary 7.5.10], we only need to proof that both (CAV*1) and (CAV*2) hold in the new state space $\mathbb{S}'$ under the condition (2.5.2). Similar to the proof of Theorem 2.1, it can be easily proved that the recurrent class R_ψ induced by ψ is equal to the whole state space $\mathbb{S}'$, and (CAV*2) holds. Thus, in what follows, we only need to prove that the policy ψ is a standard policy, i.e., both the expected first passage cost and time from any state in $\mathbb{S}'$ to the state $(1, 1, 1)$ is bounded if (2.5.2) is satisfied.

Lemma A.2. If the first passage cost from any state $s \in \mathbb{S}'$ to the set of states

$$\mathbf{s} \triangleq \{(1, 1, 1), (1, 1, 2), ..., (1, 1, B)\} \tag{A.3.4}$$

is bounded, the first passage cost from any state $s \in \mathbb{S}'$ to state $(1, 1, 1)$ is bounded.

Proof. Let $A_{s,\mathbf{s}}$ be the expected first passage cost from s to the set $\mathbf{s}$. We have

$$A_{s,\mathbf{s}} = \sum_{i=1}^{B} \varrho_{s,i} A_{s,i}, \tag{A.3.5}$$

where $A_{s,i}$ is the expected first passage cost from s to the set $\mathbf{s}$ with the condition that the first visited state in $\mathbf{s}$ is $(1, 1, i)$, and $\varrho_{s,i}$ is the probability that the first visited state from s to the set $\mathbf{s}$ is $(1, 1, i)$.

As the state transition probabilities $p_{i,j}$ in (2.2.4) are larger than zero, it is clear that $\varrho_{s,i} \in (0, 1)$. Therefore, if $A_{s,\mathbf{s}}$ is bounded $\forall s \in \mathbf{s}$, $A_{s,i}$ is bounded $\forall s \in \mathbf{s}, i \in \{1, ..., B\}$.

The average first passage cost from s to the state $(1,1,1)$ can be written as

$$\begin{aligned} U_{s,1} &= \varrho_{s,1}A_{s,1} + \sum_{i=2}^{B} \varrho_{s,i}\varrho_{i,1}\left(A_{s,i} + A_{i,1}\right) \\ &+ \sum_{j=2}^{B}\sum_{i=2}^{B} \varrho_{s,j}\varrho_{j,i}\varrho_{i,1}\left(A_{s,j} + A_{j,i} + A_{i,1}\right) \\ &+ \sum_{j=2}^{B}\sum_{i=2}^{B}\sum_{k=2}^{B} \varrho_{s,j}\varrho_{j,k}\varrho_{k,i}\varrho_{i,1}\left(A_{s,j} + A_{j,k} + A_{k,i} + A_{i,1}\right) + \cdots \\ &< \varrho_{s,1}A_{\max} + (1-\varrho_{s,1})2A_{\max} + (1-\varrho_{s,1})(1-\varrho_{\min})3A_{\max} \\ &+ (1-\varrho_{s,1})(1-\varrho_{\min})^2 4A_{\max}\cdots \end{aligned} \tag{A.3.6}$$

where $A_{i,j}$ is the expected first passage cost from $(1,1,i)$ to $(1,1,j)$, the $\varrho_{j,k}$ is the probability that the first visited state from $(1,1,j)$ to the set $\mathbf{s}$ is $(1,1,k)$, $A_{\max} = \max_{i,j=1,2,\ldots,B}\{A_{s,i}, A_{i,j}\}$, and $\varrho_{\min} = \min_{i=1,2,\ldots,B}\{\varrho_{i,1}\}$. Since $(1-\varrho_{\min}) < 1$ and $A_{\max} < \infty$, it is straightforward that the right-hand side of the inequality (A.3.6) is bounded, completing the proof. □

Using Lemma A.2, in the following, we only need to prove that if (2.5.2) holds, the average first passage cost from any state to the state set $\mathbf{s}$ is bounded, which is also a sufficient condition that guarantees the average first passage cost from any state to a specific state, say $(1,1,1)$, is bounded.

Step 3. We define a length-B row vector $\vec{d}(i), \forall i \in \mathbb{N}$, where the kth element of $\vec{d}(i)$ is the average first passage cost from state (i,i,k) to the state set $\mathbf{s}$. Analogously

to the static channel scenario (A.1.3), we have the following equation

$$\begin{aligned}\vec{d}(i) &= \vec{c}(i) + \vec{c}(1)(\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda}'))+\vec{c}(i+1)\mathbf{\Pi}\mathbf{\Lambda}' \\ &+ \vec{d}(2)(\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda}))\mathbf{\Pi}(\mathbf{\Lambda}')+\vec{c}(i+2)(\mathbf{\Pi}(\mathbf{\Lambda}))\mathbf{\Pi}(\mathbf{\Lambda}') \\ &+ \vec{d}(3)(\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda}))(\mathbf{\Pi}(\mathbf{\Lambda}))\mathbf{\Pi}(\mathbf{\Lambda}') \\ &+ \vec{c}(i+3)(\mathbf{\Pi}(\mathbf{\Lambda}))(\mathbf{\Pi}(\mathbf{\Lambda}))\mathbf{\Pi}(\mathbf{\Lambda}')+\cdots \\ &= \vec{\nu}(i)+\vec{c}(1)\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda}')+\vec{D}, \forall i>0,\end{aligned} \tag{A.3.7}$$

where $\mathbf{I}$ is the B-by-B identity matrix, $\mathbf{\Lambda}' \triangleq \operatorname{diag}\{\Lambda'_1,...,\Lambda'_B\}$, $\vec{c}(i) \triangleq c(i)\vec{1}$, $\vec{1} \triangleq \underbrace{[1,1,...,1]}_{B}$, and

$$\vec{\nu}(i) = \vec{c}(i) + \sum_{j=1}^{\infty}\vec{c}(i+j)\check{\alpha}_j, \vec{D} = \sum_{j=2}^{\infty}\vec{d}(j)\check{\beta}_j, \tag{A.3.8}$$

$$\check{\alpha}_j = (\mathbf{\Pi}\mathbf{\Lambda})^{j-1}\mathbf{\Pi}\mathbf{\Lambda}', j \geq 1 \tag{A.3.9}$$

$$\check{\beta}_j = \mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda})(\mathbf{\Pi}\mathbf{\Lambda})^{j-2}\mathbf{\Pi}\mathbf{\Lambda}', j \geq 2. \tag{A.3.10}$$

Therefore, $\vec{d}(i)$ is bounded if $\vec{\nu}(i) \prec \infty$ and $\vec{D} \prec \infty$, where $\prec$ is the symbol of element-wise less than.

From the definitions in (A.3.9) and (A.3.10) and the property of Jordan normal forms of $\check{\alpha}_j$ and $\check{\beta}_j$ [115], we have the inequalities

$$\begin{aligned}\check{\alpha}_j &\preceq \kappa_1(j-1)^B\left[\rho^{j-1}(\mathbf{\Pi}\mathbf{\Lambda})\right]_{B\times B}, \forall j \geq 1, \text{ and} \\ \check{\beta}_j &\preceq \kappa_2(j-2)^B\left[\rho^{j-2}(\mathbf{\Pi}\mathbf{\Lambda})\right]_{B\times B}, \forall j \geq 2,\end{aligned} \tag{A.3.11}$$

where κ_1, κ_2 and κ_3 are positive constant, and $\preceq$ is the symbol of element-wise less than or equal to. From [46], we have the inequality of the cost function

$$c(j) \leq \kappa\rho^{2j}(\mathbf{A}), j > 0, \tag{A.3.12}$$

where κ is a positive constant. Thus, using the inequalities (A.3.11) and (A.3.12), it is easy to prove that

$$\vec{\nu}(i) \preceq \kappa_0\vec{c}(i) \prec \infty, \tag{A.3.13}$$

if (2.5.2) holds, where κ_0 is a positive constant.

