Theses and Dissertations            1. Thesis and Dissertation Collection, all items

2022-06

# TEMPORAL CONNECTIVITY AS A MEASURE OF ROBUSTNESS IN NONORTHOGONAL MULTIPLE ACCESS WIRELESS NETWORKS

Pimentel, Benjamin A.

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/70721

# NAVAL POSTGRADUATE SCHOOL

## MONTEREY, CALIFORNIA

# DISSERTATION

**TEMPORAL CONNECTIVITY AS A MEASURE
OF ROBUSTNESS IN NONORTHOGONAL MULTIPLE
ACCESS WIRELESS NETWORKS**

by

Benjamin A. Pimentel

June 2022

Dissertation Supervisors:                           Alex Bordetsky
                                                    Ralucca Gera

**Approved for public release. Distribution is unlimited.**

THIS PAGE INTENTIONALLY LEFT BLANK

| REPORT DOCUMENTATION PAGE | | *Form Approved OMB No. 0704-0188* |
|---|---|---|

| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE<br>June 2022 | 3. REPORT TYPE AND DATES COVERED<br>Dissertation | |
|---|---|---|---|
| **4. TITLE AND SUBTITLE**<br>TEMPORAL CONNECTIVITY AS A MEASURE OF ROBUSTNESS IN NONORTHOGONAL MULTIPLE ACCESS WIRELESS NETWORKS | | **5. FUNDING NUMBERS**<br>RFPV7; RFLDD | |
| **6. AUTHOR(S)** Benjamin A. Pimentel | | | |
| **7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**<br>Naval Postgraduate School<br>Monterey, CA 93943-5000 | | **8. PERFORMING ORGANIZATION REPORT NUMBER** | |
| **9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**<br>National Security Agency, Fort George G. Meade, MD 20775 | | **10. SPONSORING / MONITORING AGENCY REPORT NUMBER** | |

**11. SUPPLEMENTARY NOTES** The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| 12a. DISTRIBUTION / AVAILABILITY STATEMENT<br>Approved for public release. Distribution is unlimited. | 12b. DISTRIBUTION CODE<br>A |
|---|---|

**13. ABSTRACT (maximum 200 words)**

   Nonorthogonal multiple access (NOMA) is recognized as an important technology to meet the performance requirements of fifth generation (5G) and beyond 5G (B5G) wireless networks. Through the technique of overloading, NOMA has the potential to support higher connection densities, increased spectral efficiency, and lower latency than orthogonal multiple access. The role of NOMA in 5G/B5G wireless networks necessitates a clear understanding of how overloading variability affects network robustness. This dissertation considers the relationship between variable overloading and network robustness through the lens of temporal network theory, where robustness is measured through the evolution of temporal connectivity between network devices (ND). We develop a NOMA temporal graph model and stochastic temporal component framework to characterize time-varying network connectivity as a function of NOMA overloading. The analysis is extended to derive stochastic expressions and probability mass functions for unidirectional connectivity, bidirectional connectivity, the inter-event time between unidirectional connectivity, and the minimum time required for bidirectional connectivity between all NDs. We test the accuracy of our analytical results through numerical simulations. Our results provide an overloading-based characterization of time-varying network robustness that is generalizable to any underlying NOMA implementation.

| 14. SUBJECT TERMS<br>5G, beyond 5G, nonorthogonal multiple access, network science, temporal network theory | 15. NUMBER OF PAGES<br>151 |
|---|---|
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT<br>Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE<br>Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT<br>Unclassified | 20. LIMITATION OF ABSTRACT<br>UU |
|---|---|---|---|

THIS PAGE INTENTIONALLY LEFT BLANK

**TEMPORAL CONNECTIVITY AS A MEASURE OF ROBUSTNESS IN NONORTHOGONAL MULTIPLE ACCESS WIRELESS NETWORKS**

Benjamin A. Pimentel
Major, United States Marine Corps
BS, United States Naval Academy, 2007
MS, Electrical Engineering, Naval Postgraduate School, 2013

Submitted in partial fulfillment of the
requirements for the degree of

**DOCTOR OF PHILOSOPHY IN INFORMATION SCIENCES**

from the

**NAVAL POSTGRADUATE SCHOOL
June 2022**

Approved by:  Alex Bordetsky
Department of
Information Sciences
Dissertation Supervisor
Dissertation Chair

Ralucca Gera
Department of
Applied Mathematics
Dissertation Supervisor

Anthony Canan
Department of
Information Sciences

Carl L. Oros
Department of
Information Sciences

Moe Win
Department of
Aeronautics and Astronautics,
Massachusetts Institute of
Technology

Approved by:  Alex Bordetsky
Chair, Department of Information Sciences

Michael E. Freeman
Vice Provost of Academic Affairs

iii

THIS PAGE INTENTIONALLY LEFT BLANK

# ABSTRACT

Nonorthogonal multiple access (NOMA) is recognized as an important technology to meet the performance requirements of fifth generation (5G) and beyond 5G (B5G) wireless networks. Through the technique of overloading, NOMA has the potential to support higher connection densities, increased spectral efficiency, and lower latency than orthogonal multiple access. The role of NOMA in 5G/B5G wireless networks necessitates a clear understanding of how overloading variability affects network robustness. This dissertation considers the relationship between variable overloading and network robustness through the lens of temporal network theory, where robustness is measured through the evolution of temporal connectivity between network devices (ND). We develop a NOMA temporal graph model and stochastic temporal component framework to characterize time-varying network connectivity as a function of NOMA overloading. The analysis is extended to derive stochastic expressions and probability mass functions for unidirectional connectivity, bidirectional connectivity, the inter-event time between unidirectional connectivity, and the minimum time required for bidirectional connectivity between all NDs. We test the accuracy of our analytical results through numerical simulations. Our results provide an overloading-based characterization of time-varying network robustness that is generalizable to any underlying NOMA implementation.

THIS PAGE INTENTIONALLY LEFT BLANK

# Table of Contents

# List of Figures

# List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

# List of Acronyms and Abbreviations

**1G**          first generation

**2G**          second generation

**3G**          third generation

**3GPP**        3rd Generation Partnership Project

**4G**          fourth generation

**5G**          fifth generation

**6G**          sixth generation

**AWGN**        additive white Gaussian noise

**B5G**         beyond fifth generation (5G)

**BS**          base station

**CDMA**        code division multiple access

**DAC**         digital-to-analog converter

**eMBB**        enhanced mobile broadband

**EMS**         electromagnetic spectrum

**FDMA**        frequency division multiple access

**IEEE**        Institute of Electrical and Electronics Engineers

**IID**         independent and identically distributed

**IMT**         International Mobile Telecommunications

**IoT**         Internet of Things

**IoBT**        Internet of Battlefield Things

**IP**          Internet Protocol

**ISR-T**       intelligence, surveillance, reconnaissance and targeting

**ITU**         International Telecommunications Union

| | |
|---|---|
| **JSD** | Jensen-Shannon divergence |
| **LTE** | Long Term Evolution |
| **MA** | multiple access |
| **MAC** | medium access control |
| **MAD** | maximum absolute deviation |
| **mAD** | mean absolute deviation |
| **MAP** | mobile access point |
| **MBS** | mobile base station |
| **MIMO** | multiple input multiple output |
| **mMIMO** | massive multiple input multiple output (MIMO) |
| **mMTC** | massive machine-type communication |
| **MU-MIMO** | multi-user MIMO |
| **MUD** | multi-user detection |
| **ND** | network device |
| **NGMA** | next generation multiple access |
| **NOMA** | nonorthogonal multiple access |
| **NR** | New Radio |
| **NTN** | non-terrestrial network |
| **OFDMA** | orthogonal frequency division multiple access |
| **OMA** | orthogonal multiple access |
| **PMF** | probability mass function |
| **QPSK** | quadrature phase shift keying |
| **RB** | resource block |
| **RF** | radio frequency |
| **RMSE** | root mean squared error |
| **SA** | situational awareness |

| | |
|---|---|
| **SC** | superposition coding |
| **SCMA** | sparse code multiple access (MA) |
| **SDMA** | space division multiple access |
| **SG** | stochastic geometry |
| **SIC** | successive interference cancellation |
| **SIF** | stand-in force |
| **SNR** | signal-to-noise ratio |
| **SWaP** | size, weight, and power |
| **TCP** | Transmission Control Protocol |
| **TDD** | time division duplex |
| **TDMA** | time division multiple access |
| **TSCC** | temporal strongly connected component |
| **TSG-RAN** | Radio Access Network Technical Specification Group |
| **TWCC** | temporal weakly connected component |
| **UAV** | unmanned aerial vehicle |
| **URLLC** | ultra-reliable and low-latency communication |

THIS PAGE INTENTIONALLY LEFT BLANK

# Acknowledgments

# CHAPTER 1:
## Introduction

Mobile wireless networks are a pervasive component of modern life. At the conclusion of 2021, the number of mobile subscribers exceeded the global population by a ratio of approximately five mobile subscriptions for every four people [1], [2]. In conjunction with core networks, wireless networks provide a foundation for the creation, sustainment, and evolution of many other networks (e.g., technological, information, and social) that are endemic to the connection-centric human experience. Thus, the methods by which wireless networks enable connectivity, and the robustness of that connectivity to parametric variations, must be studied and understood. This dissertation endeavors to make a contribution to that end through a study of network connectivity in nonorthogonal multiple access (NOMA) wireless networks. The following sections discuss the overarching context in which this research unfolds and trace the through-line of mobile wireless network development to the motivation, guiding question, and methodological approach of this research.