From (A.3.7), we further have

$$\vec{D}\left(\mathbf{I}-\sum_{i=2}^{\infty}\check{\beta}_i\right)=\vec{c}(1)\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda}')\sum_{i=2}^{\infty}\check{\beta}_i+\sum_{i=2}^{\infty}\vec{\nu}(i)\check{\beta}_i. \tag{A.3.14}$$

Using the result of sum of a geometric series of matrices, it can be obtained as

$$\sum_{i=2}^{\infty}\check{\beta}_i=\mathbf{\Pi}(\mathbf{I}-\mathbf{\Lambda})(\mathbf{I}-\mathbf{\Pi}\mathbf{\Lambda})^{-1}\mathbf{\Pi}\mathbf{\Lambda}'\prec\infty. \tag{A.3.15}$$

From (A.3.10), it is clear that each element in $\sum_{i=2}^{\infty}\check{\beta}_i$ is positive. From (A.3.7), it can be proved that $\sum_{i=2}^{\infty}\check{\beta}_i$ is part of a state-transition matrix, i.e., the sum of all the matrices behind $\vec{c}_i$ and $\vec{d}_i$, $\forall i$. Therefore, the sum of each column of $\sum_{i=2}^{\infty}\check{\beta}_i$ is less than one. Using Perron-Frobenius Theorem [116], the spectral radius of $\sum_{i=2}^{\infty}\check{\beta}_i$ is less then one, and hence $\mathbf{I}-\sum_{i=2}^{\infty}\check{\beta}_i$ is invertible. Therefore, $\vec{D}$ in (A.3.14) is bounded if the term $\sum_{i=2}^{\infty}\vec{\nu}(i)\check{\beta}_i$ on the right-hand side of (A.3.14) is.

Using the inequalities (A.3.11), (A.3.12) and (A.3.13), after some simplifications, we have

$$\sum_{i=2}^{\infty}\vec{\nu}(i)\check{\beta}_i\preceq\kappa_0\kappa\kappa_2\sum_{i=2}^{\infty}\rho^{2i}(\mathbf{A})(i-2)^{B}\vec{1}\left[\rho^{i-2}(\mathbf{\Pi}\mathbf{\Lambda})\right]_{B\times B}. \tag{A.3.16}$$

Therefore, it is clear that the right-hand side of (A.3.16) is bounded if $\rho(\mathbf{\Pi}\mathbf{\Lambda})\rho^2(\mathbf{A})<1$, yielding the proof of the bounded expected first passage cost with condition (2.5.2). Similarly, we can verify that the expected first passage time is also bounded, completing the proof.

Appendix B

Proofs for Chapter 3

B.1 Proof of Proposition 3.1

Recall that $\eta_k^0 \triangleq \eta_{k-1}$. From the definition of η_k in (3.2.11), we have

$$\eta_{j+1}^0 = \begin{cases} 1, & j = k - \eta_k^0 \\ \eta_j^0 + 1, & j = k - \eta_k^0 + 1, \cdots, k-1 \end{cases} \tag{B.1.1}$$

By using the state-updating rule (3.3.14) for $\mathbf{x}_j$, $j = (k - \eta_k^0 + 1), \cdots, k$, we have

$$\begin{cases} \mathbf{x}_{k-\eta_k^0+1} = \mathbf{A}\mathbf{x}_{k-\eta_k^0} + \mathbf{B}\mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})^0 \hat{\mathbf{x}}_{k-\eta_k^0} + \mathbf{w}_{k-\eta_k^0} \\ \mathbf{x}_{k-\eta_k^0+2} = \mathbf{A}\mathbf{x}_{k-\eta_k^0+1} + \mathbf{B}\mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})^1 \hat{\mathbf{x}}_{k-\eta_k^0} + \mathbf{w}_{k-\eta_k^0+1} \\ \qquad \vdots \\ \mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{K}(\mathbf{A}+\mathbf{B}\mathbf{K})^{\eta_k^0-1} \hat{\mathbf{x}}_{k-\eta_k^0} + \mathbf{w}_{k-1} \end{cases} \tag{B.1.2}$$

Substituting $\mathbf{x}_{k-\eta_k^0+1}$ into $\mathbf{x}_{k-\eta_k^0+2}$ and so on, it can be shown that

$$\begin{aligned} \mathbf{x}_k = (\mathbf{A}+\mathbf{B}\mathbf{K})^{\eta_k^0} \mathbf{x}_{k-\eta_k^0} + (\mathbf{A}^{\eta_k^0} - (\mathbf{A}+\mathbf{B}\mathbf{K})^{\eta_k^0}) \mathbf{e}_{k-\eta_k^0} \\ + \sum_{i=1}^{\eta_k^0} \mathbf{A}^{i-1} \mathbf{w}_{k-i}. \end{aligned} \tag{B.1.3}$$

Using the new state-updating rule (B.1.3), $\mathbf{x}_k$ can be further rewritten as

$$\begin{aligned}
\mathbf{x}_k =& (\mathbf{A}+\mathbf{BK})^{\eta_k^0}\mathbf{x}_{t_k^1}+(\mathbf{A}^{\eta_k^0}-(\mathbf{A}+\mathbf{BK})^{\eta_k^0})\mathbf{e}_{t_k^1}+\sum_{i=1}^{\eta_k^0}\mathbf{A}^{i-1}\mathbf{w}_{k-i}\\
=&(\mathbf{A}+\mathbf{BK})^{\eta_k^0}\big((\mathbf{A}+\mathbf{BK})^{\eta_k^1}\mathbf{x}_{t_k^2}+(\mathbf{A}^{\eta_k^1}-(\mathbf{A}+\mathbf{BK})^{\eta_k^1})\mathbf{e}_{t_k^2}\\
&+\sum_{i=1}^{\eta_k^1}\mathbf{A}^{i-1}\mathbf{w}_{t_k^1-i}\big)+(\mathbf{A}^{\eta_k^0}-(\mathbf{A}+\mathbf{BK})^{\eta_k^0})\mathbf{e}_{t_k^1}+\sum_{i=1}^{\eta_k^0}\mathbf{A}^{i-1}\mathbf{w}_{k-i}\\
=&(\mathbf{A}+\mathbf{BK})^{\eta_k^0+\eta_k^1}\mathbf{x}_{t_k^2}+\sum_{i=1}^{\eta_k^0}\mathbf{A}^{i-1}\mathbf{w}_{k-i}+(\mathbf{A}+\mathbf{BK})^{\eta_k^0}\sum_{i=1}^{\eta_k^1}\mathbf{A}^{i-1}\mathbf{w}_{t_k^1-i}\\
&+(\mathbf{A}^{\eta_k^0}-(\mathbf{A}+\mathbf{BK})^{\eta_k^0})\mathbf{e}_{t_k^1}+(\mathbf{A}+\mathbf{BK})^{\eta_k^0}(\mathbf{A}^{\eta_k^1}-(\mathbf{A}+\mathbf{BK})^{\eta_k^1})\mathbf{e}_{t_k^2}\\
=&\ (\mathbf{A}+\mathbf{BK})^{\eta_k^0+\eta_k^1+\cdots+\eta_k^{v-1}}\mathbf{x}_{t_k^v}+\mathbf{w}'+\mathbf{e}' \qquad \text{(B.1.4)}\\
=&\ \mathbf{w}'+\mathbf{e}',
\end{aligned}$$

where the last step is due to the fact that $\eta_k^0+\eta_k^1+\cdots+\eta_k^{v-1}\geq v$ as $\eta_k^i\geq 1,\forall i\geq 0$, and $(\mathbf{A}+\mathbf{BK})^v=\mathbf{0}$, and

$$\begin{aligned}
\mathbf{w}' =& \sum_{i=1}^{\eta_k^0}\mathbf{A}^{i-1}\mathbf{w}_{k-i}+(\mathbf{A}+\mathbf{BK})^{\eta_k^0}\sum_{i=1}^{\eta_k^1}\mathbf{A}^{i-1}\mathbf{w}_{k-i}\\
&+\cdots+(\mathbf{A}+\mathbf{BK})^{\eta_k^0+\cdots+\eta_k^{v-2}}\sum_{i=1}^{\eta_k^{v-1}}\mathbf{A}^{i-1}\mathbf{w}_{t_k^{v-1}-i},
\end{aligned} \qquad \text{(B.1.5)}$$

$$\mathbf{e}_{t_k^j}=\sum_{i=1}^{\tau_k^j}\mathbf{A}^{i-1}\mathbf{w}_{t_k^j-i}, j=1,\cdots,v, \qquad \text{(B.1.6)}$$

$$\begin{aligned}
\mathbf{e}' =& (\mathbf{A}^{\eta_k^0}-(\mathbf{A}+\mathbf{BK})^{\eta_k^0})\mathbf{e}_{t_k^1}\\
&+(\mathbf{A}+\mathbf{BK})^{\eta_k^0}(\mathbf{A}^{\eta_k^1}-(\mathbf{A}+\mathbf{BK})^{\eta_k^1})\mathbf{e}_{t_k^2}+\cdots+\\
&(\mathbf{A}+\mathbf{BK})^{\eta_k^0+\cdots+\eta_k^{v-2}}(\mathbf{A}^{\eta_k^{v-1}}-(\mathbf{A}+\mathbf{BK})^{\eta_k^{v-1}})\mathbf{e}_{t_k^v}.
\end{aligned} \qquad \text{(B.1.7)}$$

We see that $\mathbf{x}_k$ only depends on the noise terms in the time range

$$\mathcal{S}\triangleq\left[k-(\eta_k^0+\cdots+\eta_k^{v-1})-\tau_v, k-1\right]. \qquad \text{(B.1.8)}$$

To further simplify (B.1.4), we consider three complementary cases: 1) $\tau_k^i <$

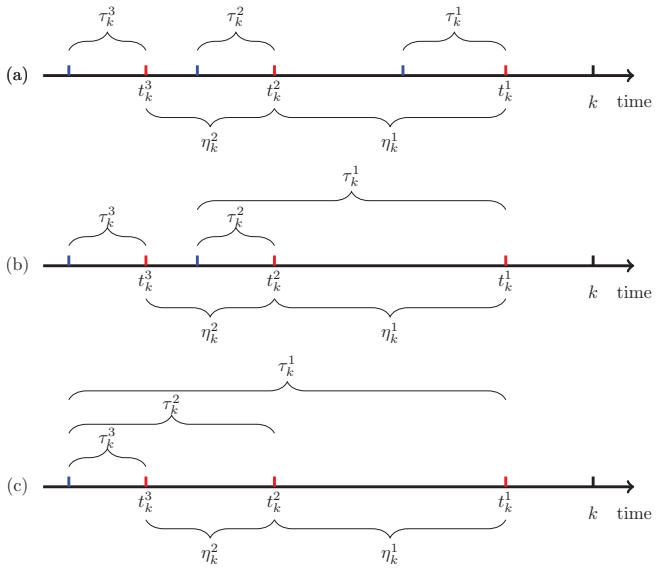


Figure B.1: Illustration of three different cases for analyzing the plant-state covariance, where red vertical bars denote successful controller's transmissions and blue vertical bars denote the most recent successful sensor's transmissions prior to the successful controller's transmissions.

$\eta_k^i, \forall i = 1, \cdots, v-1$, i.e., a sensor's successful transmission occurred between two consecutive controller's successful transmissions, as illustrated in Fig. B.1(a); 2) there exists i such that $\tau_k^i \geq \eta_k^i$ and there also exists j such that $\tau_k^j < \eta_k^j$ where $i, j \in \{1, \cdots, v-1\}$, i.e., a sensor's successful transmission did not always occur between two consecutive controller's successful transmissions, as illustrated in Fig. B.1(b). Note that from the definition of τ_k^j and η_k^j, $\tau_k^i = \eta_k^i + \tau_k^{i+1}$ if $\tau_k^i > \eta_k^i$; 3) $\tau_k^i = \eta_k^i + \tau_k^{i+1} \geq \eta_k^i$ for all $i \in \{0, \cdots, v-1\}$, i.e., a sensor's successful transmission never occur between the first and the vth controller's successful transmissions prior to the current time slot k, as illustrated in Fig. B.1(c).

For case 1), $\mathbf{e}_{t_k^i}$ contains the noise terms within time slots $t_k^i - \tau_k^i$ to $t_k^i - 1$. Since $\tau_k^i < \eta_k^i = t_k^{i+1} - t_k^i$, $\mathbf{e}_{t_k^i}$ and $\mathbf{e}_{t_k^j}$ do not contain common noise terms when $i \neq j$. Taking (B.1.6) into (B.1.4), after some simple simplifications, $\mathbf{x}_k$ can be simplified as

below with v-segment summations

$$\begin{aligned}\mathbf{x}_k = &\sum_{i=1}^{\eta_k^0+\tau_k^1} \mathbf{A}^{i-1}\mathbf{w}_{k-i} + (\mathbf{A}+\mathbf{BK})^{\eta_k^0} \sum_{i=\tau_k^1+1}^{\eta_k^1+\tau_k^2} \mathbf{A}^{i-1}\mathbf{w}_{t_k^1-i} \\ &+ \cdots + (\mathbf{A}+\mathbf{BK})^{\eta_k^0+\cdots+\eta_k^{v-2}} \sum_{i=\tau_k^{v-1}+1}^{\eta_k^{v-1}+\tau_k^v} \mathbf{A}^{i-1}\mathbf{w}_{t_k^{v-1}-i},\end{aligned} \tag{B.1.9}$$

where $\eta_k^j > \tau_k^j, \forall j = 1, \cdots, v-1$.

For case 2), the estimation-error terms $\mathbf{e}_{t_k^i}$ and $\mathbf{e}_{t_k^j}$ in (B.1.4) may contain common noise terms when $i \neq j$, and $\eta_k^j > \tau_k^j$ may not hold for $j = 1, \cdots, v-1$. Inspired by the result (B.1.9) in the first case, to calculate $\mathbf{x}_k$, we divide the time range $\mathcal{S}$ by the time slots $t_k^j - \tau_k^j$, $j = 1, \cdots, v-1$. Since $t_k^{j'} - \tau_k^{j'}$ may equal to $t_k^j - \tau_k^j$ when $j' \neq j$, $\mathcal{S}$ is divided into v' segments from left to right, and $1 \leq v' \leq v$.