## 1.1  Research Context

Before proceeding into the heart of this research, it is important to consider the context in which it is conducted. Many investigations into wireless networks occur within the domain of electrical engineering. This is evident from the significant volume of wireless network research compiled by the various communities within the Institute of Electrical and Electronics Engineers (IEEE). In contrast, this research effort takes place within the domain of information science. This distinction raises two questions: what is information science, and how does an information science investigation of wireless networks differ from electrical engineering?

### 1.1.1  What is Information Science?

Information, like most concepts and ideas, does not exist in isolation. Rather, as explained by Russell Ackoff, information exists as part of a hierarchy, beginning with data that is processed into information [3]. Pieces of information are placed in relationships to form knowledge, and knowledge is learned to yield understanding. Finally, understanding

is coupled with value-based judgements to culminate in wisdom [3]. As a result of the hierarchical relationship, information science draws on a diverse set of disciplines depending on the aspect of information under study. The broad scope of information science is well-articulated in Borko's discussion, where he contends that,

> *Information science* is that discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. It is concerned with the body of knowledge relating to the origination, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization of information. This includes the investigation of information representations in both natural and artificial systems, the use of codes for efficient message transmission, and the study of information processing devices and techniques such as computers and their programming systems. It is an interdisciplinary sciences derived from and related to such fields as mathematics, logic, linguistics, psychology, computer technology, operations research, the graphic arts, communications, library science, management, and other similar fields. It has both a pure science component, which inquires into the subject without regard to application, and an applied science component, which develops services and products. [4, p. 3]

Notice that, though Ackoff's discussion came over two decades later, Borko's enumeration of the bodies of knowledge with which information science is concerned implicitly recognizes Ackoff's hierarchy. For example, the origination and transmission of information relate to data. The interpretation, transformation, and utilization of information relate to knowledge and understanding. Borko goes on to explain that his definition is complicated "because the subject matter is complex and multidimensional, and the definition is intended to be all-encompassing" [4, p. 3]. The wide range of disciplines in information science reflect the breadth of inquiry required to consider information in relation to the other aspects of human and machine cognition.

It is certainly to Borko's credit that, despite the formulation of his definition in 1968, the flexibility in his vision has accommodated the technological advances of the past 54 years. Wireless networking is a clear example as it focuses on *artificial systems* that

2

*transmit information* represented as data signals, and is often considered through the fields of mathematics, computer technology, and communications (though the terminology has become more specialized over time). Thus, I adopt Borko's definition, recognizing its relationship to Ackoff's hierarchy, as the information science context for this research into NOMA wireless networks.

### 1.1.2   Information Science and Electrical Engineering

The expansive definition of information science might suggest that no clear distinction exists between an electrical engineering and information science approach to wireless networking. However, this is not the case. Though both are decidedly engaged in the development of theory, they differ in the questions that guide their pursuits. The electrical engineer develops theories and knowledge in pursuit of a system design objective, asking "How can I achieve design objective $x$?" The relationships described in the resulting theories are a means to an end. In contrast, the information scientist is guided by questions that seek to develop theories and knowledge as an end, such as "What is the relationship between $x$ and $y$?" As discussed in the following sections, this research attempts to span the boundaries of this distinction by developing theory as an end with the intent that it provides benefit to those who may further it as a means.

Differentiating between scientific and engineering disciplines by the nature of the questions asked places my views at odds with the end of Borko's definition, which draws a distinction between pure and applied information science. In my view, the distinction is actually between pure information science and various engineering disciplines (e.g., electrical, software). This idea extends to applied science writ large. Applied science *is* engineering, and the the applied scientist *is* an engineer. Once the purpose of theory design is bound to the design of an artifact (a product or service in Borko's words), an engineering design effort is surely underway.[1] While I suspect this idea may be met with disagreement from scientists and engineers alike, the important aspect on which to focus is the underlying commonality between the scientist and engineer. As Simon keenly observed, both are fundamentally engaged in design, and it is in this common space in which I hope this dissertation finds use to readers of many disciplines [5].

---

[1]This idea does not suggest that all engineering efforts include theory/knowledge generation. That is, all engineering is not applied science.

## 1.2 The Evolution of Mobile Wireless Networks

The evolution of radio frequency (RF) wireless networks is marked by a continual increase in network performance to meet user requirements. This holds true for mobile wireless networks as each successive generation has sought to outpace requirements through new technologies and capabilities.

### 1.2.1 The Past

In the early 1980s, the first generation (1G) of mobile wireless technology brought mobile analog telephony to the general public. The second generation (2G) made numerous advances throughout the 1990s including the transition from analog to digital transmissions, the introduction of circuit-switched data transmission (e.g., Short Message Service), and a shift to packet-switched data transmission (e.g., General Packet Radio Services), which opened the door for low data rate applications (e.g., email). However, the packet-switched data services were still supported by narrowband transmissions. In the early 2000s, third generation (3G) wireless technology advanced from narrowband to mobile broadband technology that supported Internet Protocol (IP)-based services previously restricted to fixed broadband internet connections. The fourth generation (4G) of mobile wireless technology, of which the 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) is the dominant standard, deployed circa 2010 and provided the increased throughput and decreased latency necessary to support the increasingly demanding application requirements that accompanied the transition to smart phones over the next decade [6]–[9].

### 1.2.2 The Present

The present day is witness to the initial deployment of commercial 3GPP fifth generation (5G) New Radio (NR) wireless networks, which began with South Korea's SK Telecom in April 2019 [10]. A comparison of the performance targets set by the Radiocommunication Sector of the International Telecommunications Union (ITU) for 4G and 5G networks is depicted in Figure 1.1. International Mobile Telecommunications (IMT)-Advanced (depicted in light green) corresponds to 4G networks, and IMT-2020 (depicted in dark green) corresponds to 5G networks. It is clear that 5G represents a significant step beyond 4G.
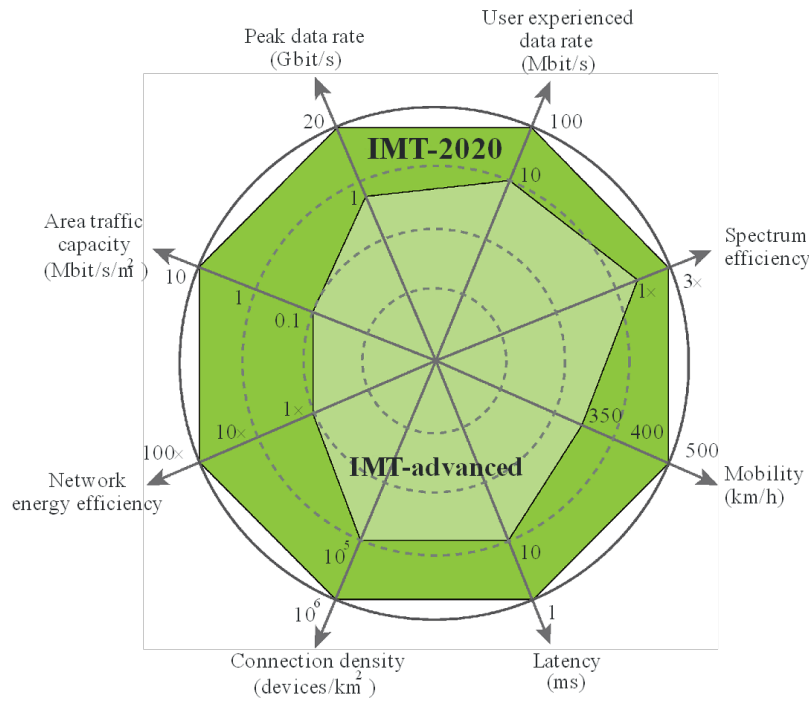
Figure 1.1. Comparison of IMT-2020 and IMT-Advanced performance performance requirements. Source: [11].

Fifth generation (5G) networks are largely distinguished from 4G through the diversity of use cases, which are broadly classified in the 2015 ITU-R Recommendation M.2083 as enhanced mobile broadband (eMBB), massive machine-type communication (mMTC), and ultra-reliable and low-latency communication (URLLC) [11]. The individual user experience dominates the eMBB use cases where the main objective is supporting high data rate applications, such as streaming 4k video, over a range of mobility and coverage area conditions. Conversely, mMTC use cases are generally characterized by large numbers of connected devices with lower data rate requirements, such as the Internet of Things (IoT), where maximizing connection density and device energy efficiency are of paramount importance. Use cases in the URLLC category may have elements of eMBB and mMTC, but are set apart by their stringent latency and reliability requirements, such as automated control systems and transportation safety [7], [12]. Each of these use case classes correspond to a different importance level for the IMT-2020 performance metrics, as depicted in Figure 1.2.
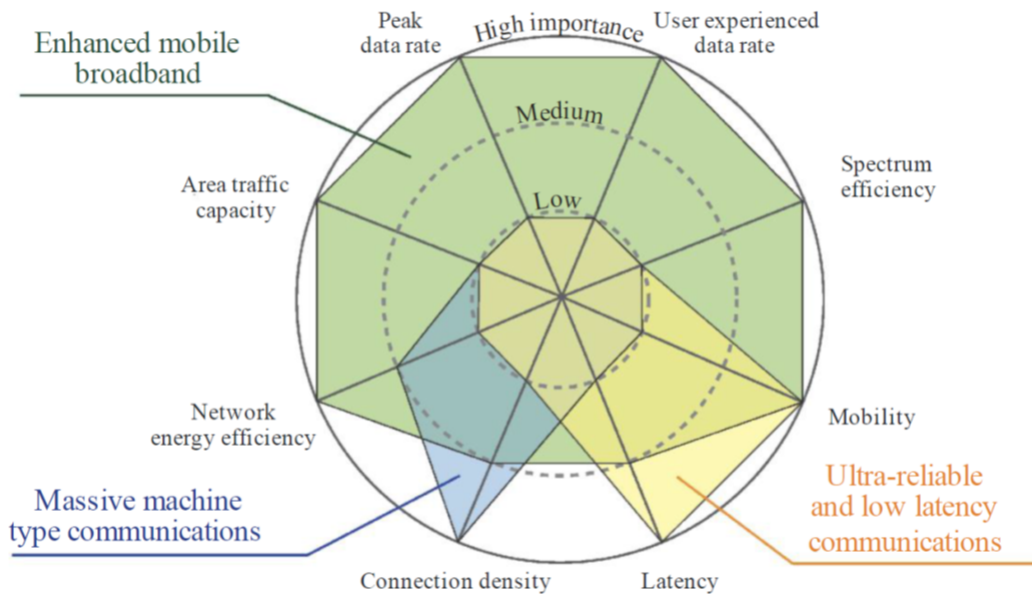
Figure 1.2. Mapping of IMT-2020 use cases to the relative importance of IMT-2020 performance metrics. Source [11].