To investigate the noise terms within the first $v'-1$ segments of $\mathcal{S}$, we assume that sensor's successful transmissions occurred in the time ranges $\left[t_k^{j'+1}+1, t_k^{j'}\right]$ and $\left[t_k^{j+1}+1, t_k^j\right]$ and there is no sensor's successful transmission in the gap between them, where $v \geq j' > j \geq 1$. Thus, $\eta_k^{j'} > \tau_k^{j'}$ and $\eta_k^j > \tau_k^j$. When $j' = j+1$, we have $t_k^{j'} - \tau_k^{j'} = t_k^{j+1} - \tau_k^{j+1} = t_k^j - \eta_k^j - \tau_k^{j+1}$ and only $\mathbf{w}'$ and the estimation-error term $\mathbf{e}_{t_k^{j'}}$ contains the noise terms within the time segment $\left[t_k^{j'} - \tau_k^{j'}, t_k^j - \tau_k^j - 1\right] = \left[t_k^j - (\eta_k^j + \tau_k^{j+1}), t_k^j - \tau_k^j - 1\right]$, therefore, the noise terms in this segment have exactly the same expressions as in (B.1.9) of case 1), i.e.,

$$(\mathbf{A}+\mathbf{BK})^{\eta_k^0+\cdots+\eta_k^{j-1}} \sum_{i=\tau_k^j+1}^{\eta_k^j+\tau_k^{j+1}} \mathbf{A}^{i-1}\mathbf{w}_{t_k^j-i}. \tag{B.1.10}$$

When $j' > j+1$, $\mathbf{w}'$ and the estimation-error terms $\mathbf{e}_{t_k^{j+1}}, \mathbf{e}_{t_k^{j+2}}, \cdots, \mathbf{e}_{t_k^{j'}}$ contains the noise terms within the time segment $\left[t_k^{j'} - \tau_k^{j'}, t_k^j - \tau_k^j - 1\right] = \left[t_k^j - (\eta_k^j + \tau_k^{j+1}), t_k^j - \tau_k^j - 1\right]$. After combining the noise terms in this range, we also have the expression (B.1.10).

To investigate the noise terms of the v'th (last) segment of $\mathcal{S}$, we assume that

the most recently successful sensor's transmission before t_k^1 is within the range of $\left[t_k^{j+1}+1, t_k^j\right]$, where $j \in \{1, \cdots, v\}$. We see that $\mathbf{w}'$ and the estimation-error terms $\mathbf{e}_{t_k^1}, \cdots, \mathbf{e}_{t_k^j}$ contains the noise terms within the time range $\left[t_k^j - \tau_k^j, k-1\right] = [k-(\eta_k^0+\tau_k^1), k-1]$. After combining the noise terms in this range contributed by $\mathbf{e}_{t_k^1}, \cdots, \mathbf{e}_{t_k^j}$ and $\mathbf{w}'$, we have exactly the same expressions as in (B.1.9) of case 1), i.e.,

$$\sum_{i=1}^{\eta_k^0+\tau_k^1} \mathbf{A}^{i-1}\mathbf{w}_{k-i}. \tag{B.1.11}$$

To sum up, different from (B.1.9) of case 1), $\mathbf{x}_k$ of case 2) has v' segment summations, i.e.,

$$\begin{aligned}\mathbf{x}_k = &\sum_{i=1}^{\eta_k^0+\tau_k^1} \mathbf{A}^{i-1}\mathbf{w}_{k-i} + \mathbb{1}(\eta_k^1 > \tau_k^1)(\mathbf{A}+\mathbf{BK})^{\eta_k^0} \sum_{i=\tau_k^1+1}^{\eta_k^1+\tau_k^2} \mathbf{A}^{i-1}\mathbf{w}_{t_k^1-i} \\ &+ \cdots + \\ &\mathbb{1}(\eta_k^{v-1} > \tau_k^{v-1})(\mathbf{A}+\mathbf{BK})^{\eta_k^0+\cdots+\eta_k^{v-2}} \sum_{i=\tau_k^{v-1}+1}^{\eta_k^{v-1}+\tau_k^v} \mathbf{A}^{i-1}\mathbf{w}_{t_k^{v-1}-i},\end{aligned} \tag{B.1.12}$$

where $\mathbb{1}(\cdot)$ is the indicator function and $\sum_{j=1}^{v-1} \mathbb{1}(\eta_k^j > \tau_k^j) = v'-1$.

For case 3), the range $\mathcal{S}$ has only one segment, which is a special case of case 2) discussed above (B.1.11), where $j = v$. Therefore, $\mathbf{x}_k$ has the expression of (B.1.11).

Therefore, the general expression of $\mathbf{x}_k$ is given in (B.1.12), and thus the state covariance $\mathbf{P}_k = \mathbb{E}[\mathbf{x}_k\mathbf{x}_k^\top]$ is obtained as (3.3.21).

B.2 Proof of Theorem 3.1

We prove that the stationary and deterministic policy π' in (3.4.5) stabilizes the plant.

It is easy to verify that the state-transition process induced by π' is an ergodic

Markov process, i.e., any state in $\mathbb{S}$ is aperiodic and positive recurrent. In the following, we prove that the average cost of the plant induced by π' is bounded.

From (3.4.5), (3.3.3) and (3.3.12), we see that the policy π' is actually a *persistent scheduling policy*, which consecutively schedules the uplink transmission until a transmission is successful and then consecutively schedules the downlink transmission until a transmission is successful, and so on. The transmission process of (s)ensor's measurement and (c)ontroller's command is illustrated as

$$\{\cdots, \overbrace{\underbrace{\mathrm{s}\cdots\mathrm{s}}_{m'},\underbrace{\mathrm{c}\cdots\mathrm{c}}_{n'}}^{\text{control cycle } (t-1)}, \overbrace{\underbrace{\mathrm{s}\cdots\mathrm{s}}_{m},\underbrace{\mathrm{c}\cdots\mathrm{c}}_{n}}^{\text{control cycle } t}, \cdots\} \tag{B.2.1}$$

where m and n are the numbers of consecutively scheduled uplink and downlink transmission, respectively.

For the ease of analysis, we define the concept of *control cycle*, which consists of M consecutive uplink transmissions and the following N consecutive downlink transmissions. It is clear that M and N follow geometric distributions with success probabilities $(1-p_s)$ and $(1-p_c)$, respectively. The values of M and N change in different control cycles independently as illustrated in (B.2.1). Thus, the uplink-downlink schedule process (B.2.1) can be treated as a sequence of control cycles.

Let S and $L \triangleq M+N$ denote the sum cost of the plant and the number of transmissions in a control cycle, respectively. We can prove that S and L of the sequence of control cycles can be treated as ergodic Markov chains, i.e., $\{\cdots, S_t, S_{t+1}, \cdots\}$ and $\{\cdots, L_t, L_{t+1}, \cdots\}$, where t is the control-cycle index. We use N' to denote the number of consecutive downlink transmissions before the current control cycle, which follows the same distribution of N. Due to the ergodicity of $\{S_t\}$ and $\{L_t\}$, the

average cost in (3.2.2) can be rewritten as

$$J = \lim_{t\to\infty} \frac{S_1 + S_2 + \cdots + S_t}{L_1 + L_2 + \cdots + L_t} = \frac{\mathbb{E}[S]}{\mathbb{E}[L]}, \tag{B.2.2}$$

where

$$\mathbb{E}[S] = \sum_{n'=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \mathbb{E}[S|N'=n', M=m, N=n] \tag{B.2.3}$$
$$\mathbb{P}[N'=n', M=m, N=n],$$

$$\mathbb{E}[L] = \sum_{n'=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} (m+n)\,\mathbb{P}[N'=n', M=m, N=n]. \tag{B.2.4}$$

Thus, the average cost J is bounded if $\mathbb{E}[S]$ is. From the policy (3.4.5) and the state-transition rules in (3.3.3) and (3.3.12), we see that ϕ is equal to $N'+1$ at the beginning of the control cycle, and increases one-by-one within the control cycle, and we have

$$\mathbb{E}[S|N'=n', M=m, N=n] = \sum_{i=1}^{m+n} c(n'+i), \tag{B.2.5}$$

and

$$\begin{aligned}\mathbb{P}[N'=n', M=m, N=n] &= \mathbb{P}[N'=n']\mathbb{P}[M=m]\mathbb{P}[N=n] \\ &= (1-p_c)p_c^{n'-1}(1-p_s)p_s^{m-1}(1-p_c)p_c^{n-1},\end{aligned} \tag{B.2.6}$$

as N', M, N are independent with each other. Let $p_0 \triangleq \max\{p_s, p_c\}$. We have

$$\mathbb{E}[S] \le \kappa \sum_{n'=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{i=1}^{m+n} c(n'+i)p_0^{n'+m+n} \tag{B.2.7}$$

$$< \kappa \sum_{n'=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(n'+m+n)c(n'+m+n)p_0^{n'+m+n} \tag{B.2.8}$$

$$< \kappa \sum_{i=1}^{\infty} i^4 c(i) p_0^i. \tag{B.2.9}$$

where $\kappa = (1-p_c)p_c^{-1}(1-p_s)p_s^{-1}(1-p_c)p_c^{-1}$, and (B.2.9) is due to the fact that the number of possible partition of $(n'+m+n)$ into three parts is less than $(n'+m+n)^3$. Since there always exists $p_0' > p_0$ and n such that $i^4 p_0^i < (p_0')^i, \forall i > n$, $\sum_{i=1}^{\infty} i^4 c(i) p_0^i < \infty$ if $\sum_{i=1}^{\infty} c(i)(p_0')^i < \infty$. Using the result that $\sum_{j=1}^{\infty}(p_0')^j c(j) < \infty$ iff $p_0'\rho^2(\mathbf{A}) < 1$

in [12] and [63], $\sum_{i=1}^{\infty} i^4 c(i) p_0^i < \infty$ if $p_0 \rho^2(\mathbf{A}) < 1$, completing the proof.

B.3 Proof of Theorem 3.2

The necessity and sufficiency are proved as follows.

B.3.1 Sufficiency

Similar to the proof of Theorem 3.1, we need to define the control cycle of the naive policy and then calculate the average cost.

Different from the Proof of Theorem 3.1, the control cycle is defined as the time slots after a effective control cycle until the end of the following effective control cycle. Here, the effective control cycle is the sequence of time slots starting from a sensor's successful transmission and ending at a controller's successful transmission, where there is no successful transmissions in between. In other words, in an effective control cycle, the sensor's measurement at the beginning of the cycle will be utilized for generating a control command, which will be implemented on the plant by the end of the cycle. The control cycle and the effective control cycle are illustrated as

$$\{\cdots, \overbrace{\underbrace{\text{sčšc}\cdots\text{sc}}_{m'},\ \underbrace{\text{šcsc}\cdots\text{sč}}_{n'},}^{\text{control cycle } (t-1)}\ \overbrace{\underbrace{\text{scsč}\cdots\text{sc}}_{m},\ \underbrace{\text{šcsc}\cdots\text{sč}}_{n},}^{\text{control cycle } t}\ \cdots\} \tag{B.3.1}$$

$$\text{effective control cycle } (t-1) \qquad \text{effective control cycle } t$$

where n and $l = m + n$ are the number of time slots of an effective control cycle and a control cycle, respectively, and š and č denotes a successful sensor's transmission and controller's transmission, respectively. Note that m and n are even numbers.

Similar to the proof of Theorem 3.1, S and $L \triangleq M + N$ denote the sum cost of the plant and the number of transmissions in a control cycle, respectively. Also, S

and L of the sequence of control cycles can be treated as ergodic Markov chains, i.e., $\{\cdots, S_t, S_{t+1}, \cdots\}$ and $\{\cdots, L_t, L_{t+1}, \cdots\}$, where t is the control-cycle index. Due to the ergodicity of $\{S_t\}$ and $\{L_t\}$, the average cost in (3.2.2) can be rewritten as (B.2.2), where

$$\mathbb{E}[S] = \sum_{\frac{n'}{2}=1}^{\infty} \sum_{\frac{m}{2}=0}^{\infty} \sum_{\frac{n}{2}=1}^{\infty} \mathbb{E}[S|N'=n', M=m, N=n] \tag{B.3.2}$$
$$\mathbb{P}[N'=n', M=m, N=n],$$

$$\mathbb{E}[L] = \sum_{\frac{n'}{2}=1}^{\infty} \sum_{\frac{m}{2}=0}^{\infty} \sum_{\frac{n}{2}=1}^{\infty} l\mathbb{P}[N'=n', M=m, N=n], \tag{B.3.3}$$

where $M+N$ and N are the length of the current control cycle and effective control cycle, respectively, and N' is the length of the previous effective control cycle. It is clear that N' is independent with M and N.

Thus, the average cost J is bounded if $\mathbb{E}[S]$ is. From the naive policy and the definition of the control cycle and the effective control cycle, we can derive the following probability density functions as

$$\begin{aligned}
&\mathbb{P}[M=m, N=n] \\
&= \begin{cases}
(1-p_s)(1-p_c)p_c^{n/2-1}p_s^{n/2-1}, & m=0, n=2,4,6 \\
(1-p_s)\left(p_s^{m/2} + (1-p_s)\sum_{i=1}^{m/2} p_s^{i-1}p_c^{m/2+1-i}\right) & \\
\qquad \times(1-p_c)p_c^{n/2-1}p_s^{n/2-1}, & m,n=2,4,6,\cdots
\end{cases}
\end{aligned} \tag{B.3.4}$$

and thus

$$\begin{aligned}
\mathbb{P}[N'=n'] = \mathbb{P}[N=n'] &= \sum_{m=0}^{\infty} \mathbb{P}[M=m, N=n'] \\
&= (1-p_sp_c)p_c^{\frac{n'}{2}-1}p_s^{\frac{n'}{2}-1}, \ n'=2,4,6,\cdots
\end{aligned} \tag{B.3.5}$$

Then, it can be proved that

$$\mathbb{P}[N' = n'] \leq \kappa_1 p_0^{n'}, \ \forall n' = 2, 4, 6, \cdots$$

$$\mathbb{P}[M = m, N = n] \leq \kappa_2 (1 + m/2) p_0^{m/2} p_0^n, \forall m = 0, 2, 4, \cdots, \quad n = 2, 4, 6, \cdots \tag{B.3.6}$$

where $p_0 = \max\{p_s, p_c\}$, $\kappa_1 = (1 - p_s p_c) p_c^{-1} p_s^{-1}$, and $\kappa_2 = (1 - p_s)(1 - p_c) p_c^{-1} p_s^{-1}$.

Since $\mathbb{E}[S|N' = n', M = m, N = n] = \sum_{i=1}^{m+n} c(n' + i)$, we have

$$\mathbb{E}[S] = \sum_{\frac{n'}{2}=1}^{\infty} \sum_{\frac{m}{2}=0}^{\infty} \sum_{\frac{n}{2}=1}^{\infty} \mathbb{E}[S|N' = n', M = m, N = n]$$

$$\mathbb{P}[N' = n'] \mathbb{P}[M = m, N = n], \tag{B.3.7}$$

$$\leq \kappa_1 \kappa_2 \sum_{\frac{n'}{2}=1}^{\infty} \sum_{\frac{m}{2}=0}^{\infty} \sum_{\frac{n}{2}=1}^{\infty} \sum_{i=1}^{m+n} c(n' + i)(1 + \frac{m}{2}) p_0^{n' + \frac{m}{2} + n} \tag{B.3.8}$$

$$< \kappa_1 \kappa_2 \sum_{\frac{n'}{2}=1}^{\infty} \sum_{\frac{m}{2}=0}^{\infty} \sum_{\frac{n}{2}=1}^{\infty} (n' + m + n)^2 c(n' + m + n) p_0^{\frac{n' + m + n}{2}} \tag{B.3.9}$$

$$< 4 \kappa_1 \kappa_2 \sum_{i=2}^{\infty} i^5 c(2i) p_0^i. \tag{B.3.10}$$

Since there always exists $p_0' > p_0$ and $\bar{n}$ such that $i^5 p_0^i < (p_0')^i, \forall i > \bar{n}$, $\sum_{i=2}^{\infty} i^5 c(2i) p_0^i < \infty$ if $\sum_{i=2}^{\infty} c(2i)(p_0')^i < \infty$. Also, we have $\sum_{i=2}^{\infty} c(2i)(p_0')^i < \sum_{i=1}^{\infty} c(i) \sqrt{p_0'}^i$. Using the result that $\sum_{j=1}^{\infty} \sqrt{p_0'}^j c(j) < \infty$ iff $\sqrt{p_0'} \rho^2(\mathbf{A}) < 1$ in [12] and [63], $\sum_{i=2}^{\infty} i^5 c(2i) p_0^i < \infty$ if $\sqrt{p_0} \rho^2(\mathbf{A}) < 1$, completing the proof of sufficiency.