### 1.2.3 The Future

While 5G networks continue their initial deployment, the research and development of beyond 5G (B5G)/sixth generation (6G) wireless networks is underway [13]. The vision for 6G wireless networks generally calls for an order of magnitude improvement over 5G in the key performance metrics of throughput, latency, connection density, reliability, and mobility. Example use cases for 6G networks include holographic communications, tactile and haptic internet applications (e.g., remote robotic surgery), and the Internet of Everything [14]. Supporting these diverse use cases requires a versatile network capable of terabit/second throughput, sub-millisecond latency, and connection densities of 10 million devices per $km^2$.

## 1.3 Motivation

The network performance gains in each generation of mobile wireless technology discussed in Section 1.2 are supported by progress in a variety of technologies, including medium access control (MAC). While the advances in MAC technology for 1G through 4G generally correspond to a primary objective (e.g., transition from circuit-switching to packet switching, transition from narrowband to broadband), the myriad ambitious performance requirements for 5G and B5G/6G require a MAC design that supports multiple objectives.

Among the variety of nascent MAC approaches for next generation multiple access (NGMA) is NOMA. The central concept in NOMA is *overloading*, which is the assignment of more than one network device (ND) to transmit or receive RF signals using the same time-frequency resources, or resource blocks (RBs). Through overloading, NOMA has the potential to support the increased throughput, higher connection densities, increased spectral efficiency, and lower latency of future 5G and B5G/6G networks [15]–[17].

The 3GPP Radio Access Network Technical Specification Group (TSG-RAN) conducted a formal study of NOMA for inclusion in 3GPP NR in 2018 [18]. The results of the study, and recorded discussion from TSG-RAN Meeting 82, suggest that meeting participants did not believe that NOMA demonstrated sufficient performance gains, when compared to multiple input multiple output (MIMO) technologies, to warrant near-term standardization in NR [19]–[21]. Specifically, massive MIMO (mMIMO) and multi-user MIMO (MU-MIMO) are widely accepted multi-antenna approaches to meet IMT-2020 requirements in the beam-centric NR design [7], [21]. While this perspective may be valid for mobile wireless networks supported by fixed infrastructure, it fails to consider the increasing attention given to non-terrestrial network (NTN) deployments that are partially or entirely composed of mobile base stations (MBSs), such as low-altitude unmanned aerial vehicles (UAVs) [22]–[27]. The MBS-to-ground communications channel is generally characterized by a strong line-of-sight component that increases channel correlation and reduces the requisite diversity for spatial multiplexing [24]. Additionally, the size, weight, and power constraints of some MBSs may prevent them from supporting mMIMO arrays in the sub-6 GHz frequency range.

Consider the 64-element mMIMO array from [28] operating at a carrier frequency of 6 GHz or below. This frequency range is important to support a larger coverage area from the MBS, and to provide connectivity with NDs that may not have line-of-sight to the

MBS [28]. Approximating the speed of light as $3 \times 10^8$ m/s, the wavelength of a 6 GHz carrier is 50 mm. An $8 \times 8$ element planar array with $1/2$ wavelength elements and $1/2$ wavelength separation between elements requires 375 mm $\times$ 375 mm, or approximately 1.5 ft$^2$, of space on the MBS. This space requirement will increase with the addition of antenna elements or reduction in frequency. Additionally, realizing the full MU-MIMO potential of the array that is enabled by digital beamforming requires a digital-to-analog converter (DAC) for every antenna element. This fully digital configuration allows the array to simultaneously support the same number of network devices as array elements in each RB, but at the cost of increased space and power consumption from the DACs [7], [28]. While these constraints may be easily accommodated by a static base station, they could prove challenging to implement on smaller MBSs with other competing objectives such as maximizing time-on-station. Conversely, implementing NOMA does not impose additional array space constraints on the MBS.

NTN wireless networks supported by MBSs could be deployed in a number of scenarios such as augmentation to fixed infrastructure experiencing high traffic volume (large public event), providing service to a location at which fixed infrastructure has failed (natural disaster response), or providing service in a location not served by fixed infrastructure (military operations[2]) [26], [29], [30]. In these scenarios, NOMA is particularly suited for meeting massive access and low latency requirements for IoT and federated learning applications that are unable to access centralized computing resources [16], [31]. Although currently lacking momentum in the 3GPP NR standardization process, these use-cases coupled with the potential challenges of mMIMO implementation in small MBSs suggest that NOMA has a role to play in the evolution of 5G and B5G/6G wireless networks.

## 1.4   Research Overview

This section provides an overview of the research gap, purpose, guiding question, and approach to theory development.

---

[2]An expanded discussion of specific applicability to the Marine Corps is provided in the appendix.

### 1.4.1 Research Gap

Information theoretic treatments of NOMA consider communications channel capacity boundaries in the Broadcast and Multiple Access Channels. These exploratory studies provide an understanding of the best achievable transmissions rates in which the number of users sharing a RB is arbitrarily large [32]–[34]. However, researchers exploring physical implementations of NOMA have noted many practical considerations such as high receiver complexity, variable channel conditions, minimum required transmission rates, and codebook collision that suggest the number of NDs that can effectively use the same RB is bounded and variable [20], [35]–[39]. Correspondingly, much of the NOMA literature investigates methods by which it can be employed to achieve IMT-2020 performance targets under practical constraints; however, *there is a general lack of analysis on the relationship between bounded variable overloading and network connectivity* [15], [18], [33], [40]–[43]. This research aims to address that gap.

### 1.4.2 Research Purpose and Guiding Question

As alluded to in Section 1.1, this research aims to span the boundary between the information scientist and electrical engineer. Thus, the purpose of this research is to contribute to the understanding of NOMA wireless networks through the development of a theoretical model and analysis that inform their practical design. To that end, this research is guided by the following question:

*How does variable overloading affect robustness in NOMA wireless networks?*

This question is approached through the theoretical lens of network science, with a particular emphasis on temporal network theory, and considers robustness in terms of the time-varying connectivity between NDs. In addition to achieving the expressed aims, we hope that the results of this study will find applicability in a variety of isomorphic contexts across the interdisciplinary fields of which information science is composed.

### 1.4.3 Theoretical Development

The evolutionary development of theory is at the core of this research effort. Informed by the research gap in Section 1.4.1 and the design of NOMA networks to intentionally

interfere, we consider the relationship between NOMA overloading and robustness from the perspective of connectivity rather than throughput.

The connectivity between entities is a prominent focus of network science, which explores network connectivity through graph-theoretic representations. In this context, we consider the role of overloaded RB allocation in network connectivity and, taking inspiration from interdependent network theory, conceptualize a graph-based NOMA system model that represents the relationships between NDs, RBs, and a single base station (BS) in a NOMA wireless network. Drawing from random graph theory and temporal network theory, we develop the system model into a mathematically formalized temporal network ensemble that is suitable for time-varying connectivity analysis. The system model and temporal network ensemble offer a novel theoretical perspective to investigate the relationship between variable overloading and robustness, and provide the basis for subsequent theoretical contributions.

Building on the theoretical foundation of the system model, we proceed to consider the implications for RB allocation and the corresponding effect on the probability of connectivity between NDs. We recognize the system model assumptions and parameters lend RB allocation to representation as a Bernoulli random process, and this insight drives the stochastic analysis of connectivity between NDs. Probabilistic connectivity is initially considered through ND membership in temporal network components, resulting in the development of a stochastic temporal component framework. This framework provides a theoretical characterization of the expected type of connectivity between NDs over time. However, it becomes clear at the conclusion of the temporal component framework analysis that queued messages from one ND to another (which are not represented in the temporal component framework) will have a significant impact on temporal connectivity. This realization shifts the research focus from temporal components to directed connectivity between ND pairs, and leads directly to the development of a mathematical model of connectivity that accounts for queued messages.

After formalizing the mathematical model of connectivity, we consider many different questions that each have the potential to support NOMA network design for the motivating use-cases.

1. What is the probability that ND $i$ connects to ND $j$ at time $t$?
2. When should we expect that ND $i$ *first* connects to ND $j$?

10

3. How much time passes between a connection from ND $i$ to ND $j$ and the next?
4. When should we expect that ND $i$ and ND $j$ are first bidirectionally connected?
5. How quickly can all NDs in the network connect, and how likely is this to occur?

The first and third questions support design for ND density and duty cycle, while the second, fourth and fifth questions all support design for decentralized algorithm convergence among NDs. The results from these investigations are expressed as stochastic mathematical equations that each contribute new theoretical insights about the impact of variable overloading on NOMA network robustness.