B.3.2 Necessity

To prove the necessity, we consider two ideal cases: the sensor's transmission is perfect, i.e., $p_s = 0$, and the controller's transmission is perfect, i.e., $p_c = 0$. In these cases, the stability conditions are the necessary condition that the plant can be stabilized by the naive policy. The proof requires the analysis of average cost

of control cycles which follows similar steps in the proof of sufficiency. Since the average cost $J = \frac{\mathbb{E}[S]}{\mathbb{E}[L]}$ and $\mathbb{E}[L]$ is bounded straightforwardly, we only need to prove the necessary condition that the average sum cost of a control cycle is bounded, i.e., $\mathbb{E}[S]$.

In the ideal cases, we have

$$\mathbb{E}[S] = \begin{cases} (1-p_c)\sum_{j=1}^{\infty}\sum_{i=2}^{2j+1} c(i)p_c^{j-1}, & p_s = 0 \\ (1-p_s)\sum_{j=1}^{\infty}\sum_{i=2}^{2j+1} c(i)p_s^{j-1}, & p_c = 0 \end{cases} \tag{B.3.11}$$

Therefore, if $\mathbb{E}[S]$ is bounded, we have

$$\begin{aligned} &\sum_{i=1}^{\infty} c(2i)p_c^i < \infty, \sum_{i=1}^{\infty} c(2i+1)p_c^i < \infty \\ &\sum_{i=1}^{\infty} c(2i)p_s^i < \infty, \sum_{i=1}^{\infty} c(2i+1)p_s^i < \infty \end{aligned} \tag{B.3.12}$$

and hence

$$\sum_{i=1}^{\infty} c(i)\sqrt{p_c}^i < \infty, \sum_{i=1}^{\infty} c(i)\sqrt{p_s}^i < \infty. \tag{B.3.13}$$

Using the result that $\sum_{j=1}^{\infty}\sqrt{p_0}^j c(j) < \infty$ iff $\sqrt{p_0}\rho^2(\mathbf{A}) < 1$ in [12] and [63], the necessary condition that the average cost inducted by the naive policy is bounded, is $\sqrt{p_s}\rho^2(\mathbf{A}) < 1$ and $\sqrt{p_c}\rho^2(\mathbf{A}) < 1$, completing the proof of necessity.

B.4 Proof of Theorem 3.3

The necessity and sufficiency are proved as follows.

B.4.1 Sufficiency

We construct a persistent-scheduling-like policy including three phases: 1) the sensor's transmission is consecutively scheduled until it is successful, and then 2) the

controller's transmission is consecutively scheduled until a successful transmission, and then 3) none of the sensor nor the controller is scheduled for transmission in the following $v-1$ time slots, i.e., all the commands contained in the successfully received control packet will be implemented by the actuator, and then phase 1) and so on.

Then, following the similar steps of the proof of Theorem 3.1, it can be proved that the persistent-scheduling-like policy stabilizes the plant if (3.4.4) holds.

B.4.2 Necessity

The proof is conducted by considering two virtual cases: 1) the sensor's transmission is continuously scheduled, while there is a virtual control input $\mathbf{u}_k$ at each time slot that ideally resets $\mathbf{x}_k$ to $\mathbf{0}$ if the sensor's transmission is successful at k, and is $\mathbf{0}$ otherwise; 2) the controller's transmission is continuously scheduled, while the controller applies a virtual estimator that has perfect estimation of the plant states in each time slots.

It can be readily proved that the two virtual cases result in lower average costs than any feasible uplink-downlink scheduling policy. Then, following the similar steps in the proof of Theorem 3.1 and 3.2, it can be shown that if the average cost of case 1) is bounded, $p_s < 1/\rho^2(\mathbf{A})$ must be satisfied, and if the average cost of case 2) is bounded, $p_c < 1/\rho^2(\mathbf{A})$ must be satisfied, completing the proof of necessity.

Bibliography

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5g," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.

[2] W. Liu, X. Zhou, S. Durrani, and P. Popovski, "A novel receiver design with joint coherent and non-coherent processing," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3479–3493, 2017.

[3] Y. Wang, W. Liu, X. Zhou, and G. Liu, "On the performance of splitting receiver with joint coherent and non-coherent processing," *IEEE Transactions on Signal Processing*, vol. 68, pp. 917–930, 2020.

[4] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[5] X. Yu and Y. Xue, "Smart grids: A cyber–physical systems perspective," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 1058–1070, May. 2016.

[6] A. Aswani, N. Master, J. Taneja, D. Culler, and C. Tomlin, "Reducing transient and steady state electricity consumption in hvac using learning-based model-predictive control," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 240–253, Jan. 2012.

[7] A. Willig, "Recent and emerging topics in wireless industrial communications: A selection," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 2, pp. 102–124, May. 2008.

[8] ISA100. [Online]. Available: https://www.isa.org/isa100

[9] WirelessHART. [Online]. Available: https://en.hartcomm.org

[10] Wireless Industrial Networking Alliance. [Online]. Available: https://www.wina.org

[11] ZigBee Alliance. [Online]. Available: https://www.zigbee.org

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

[13] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[14] J. Humpherys, P. Redd, and J. West, "A fresh look at the kalman filter," *SIAM review*, vol. 54, no. 4, pp. 801–823, 2012.

[15] A. Giovanidis, G. Wunder, and J. Bühler, "Optimal control of a single queue with retransmissions: Delay-dropping tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3736–3746, Jul. 2009.

[16] R. Dorf and R. Bishop, "Modern control systems," *Pearson Education*, 2008.

[17] J. O'Reilly, "The discrete linear time invariant time-optimal control problem — an overview," *Automatica*, vol. 17, no. 2, pp. 363–370, 1981.

[18] M. Athans, "The role and use of the stochastic linear-quadratic-gaussian problem in control system design," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 529–552, Dec. 1971.

[19] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, no. 7, pp. 733–764, 2003.

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

[21] S. Deshmukh, B. Natarajan, and A. Pahwa, "State estimation over a lossy network in spatially distributed cyber-physical systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3911–3923, Aug. 2014.

[22] W. Liu, D. E. Quevedo, Y. Li, K. H. Johansson, and B. Vucetic, "Remote state estimation with smart sensors over markov fading channels," *arXiv preprint arXiv:2005.07871*, 2020.

[23] W. Liu, P. Popovski, Y. Li, and B. Vucetic, "Wireless networked control systems with coding-free data transmission for industrial iot," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1788–1801, Mar. 2020.