**Theory Development Cycle**

Although the presentation of theory development and simulation validation in Chapters 5 and 6 is sequential, this process occurs in a cyclic manner similar to the spiral development method for computer software. Each cycle is composed of an analytical and simulation phase. The analytical phase considers the enumerated questions (about evolving connectivity between NDs) in terms of the NOMA system model formalism and derives a stochastic expression, or probability mass function (PMF), as a result. Each of these analytical results is a theory of how temporal connectivity evolves in the network, as a function of NOMA overloading. The simulation phase tests the internal validity of the theory by instantiating the mathematical formalism in the MATLAB® technical computing environment and conducting trials to determine if the connectivity behaves as the theoretical expressions predict. The results and insights gained during one theory development cycle often generate the question that begins the next. In this way, the research effort is self-sustaining as each new cycle leads to new areas of exploration.

The following section outlines the specific theoretical contributions of this research, along with the chapter and section in which they occur.

## 1.5 Contributions

The theoretical contributions resulting from this research are summarized in Table 1.1. We believe the NOMA graph model, temporal network ensemble, and stochastic characterizations of time-varying network connectivity simultaneously contribute to the bodies of wireless networking and temporal network theory research. Sections marked by (*) denote

Table 1.1. Theoretical Contributions of Research.

| Section | Contribution |
|---|---|
| Ch. 3.1* | NOMA Graph Model |
| Ch. 3.2* | Temporal Network Ensemble |
| Ch. 4* | Stochastic Temporal Component Framework |
| Ch. 5.2† | Probability of Temporal Connectivity in Each Frame |
| Ch. 5.3† | Probability of Time to Initial Unidirectional Connectivity |
| Ch. 5.4 | Probability of Time Window Between Unidirectional Connectivity |
| Ch. 5.5 | Probability of Time to Initial Bidirectional Connectivity |
| Ch. 5.6* | Probability of Minimum Time to Complete Bidirectional Connectivity |

work that has been refereed and accepted at the 2022 Hawaii International Conference on System Sciences, and sections marked by (†) denote work that has been refereed and accepted at the 2022 IEEE International Conference on Communications.

## 1.6 Organization

The remainder of this dissertation is organized as follows. In Chapter 2, an overview of multiple access (MA), temporal network theory, and related work is provided. In Chapter 3, a temporal graph model of a NOMA wireless network is proposed, a temporal network ensemble is developed, and the relationship to existing graph models is discussed. In Chapter 4, the probability of RB allocation is considered, and an analytical framework for temporal component analysis is developed. In Chapter 5, the relationship between NOMA overloading and temporal connectivity between NDs is explored. In Chapter 6, the consolidated simulation results from each theory development cycle are presented. In Chapter 7, the dissertation is concluded with a brief review of contributions, limitations, and suggestions for future work. A focused discussion of the Marine Corps relevance of this research is provided in the appendix.

# CHAPTER 2:
# Background and Related Work

This chapter reviews aspects of wireless multiple access, temporal network theory, and related work to develop the requisite context for the remainder of the dissertation. First, a brief overview of orthogonal multiple access schemes is provided to motivate the discussion of NOMA. Subsequently, the definition of NOMA adopted in this research is presented followed by an introduction to fundamental NOMA concepts. Next, temporal network theory is introduced with a specific emphasis on temporal components. Finally, a review of prior work related to this dissertation is provided.

## 2.1   Orthogonal Multiple Access

A basic function of MAC protocols is the management of resources in the time-, frequency-, code-, and space-domain by which information is transmitted [15]. Medium access protocols designed to accommodate shared access to transmission resources by multiple NDs are called MA schemes.

Multiple access technologies that separate ND transmissions in time, frequency, code, space, or some combination thereof are classified as orthogonal multiple access (OMA) schemes. In general, OMA is designed to prevent interference between NDs so that all transmissions can be successfully received and decoded.

Many different OMA approaches exist, including frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), space division multiple access (SDMA), and orthogonal frequency division multiple access (OFDMA) [8]. In FDMA, the available frequency spectrum is divided into channels that are allocated to individual NDs for transmission and reception. Similarly, in TDMA, each ND is assigned a specific time period during which it can transmit or receive over the entire frequency band. In CDMA, NDs transmit at the same time across the entire frequency band, and orthogonal spreading codes are employed to disambiguate their interfering signals. Similarly, SDMA distinguishes ND signals through directional spatial beams transmitted at the same time and frequency. OFDMA is a blend of FDMA and TDMA that combines RF channels with

time periods to create time-frequency RBs, which are allocated to NDs for transmission and reception [8], [33], [44]. A comparative illustration of these OMA schemes and an OFDMA RB are depicted in Figure 2.1.
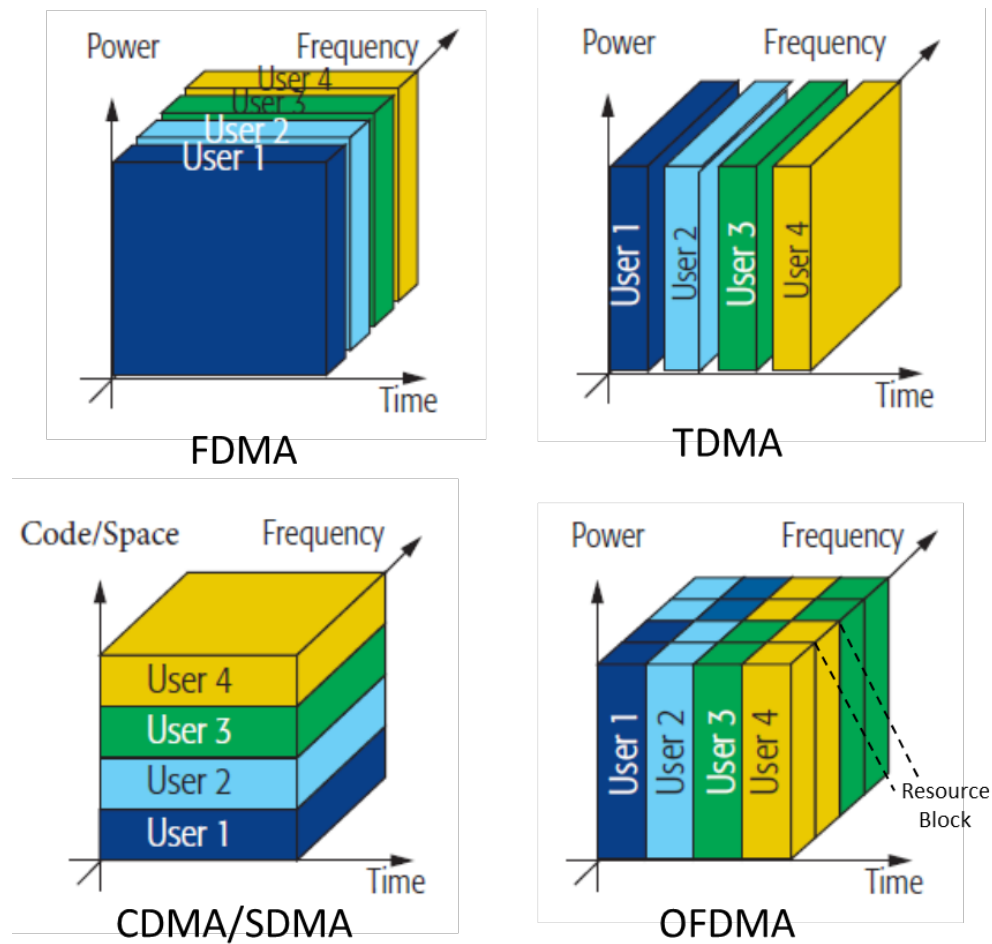


Figure 2.1. Comparison of different OMA schemes. Adapted from [44].

Despite the success of OMA approaches in managing interference while sharing transmissions resources, the scalability of OMA to support increasing numbers of NDs is inherently limited by the orthogonality constraint (i.e., the total number of transmission resources) [33].

## 2.2 Nonorthogonal Multiple Access

As alluded to in Chapter 1.3, NOMA is a family of MA approaches that allow multiple NDs to transmit or receive RF signals using the same transmission resources. Depending on the MA technology, transmission resources include some subset of time, frequency, code, and space. Simultaneous use of the same transmission resources (employed by a given MA technology) by more than one ND is called *overloading*, and the number of NDs sharing these resources is the overloading ratio. A conceptual illustration of NOMA is provided in Figure 2.2.
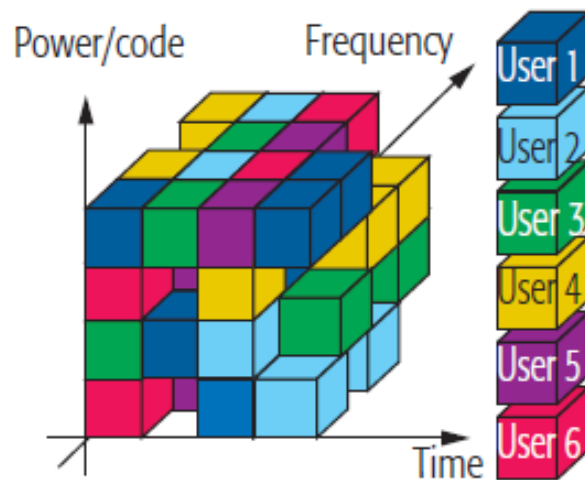


Figure 2.2. Conceptual view of NOMA schemes that is applicable for power-domain or code-domain. Adapted from [44].