[24] W. Liu, G. Nair, Y. Li, D. Nesic, B. Vucetic, and H. V. Poor, "On the latency, rate and reliability tradeoff in wireless networked control systems for iiot," *IEEE Internet of Things Journal*, vol. 8, no. 2, pp. 723–733, Jan. 2021.

[25] W. Liu, D. E. Quevedo, K. H. Johansson, B. Vucetic, and Y. Li, "Remote state estimation of multiple systems over multiple markov fading channels," *arXiv preprint arXiv:2104.04181*, 2021.

[26] A. Cervin, D. Henriksson, B. Lincoln, J. Eker, and K.-E. Arzen, "How does control timing affect performance? analysis and simulation of timing using

jitterbug and truetime," *IEEE Control Systems Magazine*, vol. 23, no. 3, pp. 16–30, Jun. 2003.

[27] F.-L. Lian, J. Moyne, and D. Tilbury, "Network design consideration for distributed control systems," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 2, pp. 297–307, Mar. 2002.

[28] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 1041–1046, Apr. 2013.

[29] W. Heemels, K. H. Johansson, and P. Tabuada, "An introduction to event-triggered and self-triggered control," in *Proc. IEEE Conference on Decision and Control (CDC)*, Dec. 2012.

[30] C. Peng, D. Yue, and M.-R. Fei, "A higher energy-efficient sampling scheme for networked control systems over ieee 802.15. 4 wireless networks," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1766–1774, Oct. 2016.

[31] W. M. H. Heemels, A. R. Teel, N. Van de Wouw, and D. Nešić, "Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance," *IEEE Transactions on Automatic control*, vol. 55, no. 8, pp. 1781–1796, Aug. 2010.

[32] K. Gatsis, H. Hassani, and G. J. Pappas, "Latency-reliability tradeoffs for state estimation," *IEEE Transactions on Automatic Control*, Mar. 2021.

[33] P. Park, C. Fischione, A. Bonivento, K. H. Johansson, and A. Sangiovanni-Vincent, "Breath: An adaptive protocol for industrial control applications using wireless sensor networks," *IEEE Transactions on Mobile Computing*, Jun. 2011.

[34] J.-H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 4, pp. 609–619, Aug. 2004.

[35] J. Ploennigs, V. Vasyutynskyy, and K. Kabitzsch, "Comparative study of energy-efficient sampling approaches for wireless control networks," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 3, pp. 416–424, Aug. 2010.

[36] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Optimal power management in wireless control systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1495–1510, Jun. 2014.

[37] C. E. Garcia, D. M. Prett, and M. Morar, "Model predictive control: Theory and practice: A survey," *Automatica*, vol. 25, no. 3, pp. 335–348, 1989.

[38] E. Henriksson, D. E. Quevedo, E. G. Peters, H. Sandberg, and K. H. Johansson, "Multiple-loop self-triggered model predictive control for network scheduling and control," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 6, pp. 2167–2181, Nov. 2015.

[39] Z.-H. Pang, G.-P. Liu, D. Zhou, and D. Sun, "Data-based predictive control for networked nonlinear systems with network induced-delay and packet dropout," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1249–1257, Feb. 2016.

[40] D. E. Quevedo, E. I. Silva, and G. C. Goodwin, "Packetized predictive control over erasure channels," in *Proc. American Control Conference*, Jul. 2007.

[41] E. G. Peters, D. E. Quevedo, and J. Østergaard, "Shaped gaussian dictionaries for quantized networked control systems with correlated dropouts," *IEEE Transactions on Signal Processing*, vol. 64, no. 1, pp. 203–213, Jan. 2016.

[42] V. Gupta and F. Luo, "On a control algorithm for time-varying processor availability," *IEEE Transactions on Automatic Control*, vol. 3, no. 58, pp. 743–748, Mar. 2013.

[43] W. Liu, D. E. Quevedo, Y. Li, and B. Vucetic, "Anytime control with markovian computation and communication resources," *arXiv preprint arXiv:2012.00962*, 2020.

[44] A. S. Matveev and A. V. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 670–676, Apr. 2003.

[45] S. Weerakkody, Y. Mo, B. Sinopoli, D. Han, and L. Shi, "Multi sensor scheduling for state estimation with event based, stochastic triggers," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2695–2701, Sep. 2016.

[46] S. Wu, X. Ren, S. Dey, and L. Shi, "Optimal scheduling of multiple sensors over shared channels with packet transmission constraint," *Automatica*, vol. 96, pp. 22–31, 2018.

[47] A. S. Leong, S. Dey, and D. E. Quevedo, "Sensor scheduling in variance based event triggered estimation with packet drops," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1880–1895, Apr. 2017.

[48] C. Yang, J. Zheng, X. Ren, W. Yang, H. Shi, and L. Shi, "Multi-sensor kalman filtering with intermittent measurements," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 797–804, Mar. 2017.

[49] L. Shi and H. Zhang, "Scheduling two gauss–markov systems: An optimal solution for remote state estimation under bandwidth constraint," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2038–2042, Apr. 2012.

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

[51] M. H. Mamduhi, M. Kneissl, and S. Hirche, "Decentralized event-triggered medium access control for networked control systems," in *Proc. IEEE Conference on Decision and Control (CDC)*, Dec. 2016.

[52] M. H. Mamduhi, M. Vilgelm, W. Kellerer, and S. Hirche, "Prioritized contention resolution for random access networked control systems," in *Proc. IEEE Conference on Decision and Control (CDC)*, Dec. 2017.

[53] D. Han, J. Wu, Y. Mo, and L. Xie, "On stochastic sensor network scheduling for multiple processes," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6633–6640, Dec. 2017.

[54] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Random access design for wireless control systems," *Automatica*, vol. 91, pp. 1–9, 2018.

[55] W. Liu, X. Zhou, S. Durrani, H. Mehrpouyan, and S. D. Blostein, "Energy harvesting wireless sensor networks: Delay analysis considering energy costs of sensing and transmission," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4635–4650, Jul. 2016.

[56] B. Demirel, A. Aytekin, D. E. Quevedo, and M. Johansson, "To wait or to drop: On the optimal number of retransmissions in wireless control," in *Proc. European Control Conference (ECC)*, Jul. 2015.

[57] V. Gupta, "On estimation across analog erasure links with and without acknowledgements," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2896–2901, Dec. 2010.

[58] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.

[59] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[60] E. Lähetkangas, K. Pajukoski, J. Vihriälä, G. Berardinelli, M. Lauridsen, E. Tirola, and P. Mogensen, "Achieving low latency and energy consumption by 5G TDD mode optimization," in *Proc. International Conference on Communications Workshops (ICC)*, Jun. 2014.

[61] K. Antonakoglou, X. Xu, E. Steinbach, T. Mahmoodi, and M. Dohler, "Towards haptic communications over the 5G tactile Internet," *IEEE Communications Surveys & Tutorials*, Fourthquarter 2018.

[62] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Mar. 2012.

[63] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.

[64] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017.

[65] L. Shi and L. Xie, "Optimal sensor power scheduling for state estimation of Gauss-Markov systems over a packet-dropping network," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2701–2705, May. 2012.

[66] C. Yang, J. Wu, W. Zhang, and L. Shi, "Schedule communication for decentralized state estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 10, pp. 2525–2535, May. 2013.

[67] P. S. Maybeck, *Stochastic Models, Estimation, and Control*. Academic Press, 1982, vol. 3.

[68] M. B. Rhudy and Y. Gu, "Online stochastic convergence analysis of the kalman filter," *International Journal of Stochastic Analysis*, 2013.

[69] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

[70] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 1, pp. 163–171, Feb. 1995.

[71] J. Razavilar, K. R. Liu, and S. I. Marcus, "Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 484–494, Mar. 2002.