In contrast to OMA, NOMA overloading explicitly induces RF interference between NDs by design. Nonorthogonal multiple access techniques mitigate the intentional overloading interference through multi-user detection (MUD) methods broadly categorized as power-domain or code-domain. Power-domain approaches differentiate between NDs using a combination of direct superposition coding (SC) at the transmitter and successive interference cancellation (SIC) at the receiver [45]. Code-domain schemes generally employ more complex forms of SC, such as nonorthogonal low cross-correlation spreading sequences or interleaving, and often achieve MUD through iterative detection methods (such as the message passing

algorithm) that may or may not be paired with SIC [46]–[50].

Similar to the views expressed in [34], *this research broadly defines NOMA as any MA approach that enables overloading*. This includes those already defined as NOMA in the literature, such as power-domain NOMA and code-domain NOMA, and other techniques such as rate-splitting [51]. Despite the wide variety of NOMA approaches, *overloading* is the central concept that opens the door for performance gains in aggregate throughput, connection density, and spectral efficiency; thus, the overloading ratio is the key parameter of interest, and the theoretical developments of this research are intended to be applicable to any MA scheme for which an overloading ratio can be defined.

Given the generalizable intent of this inquiry, the subsequent sections sketch the concepts of SC and SIC from the information theoretic perspective of the downlink (broadcast) channel. They are not intended as a comprehensive review, but rather to provide intuition for some of the fundamental concepts that permeate (individually or in concert) many NOMA implementations across categories.

### 2.2.1   Multiuser Capacity in the Broadcast Channel

This section introduces SC and SIC via the capacity analysis of the broadcast (downlink) channel. Unless otherwise attributed, all material in Section 2.2.1 is derived from Chapter 6 of [32] and the example in the last section is taken from [52].

**Channel Capacity and Achievable Transmission Rate**

Like all other wireless communications approaches, NOMA schemes are constrained by Shannon's channel capacity theorem [53], which states

$$C = W \log \left( 1 + \frac{P}{N_0 W} \right) \tag{2.1}$$

where $C$ is the channel capacity in bits/second, $W$ is bandwidth in Hz, $P$ is the received signal power in Watts, and $N_0$ is the power spectral density of additive white Gaussian noise (AWGN) in Watts/Hz. The achievable transmission rate for any communications channel cannot exceed this bound, including in multiuser channels that support communications between more than two NDs. However, Tse and Vishwanath show in [32] that the sum rate

capacity of multiuser communications channels employing NOMA schemes can exceed that of OMA. This is achieved through a combination of SC and SIC.

Consider the downlink communications channel from a single BS to multiple NDs represented in baseband discrete-time by

$$y_k[m] = h_k x[m] + w_k[m], \tag{2.2}$$

where $y_k[m]$ is the received signal by ND $k$ at time $m$, $h_k$ is the complex channel gain for ND $k$, $x[m]$ is the transmitted signal from the BS, and $w_k[m]$ is AWGN [32]. Given the BS has an average power constraint $P$ for each transmission period then, by the channel capacity in Equation (2.1), the transmission rate of single ND $k$ is bounded by

$$R_k < W \log\left(1 + \frac{P|h_k|^2}{N_0 W}\right), \tag{2.3}$$

where $R_k$ is the rate of ND $k$ in bits/second [32]. Extending this to the case of two NDs with equal channel gains that share orthogonal transmissions resources yields the achievable capacity region described in [32] as

$$R_1 + R_2 < W \log\left(1 + \frac{P|h_1|^2}{N_0 W}\right). \tag{2.4}$$

This region, illustrated in Figure 2.3, includes all achievable rate pairs for both NDs and the boundary demonstrates that the single ND channel capacity limit is the best achievable rate. However, observing that both NDs have equal channel gains, Tse and Vishwanath explore the idea that each ND is capable of decoding the signal of the other. In this case, if both signals are simultaneously transmitted, NDs can decode and subtract interfering signals from the aggregate to decode their own signal with less interference.
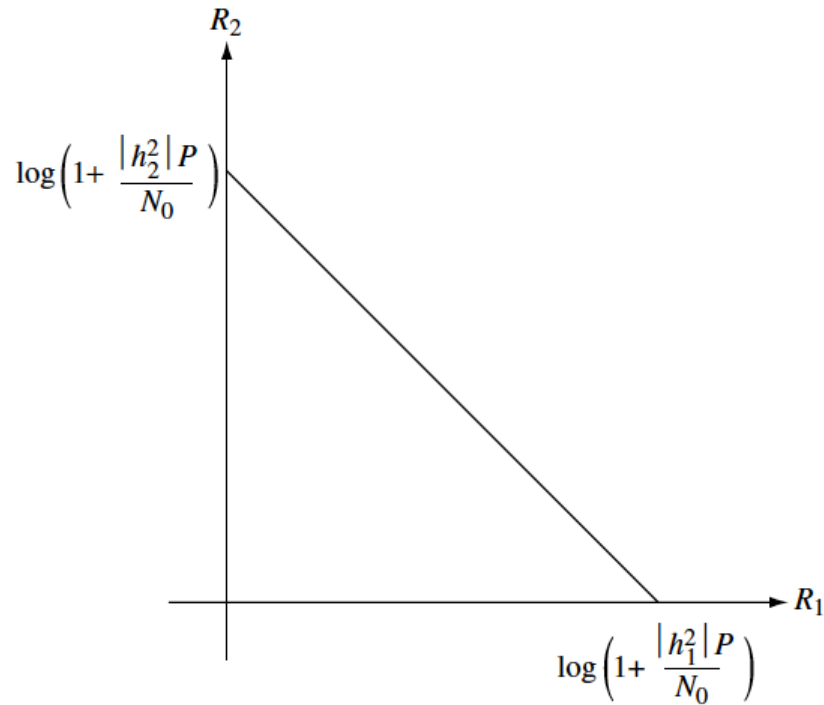
Figure 2.3. Downlink capacity region of two NDs with $|h_1| = |h_2|$ and $W = 1$ Hz. Source: [32].

**Superposition Coding and Successive Interference Cancellation**

Summing the information of two or more signals into a single transmission is called superposition coding. In the case of two NDs, the transmitted signal across the entire bandwidth from [32] is

$$x[m] = x_1[m] + x_2[m], \tag{2.5}$$

and the transmission power is split between two NDs, $P_1$ and $P_2$ where, $P_1 + P_2 = P$. An example of superposition coding using a quadrature phase shift keying (QPSK) modulation constellation is depicted in Figure 2.4.
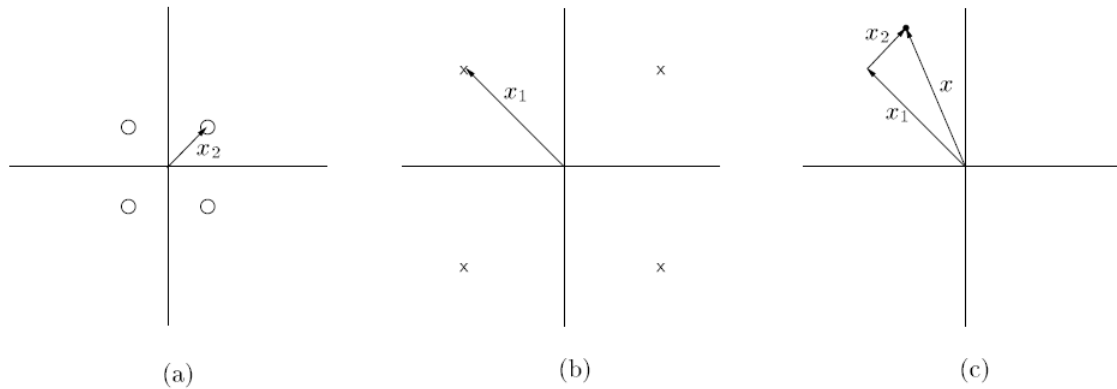
Figure 2.4. Superposition coding example. The QPSK constellation of $x_2$ (a) is superimposed on the constellation of $x_1$ (b) to produce the combined signal (c). Source: [32].

Successive interference cancellation is the process of decoding a signal from an aggregate received signal (composed of other interfering signals and noise), subtracting the decoded signal from the aggregate received signal, and proceeding to decode another signal from the remaining aggregate received signal. In this way, signals are successively canceled from the aggregate received signal, thereby reducing the interference until the desired signal is decoded. An example of a superposed signal decoded through SIC is illustrated in Figure 2.5.

19
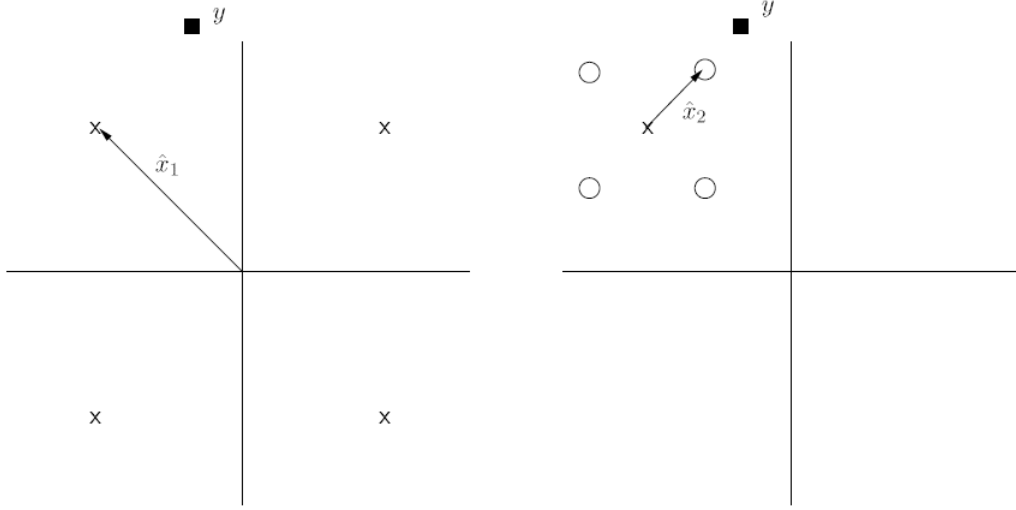
Figure 2.5. Successive interference cancellation example. The received signal of $\hat{x}_1$ is decoded first followed by the received signal of $\hat{x}_2$. Source: [32].

The rate equations describing multiple access schemes employing SC and SIC differ from those that do not. Network device 2 treats the signal of ND 1 as noise, and the achievable rate from [32] is

$$R_2 = W \log \left( 1 + \frac{P_2 |h_2|^2}{P_1 |h_1|^2 + N_0 W} \right).$$  (2.6)

Network device 1 performs SIC: it decodes the signal $x_2[m]$, treating the rest of the signal as noise, subtracts $x_2[m]$ from the aggregate received signal, and then decodes the desired signal $x_1[m]$ without interference, yielding the achievable rate from [32] as

$$R_1 = W \log \left( 1 + \frac{P_1 |h_1|^2}{N_0 W} \right).$$  (2.7)

The successful use of SIC with SC requires that each user signal is decoded in the order of increasing complex channel gains (i.e., decoding must begin with the ND that has the weakest channel gain). The transmitter selects rates for each ND that are decodable based on the allocated transmission power and channel gain. Thus, if the ND with the weakest channel gain is able to decode its signal, then all other NDs are also able to decode the signal and subtract it from the aggregate. Equations (2.6) and (2.7) represent this process in

20

the two-device case, and it extends for multiple rounds of cancellation in the multi-device case. Notice this requires the transmitter and all NDs to know the complex channel gains between each device and the transmitter.

**Comparison of OMA and NOMA**

The achievable rate equations in (2.6) and (2.7) provide a means to understand whether employing SC and SIC for nonorthogonal resource sharing can provide any gain with respect to orthogonal schemes. The transmission rates of two NDs employing an orthogonal resource sharing scheme, as described in [32], are

$$R_1 = \alpha W \log \left( 1 + \frac{P_1 |h_1|^2}{\alpha N_0 W} \right), \tag{2.8}$$

and

$$R_2 = (1 - \alpha) W \log \left( 1 + \frac{P_2 |h_2|^2}{(1 - \alpha) N_0 W} \right). \tag{2.9}$$

In addition to the power constraint, $P = P_1 + P_2$, the transmission rates of orthogonal schemes must account for resource sharing, so let $\alpha$ denote the allocation of resources between NDs, where $0 \leq \alpha \leq 1$. In this case, $\alpha$ corresponds to bandwidth, but it generalizes to any orthogonal transmission resources that must be shared (e.g., time). The total transmission bandwidth available ($W$) is fixed, so the allocation among the two NDs is $\alpha$ and $1 - \alpha$, respectively. A contrast of resource sharing between OMA and NOMA during a single transmission is shown in Figure 2.6 where $|h_1|^2/N_0$ and $|h_2|^2/N_0$ are constant during the transmission and $|h_1|^2 > |h_2|^2$. Figure 2.6(a) shows the power split and orthogonal bandwidth allocation to each ND, while Figure 2.6(b) shows the power split and nonorthogonal bandwidth allocation.
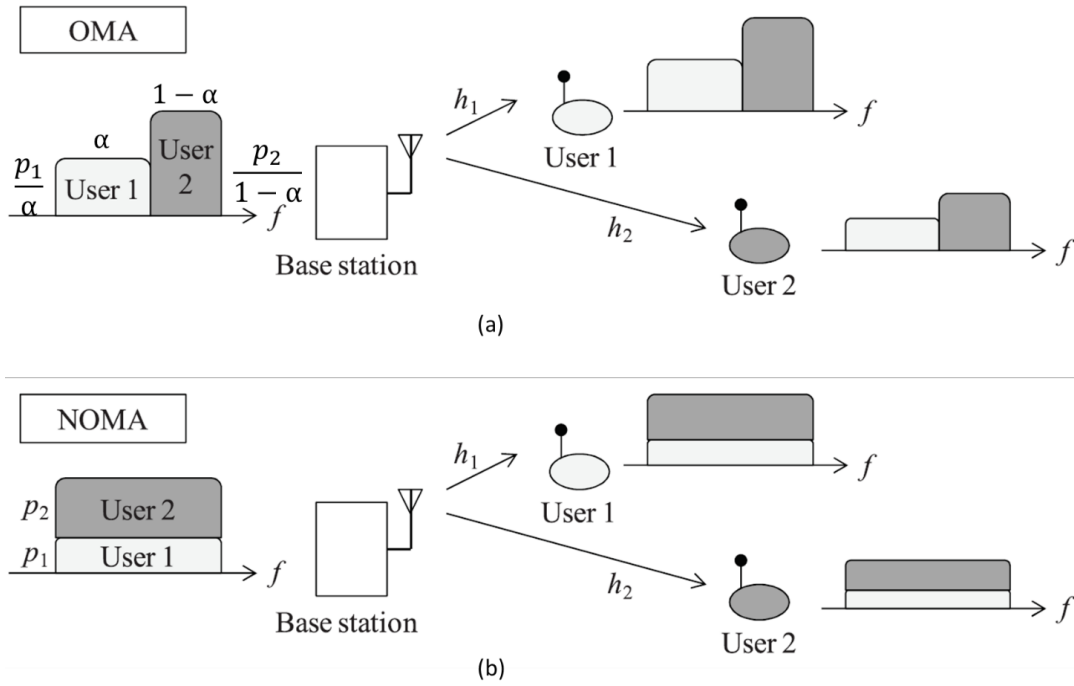
Figure 2.6. Comparison of resource sharing in OMA and NOMA: (a) power sharing and orthogonal resource sharing; (b) power sharing and nonorthogonal resource sharing. Source: [54].

The corresponding comparison of achievable rate pairs $(R_1, R_2)$ for OMA and NOMA in the symmetric and asymmetric communications channel is depicted in Figure 2.7. The OMA rates pairs are calculated by jointly optimizing the power split and resource sharing, while the NOMA rate pairs are only constrained by the power split.
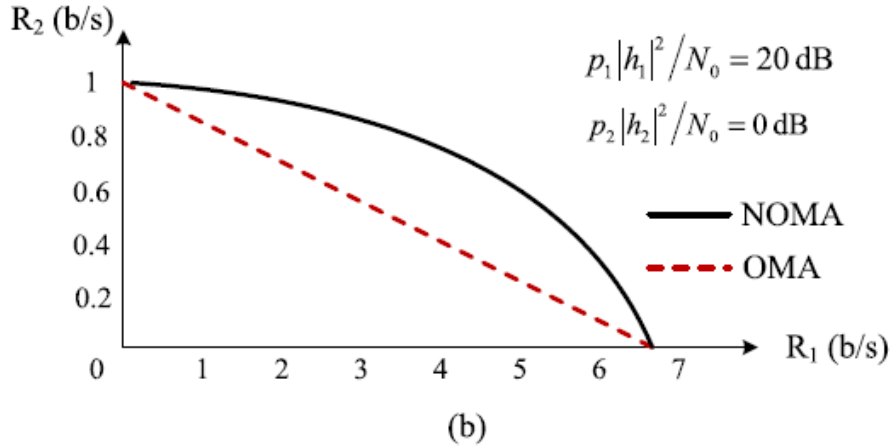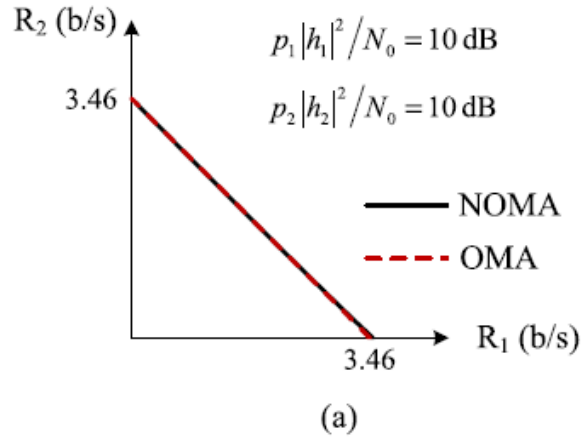
Figure 2.7. Channel capacity comparison of OMA and NOMA in the downlink AWGN channel: (a) symmetric channel; (b) asymmetric channel. Sources: [33], [54].

In the case of a symmetric communications channel, with equal complex channel gains $|h_1|^2 = |h_2|^2$, the boundary of the achievable rate region is the same for OMA and NOMA. In the case of an asymmetric communications channel, with unequal complex channel gains $|h_1|^2 > |h_2|^2$, the boundary of the achievable rate region for NOMA exceeds OMA in all cases where both users have a rate greater than zero. Thus, NOMA does not provide any gain in the symmetric channel, but does provide gains in the asymmetric channel. This is because the difference in channel gains stratifies the NDs into the bandwidth-limited and

23

power-limited signal regions. The bandwidth-limited device with the stronger channel gain benefits by accessing the entire bandwidth while sacrificing some of the power allocation from the transmitter. The power-limited device with the weaker channel gain benefits from a larger power allocation by the transmitter, thereby increasing the effective bandwidth, and a smaller interfering signal from the bandwidth-limited user due to the power split constraint [52].