[72] R. Ford, S. Rangan, E. Mellios, D. Kong, and A. Nix, "Markov channel-based performance analysis for millimeter wave mobile networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2019.

[73] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for mimo channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, Oct. 2004.

[74] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May. 2010.

[75] Y. Li, M. C. Gursoy, and S. Velipasalar, "Throughput of harq-ir with finite blocklength codes and qos constraints," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017.

[76] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

[77] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid arq under a resource constraint," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1900–1913, Mar. 2019.

[78] *Markov Decision Processes (MDP) Toolbox for MATLAB*, 2015. [Online]. Available: https://au.mathworks.com/matlabcentral/fileexchange/25786-markov-decision-processes-mdp-toolbox

[79] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," in *Proc. Association for Uncertainty in AI (AUAI)*, 1995.

[80] P. Park, S. C. Ergen, C. Fischione, C. Lu, and K. H. Johansson, "Wireless network design for control systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 978–1013, Second Quarter, 2018.

[81] C. Perera, C. H. Liu, and S. Jayawardena, "The emerging Internet of Things marketplace from an industrial perspective: A survey," *IEEE Transactions on Emerging Topics in Computing*, vol. 3, no. 4, pp. 585–598, Jan. 2015.

[82] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the Internet of Things and Industry 4.0," *IEEE Industrial Electronics Magazine*, vol. 11, no. 1, pp. 17–27, Mar. 2017.

[83] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel, A. Puschmann, A. Mitschele-Thiel, M. Muller, T. Elste, and M. Windisch, "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, Feb. 2017.

[84] O. Bello and S. Zeadally, "Intelligent device-to-device communication in the Internet of Things," *IEEE Systems Journal*, vol. 10, no. 3, pp. 1172–1182, Jan. 2014.

[85] D. Zhang, P. Shi, Q.-G. Wang, and L. Yu, "Analysis and synthesis of networked control systems: A survey of recent advances and challenges," *ISA Transactions*, vol. 66, pp. 376–392, Jan. 2017.

[86] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, Jun. 2016.

[87] L. Lyu, C. Chen, S. Zhu, N. Cheng, B. Yang, and X. Guan, "Control performance aware cooperative transmission in multiloop wireless control systems for industrial IoT applications," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3954–3966, Sep. 2018.

[88] P. Gil, A. Santos, and A. Cardoso, "Dealing with outliers in wireless sensor networks: An oil refinery application," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1589–1596, Nov. 2013.

[89] Y. Wang, S. X. Ding, D. Xu, and B. Shen, "An h∞ fault estimation scheme of wireless networked control systems for industrial real-time applications," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 6, pp. 2073–2086, Apr. 2014.

[90] V. Liberatore and A. Al-Hammouri, "Smart grid communication and co-simulation," in *Proc. IEEE EnergyTech*, May. 2011.

[91] R. Hult, G. R. Campos, E. Steinmetz, L. Hammarstrand, P. Falcone, and H. Wymeersch, "Coordination of cooperative autonomous vehicles: Toward safer and more efficient road transportation," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 74–84, Nov. 2016.

[92] C. Yang, J. Wu, X. Ren, W. Yang, H. Shi, and L. Shi, "Deterministic sensor selection for centralized state estimation under limited communication resource," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2336–2348, May. 2015.

[93] G.-P. Liu, Y. Xia, J. Chen, D. Rees, and W. Hu, "Networked predictive control of systems with random network delays in both forward and feedback channels," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 3, pp. 1282–1297, Jun. 2007.

[94] B. Demirel, V. Gupta, D. E. Quevedo, and M. Johansson, "On the trade-off between communication and control cost in event-triggered dead-beat control," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2973–2980, Jun. 2017.

[95] P. K. Mishra, D. Chatterjee, and D. E. Quevedo, "Stabilizing stochastic predictive control under bernoulli dropouts," *IEEE Transactions on Automatic Control*, vol. 63, no. 6, pp. 1579–1590, Jun. 2018.

[96] Z. Shen, A. Khoryaev, E. Eriksson, and X. Pan, "Dynamic uplink-downlink configuration and interference management in td-lte," *IEEE Communications Magazine*, vol. 50, no. 11, pp. 51–59, Nov. 2012.

[97] F. Boccardi, J. Andrews, H. Elshaer, M. Dohler, S. Parkvall, P. Popovski, and S. Singh, "Why to decouple the uplink and downlink in cellular networks and how to do it," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 110–117, Mar. 2016.

[98] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[99] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels - a survey of principles and applications," *IEEE Signal Processing Magazine*, vol. 25, no. 5, pp. 57–80, Sep. 2008.

[100] V. Gupta, B. Sinopoli, S. Adlakha, A. Goldsmith, and R. Murray, "Receding horizon networked control," in *Proc. Allerton Conference on Commununications, Control & Computing*, 2006.

[101] *Grain Storage*, Grain research and development corporation, 2017. [Online]. Available: https://grdc.com.au/

[102] K. Huang, W. Liu, Y. Li, B. Vucetic, and A. Savkin, "Optimal downlink-uplink scheduling of wireless networked control for Industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1756–1772, Mar. 2020.

[103] F. Mager, D. Baumann, R. Jacob, L. Thiele, S. Trimpe, and M. Zimmerling, "Feedback control goes wireless: Guaranteed stability over low-power multi-hop networks," in *Proc. IEEE International Conference on Cyber-Physic Systems (ICCPS)*, Apr. 2019.

[104] A. Aijaz, A. Stanoev, and U. Raza, "Gallop: Toward high-performance connectivity for closing control loops over multi-hop wireless networks," in *Proc.*

International Conference on Real-Time Networks and Systems (RTNS), Nov. 2019.

[105] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid ARQ in wireless networked control," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3490–3504, May. 2020.

[106] K. Huang, W. Liu, Y. Li, and B. Vucetic, "To retransmit or not: Real-time remote estimation in wireless networked control," in *Proc. IEEE International Conference on Communications (ICC)*, May,2019.

[107] W. Liu, P. Popovski, Y. Li, and B. Vucetic, "Real-time wireless networked control systems with coding-free data transmission," in *Proc. IEEE Conference and Exhibition on Global Telecommunications (GLOBECOM)*, Dec. 2019.

[108] S. Hara, A. Ogino, M. Araki, M. Okada, and N. Morinaga, "Throughput performance of saw-arq protocol with adaptive packet length in mobile packet data transmission," *IEEE Transactions on Vehicular Technology*, vol. 45, no. 3, pp. 561–569, Aug. 1996.

[109] G. Alnwaimi and H. Boujemaa, "Adaptive packet length and mcs using average or instantaneous snr," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10 519–10 527, Nov. 2018.

[110] H. C. Tijms, *A First Course in Stochastic Models*. John Wiley and sons, 2003.

[111] V. Tripathi, E. Visotsky, R. Peterson, and M. Honig, "Reliability-based type II hybrid ARQ schemes," in *Proc. IEEE International Conference on Communications (ICC)*, May. 2003.

[112] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal dos attack scheduling in wireless networked control system," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 3, pp. 843–852, May. 2016.

[113] A. Cetinkaya, H. Ishii, and T. Hayakawa, "Networked control under random and malicious packet losses," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2434–2449, May. 2017.

[114] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons, 2009, vol. 504.

[115] S. H. Weintraub, "Jordan canonical form: Theory and practice," *Synthesis Lectures on Mathematics and Statistics*, vol. 2, no. 1, pp. 1–108, 2009.

[116] J. P. Keener, "The perron-frobenius theorem and the ranking of football teams," *SIAM Review*, vol. 35, pp. 80–93, 1993.