The results for AWGN uplink channels are similar to the downlink, where NOMA is shown to be optimal in the symmetric and asymmetric cases, with many segments of the boundary exceeding OMA in both cases. However, the rate equations differ due to the power constraint imposed at each ND rather than a central BS.

### 2.2.2   NOMA Summary

In summary, the central idea of NOMA is overloading, in which multiple NDs simultaneously use the same transmission resources across all degrees of freedom. Superposition coding and SIC are two of the fundamental concepts that enable overloading across a variety of NOMA implementations. This research contributes to a generalized understanding of the relationship between overloading and time-varying network robustness by abstracting the physical NOMA implementation through network science. The following section introduces the specific elements of network science that are pertinent to this research.

## 2.3   Temporal Network Theory

The analysis and results of this research are developed within the broad theoretical framework of network science. In general, network science aim to understand the relationship between the structure and function of a network [55], [56]. Networks represent relationships, or connections, among entities. Correspondingly, the connectedness of a network is a structural property of particular interest, and a common measure employed to map from network structure to network function [55], [56]. Beginning with the representation of networks as graphs, and extending through static networks to temporal networks, the following discussion provides definitions of network connectedness and components that are relevant to this investigation of NOMA wireless networks.

### 2.3.1 Static Networks

Network science is largely underpinned by the mathematics of graph theory, where a network is represented by a graph $G$ composed of a non-empty set of vertices, $v \in V$, and a set of edges, $e \in E$ [55], [57]. Vertices represent entities, edges represent the relationships between those entities, and the entire network is denoted by $G(V, E)$. In the parlance of network science, vertices are called *nodes* and edges are called *links*. An example graph is illustrated in Figure 2.8. The nodes in the graph, represented by the circles, could represent people, airports, or internet routers. The links in the graph, represented by lines between the circles, might represent friendships between people, routes between airports, or logical IP connections between internet routers.
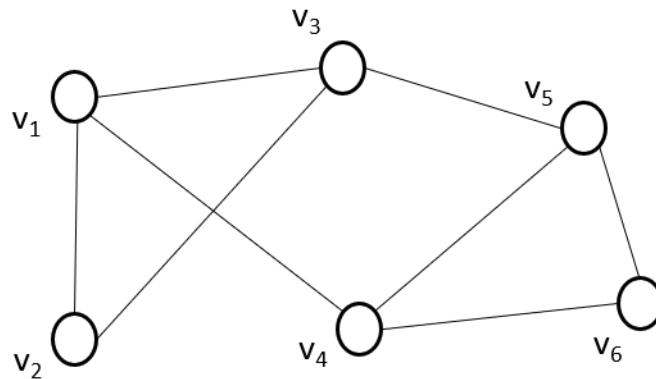


Figure 2.8. Example graph composed of six vertices and eight edges.

A network in which a link between two nodes represents a bidirectional relationship is called an *undirected* network. Conversely, networks in which a link between two nodes represents a directional relationship are called *directed* networks. Directed networks are visually represented with an indicator of the relationship direction, such as an arrow, whereas undirected networks are often depicted the same as in Figure 2.8.

**Node Adjacency**

Two nodes that are joined by a link are *adjacent* to each other. For example, in Figure 2.8, $v_1$ and $v_2$ are adjacent. The links between all nodes in the network can be mathematically represented through a *adjacency matrix*. The adjacency matrix of an undirected network **A**

is an $|V| \times |V|$ matrix with elements, denoted $A_{ij}$, defined by

$$A_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}, \qquad (2.10)$$

where $|\cdot|$ denotes the set cardinality [55]. The adjacency matrix for the undirected network in Figure 2.8 is

$$\mathbf{A} = \begin{cases} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{cases}. \qquad (2.11)$$

The main diagonal of $\mathbf{A}$ is composed of all zeros since no nodes are connected to themselves (i.e., there are no self-links), and the matrix is symmetric about the main diagonal because all links are bidirectional (i.e., the network is undirected).

The adjacency matrix $\mathbf{A}$ for a directed network is similar to an undirected network, but may not be symmetric. Assuming that nodes do not connect to themselves, the main diagonal will still be all zeros. However, the links in the network are *from* one node *to* another resulting in the modified definition for the elements of $\mathbf{A}$ [55], given by

$$A_{ij} = \begin{cases} 1 & \text{if there is a link } from\ v_i\ to\ v_j \\ 0 & \text{otherwise} \end{cases}. \qquad (2.12)$$

Notice that Equation (2.12) is the more general definition since Equation (2.10) can be considered a special case of Equation (2.12) in which all links are bidirectional.

**Walks and Paths**

Inherent in the concept of a network is the idea of movement. An airline network describes how a person might travel from one destination to another. A friendship network might describe how information or affection moves from one person to another. Regardless of the application, it is important to formally characterize the ability to move through or across

the network. This is accomplished through the concepts of walks and paths.

A *walk* is a series of nodes in the network where each consecutive pair of nodes are adjacent [57]. If the walk begins and ends at the same node, it is called a closed walk. If the walk ends at any node other than the one at which it began, the walk is open [57]. The number of links traversed during the walk is the length of the walk. Both nodes and links may be revisited during a walk, so the walk length is not limited.

A *path* is a walk in which no nodes are repeated [57]. Similar to walks, the number of links traversed on the path is the path length. A path between two nodes, $i$ and $j$, with a path length less than or equal to all other path lengths is a *shortest path* [57]. Paths, and shortest paths, often have more utility than walks in network science since real networks often seek to minimize the length of a walk (e.g. to reduce travel time in a transportation network, or latency in a communications network) [55]. Thus, the remaining discussion will only consider the role of paths in network connectedness.

**Connectedness and Components**

Two nodes $i$ and $j$ in an undirected network are connected if there is a path between them [55], [56], [58]. An undirected network is *connected* if there is a path between every pair of nodes in the network; otherwise, the network is disconnected [55], [56], [58]. The network in Figure 2.8 is connected because each node has a path to every other node in the network. If the links connecting the node pairs $(v_3, v_5)$ and $(v_1, v_4)$ are removed, the network is disconnected.

Disconnected undirected networks are characterized by connected subgraphs, called components, which are maximal subsets of nodes in which all nodes have a path to all others [55], [56]. Continuing with the example that the adjacency between $(v_3, v_5)$ and $(v_1, v_4)$ is removed, the graph is composed to two subgraphs with node sets $V_1 = \{v_1, v_2, v_3\}$ and $V_2 = \{v_4, v_5, v_6\}$, as shown in Figure 2.9.
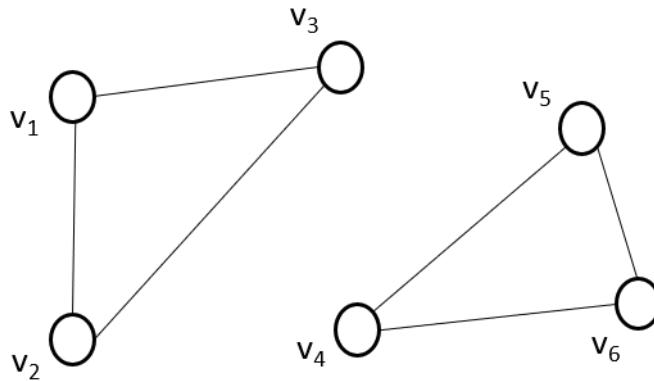
Figure 2.9. Example disconnected graph resulting from the removal of connections between node pairs. The disconnected graph is composed of two subgraphs with node sets $V_1 = \{v_1, v_2, v_3\}$ and $V_2 = \{v_4, v_5, v_6\}$.

The concepts of connectedness and network components extend to directed networks, in which a link represents a directional relationship (sometimes called an arc). In directed networks, connectedness is divided into two categories, strong and weak. As described in [55], [56], two nodes $i$ and $j$ in a directed network are *strongly* connected if there is a path from $i$ to $j$ and a path from $j$ to $i$. Similarly, two nodes $i$ and $j$ in a directed network are *weakly* connected if there is a path from $i$ to $j$ *or* a path from $j$ to $i$. The categories of node-level connectedness also give rise to multiple types of components in directed networks, defined in [55] and [56] as:

- *Strongly Connected Components*: The strongly connected component of a node $i$ is the set of nodes $j$ such that there exists a directed path from $i$ to $j$ and a directed path from $j$ to $i$, $\forall j$.
- *Weakly Connected Components*: The weakly connected component of a node $i$ is the set of nodes $j$ such that there exists a directed path from $i$ to $j$ or a directed path from $j$ to $i$, $\forall j$.

Notice that the weakly connected component is the same as a component that is derived by treating the directed network as an undirected network.

**Bipartite Networks**

One particular class of network that is important in this research is a *bipartite* network [55], [56]. As the name suggests, a bipartite network is composed of two sets of nodes that are partitioned from each other. Nodes in each set can only form links with nodes in the opposite set. Bipartite networks are described by the notation $G(U, V, E)$, where $U$ is one node set, $V$ is the other node set, and $E$ is the set of links between nodes in $U$ and $V$. An example bipartite network is shown in Figure 2.10.
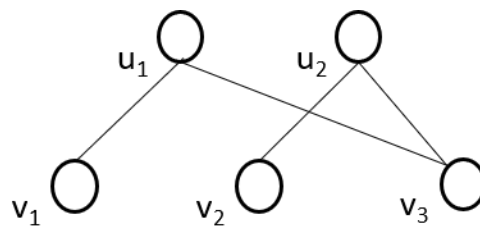


Figure 2.10. Example bipartite network composed of two node sets $U = \{u_1, u_2\}$ and $V = \{v_1, v_2, v_3\}$, where links can only occur between nodes from different note sets.

A common example of a bipartite network is the network describing actors and the films in which they have appeared [55]. The actors form the nodes of one set, the films form the nodes of the other node set, and the links connect actor nodes to the film nodes in which they have appeared. The role of bipartite networks as they relate to NOMA wireless networks is discussed in Chapter 3.1.

## 2.3.2   Temporal Networks

The ideas of connectedness and components discussed thus far apply to *static* networks. Static networks are those in which the network structure *does not* change over time. Conversely, temporal networks (or time-varying networks) are those in which the network structure *does* change over time. Temporal network theory extends static network theory to incorporate the effects of a time-varying network structure, and the time-varying nature of wireless communication networks lends itself to temporal representation and analysis [59]–[62].

**Temporal Network Representations**

Unlike a static network, two nodes in a temporal network that are connected at time $t_1$ may not be connected at a later time, $t_2$. The connectivity dynamics of temporal networks require conceptualizations and representations that differ from static networks. Many different representations exist including contact sequences, graph sequences, and dynamic networks [59].

Contact sequence, or event-based, representations of temporal networks capture the changing network structure through a time-ordered event list that includes the interacting nodes, the time at which the interaction began, and the duration for which the interaction lasted [59], [60]. This method is well-suited for temporal networks that change in continuous time, but can also be applied to discrete-time networks by omitting the duration. Additionally, the event-based representation enables clear visibility on interactions that include more than two nodes (i.e. hypergraphs and simplicial complexes) [60]. While the simplicity of a 4-column database may be appealing for computational purposes, this representations suffers from potentially poor visualization ability [59].

The graph sequence, or snapshot representation captures the network evolution through a discrete-time sequence of static networks [59], [60]. Mathematically, this is represented by

$$\mathcal{G} = \{G(t_1), G(t_2), \ldots, G(t_{max})\}, \tag{2.13}$$

where $t_{max}$ is the total number of the discrete-time network snapshots, and each snapshot $t_i$ corresponds to the node set, $V$, and link set, $E$, as it existed at $t_i$ [60]. Equation (2.13) is also similar to the multi-slice representation of temporal networks in the sub-field of multilayer networks, but without interlinks between the same nodes in each slice [56]. The snapshot representation is fitting for networks with discrete-time dynamics (such as a time-slotted wireless communication network), and benefits from the ability to apply static network analysis methods to each individual snapshot. For instance, each graph snapshot in Equation (2.13) could be represented by the corresponding adjacency matrix, described in [60] as

$$\mathcal{A} = \{\mathbf{A}(t_1), \mathbf{A}(t_2), \ldots, \mathbf{A}(t_{max})\}. \tag{2.14}$$

Event-based and snapshot representations of temporal networks both consider time-varying

connectivity among a fixed number of nodes. That is, only the link set changes. Dynamic networks, as defined by Holme in [59], consider the interplay of changes in both the link set and node set [59]. While a subtle difference, this class of networks investigates the co-evolution of the network structure dynamics and network function dynamics. However, the rate of change for each set of dynamics must be considered. If the network function changes rapidly with respect to the network structure, the investigation effectively reduces back to an event or snapshot based representation.

**Temporal Walks and Paths**

Unlike connectivity between two nodes in statics networks, temporal network connectivity also considers the causal impact of time. Consider the snapshot representation of an undirected temporal network consisting of two discrete times in Figure 2.11. Over the course of both snapshots, $t_1$ and $t_2$, Node 4 has a path to Node 1, but Node 1 does not have a reciprocal path to Node 4.[3] The path from Node 1 only extends to Node 3. This results from the order in which the links between the nodes occur. This provides an intuitive sense for the concepts of temporal walks and temporal paths.

A temporal walk from node $i$ to node $j$ is a time-ordered sequence of contacts (or events) that allows node $i$ to reach node $j$ [60]. Similar to static network paths, a temporal path is a temporal walk in which no node is visited more than once. Formally, in terms of a snapshot representation, a temporal path is a sequence of $L$ links $[e_1, e_2, ..., e_{L-1}, e_L]$ between two nodes $i$ and $j$ that are traversed over a sequences of snapshots that is also of length $L$, $[t_1, t_2, ..., t_L]$ [60]. The existence of a temporal path implies the existence of a temporal walk, so the remaining discussion will focus solely on temporal paths.

It is important to note that there is some disagreement in the literature about the number of links that can be traversed in a single snapshot. The authors in [63] explicitly allow traversal of multiple links during a single snapshot (similar to the example in Figure 2.11) by defining the sequence of snapshot times as *non-decreasing* ($t_1 \leq t_2 \leq ... \leq t_L$). The authors of [58] and [64] cite [63] in the development of their definition, but tighten the constraint to an *increasing* sequences of snapshots that is also of length $L$, $t_1 < t_2 < ... < t_L$. The stipulation that the time sequence of snapshots be increasing and equal to the number of links traversed

---

[3]This example assumes more than one link can be traversed during a single snapshot.
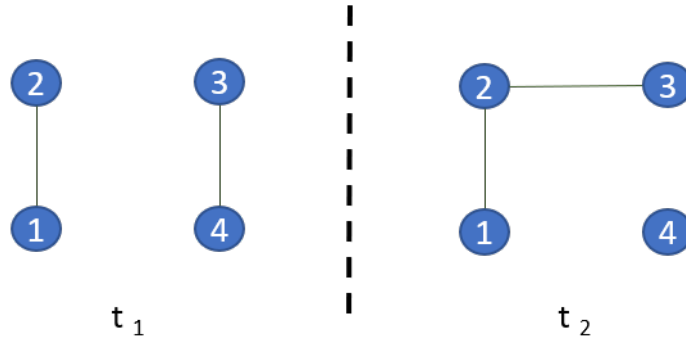
Figure 2.11. Two snapshots of temporal network during which Node 4 is able to reach Node 1, but the reverse path does not exist.

means that only one link can be traversed in each time-step, and that is the definition adopted in this research. The reasoning for this choice and the implications for temporal component definitions are discussed in Chapter 4.

**Temporal Connectedness and Components**

Nicosia et al. [58], [64] employ the concepts of temporal walks and paths to extend the static network notions of connectedness and components to temporal networks, defining node $i$ as *temporally connected* to a node $j$ if there exists a temporal path from $i$ to $j$. Given that temporal connectedness is not a symmetric relation (i.e., a temporal path from node $i$ to $j$ does not imply a temporal path from node $j$ to $i$), they also define strongly and weakly connected nodes for temporal networks.

"Two nodes $i$ and $j$ of a time-varying graph are *strongly connected* if $i$ is temporally connected to $j$ and also $j$ is temporally connected to $i$" [58, p. 4]. Similarly, "two nodes $i$ and $j$ of a time-varying graph are *weakly connected* if $i$ is temporally connected to $j$ and also $j$ is temporally connected to $i$ in the underlying undirected time-varying graph" [58, p. 4]. These definitions are intentionally designed by Nicosia et al. so that two nodes that are strongly or weakly connected in a temporal graph are also strongly or weakly connected in the aggregated static graph [58], [64]. This sets up the following temporal component definitions from [58]:

- *Temporal strongly connected component* (TSCC): A set of nodes of a time-varying

graph in which each node of the set is temporally strongly connected to all other nodes.

- *Temporal weakly connected component* (TWCC): A set of nodes of a time-varying graph in which each node of the set is temporally weakly connected to all other nodes.

Finally, Nicosia et al. define the *affine graph* of a time-varying network as the underlying undirected static graph (with the same nodes as the time-varying graph) associated with a specified time interval (or observation interval) in a time-varying graph such that two nodes $i$ and $j$ are connected only if they are strongly connected in the time-varying graph [58], [64]. Said differently, the affine graph is the static projection of all temporal strongly connected nodes pairs in a time interval.

### 2.3.3 Temporal Network Theory Summary

In summary, the structural connectivity of a network is a common measure of functional network robustness in network science. Connectivity in static networks is represented through node adjacency and the related concepts of walks, paths, and components. These ideas are extended in temporal network theory to account for the causal impacts of time-varying node adjacency, and are applied to the analysis of NOMA network robustness in this research. The following section provides an overview of prior work related to this research.

## 2.4 Related Work

The role of NOMA in NGMA suggests that it is important to have a clear understanding of how overloading variability affects network robustness. As noted in Chapter 1.4, the NOMA literature generally focuses on developing NOMA implementations that achieve IMT-2020 performance objectives. However, several researchers have considered overloading through standard physical and network layer measures of robustness while others have approached NOMA through stochastic geometry (SG). Additionally, the connectivity of wireless networks has been previously considered from through the lenses of static and temporal network theory.

The following sections include expanded versions of previously published material in

B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal mul-

tiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

and

B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917. © 2022 IEEE.

### 2.4.1   NOMA Robustness - Wireless Network Analysis

The link level simulation results presented by Huawei and HiSilicon at the TSG-RAN (Working Group 1, Meeting #85) compare uplink sparse code MA (SCMA) (a code-domain NOMA implementation) to OFDMA by assessing the effect of variable overloading on block error rate at multiple signal-to-noise ratios (SNRs) [66]. The measure of robustness is the achievable decrease in SNR while maintaining a constant spectral efficiency and error rate (i.e., achieving the same throughput for less transmit power, or in worse channel conditions). However, due to the selected codebook structure of SCMA, the investigation is limited to a maximum of 300% overloading and does not consider scenarios in which the ratio of NDs to RBs exceeds the maximum overloading ratio.

In [38], Shirvanimoghaddam et al. consider the robustness of the increased connection density enabled by NOMA. In the context of a mMTC scenario, they propose a random access codebook-based (code-domain) uplink NOMA implementation, and define network stability in terms of the queue size of NDs attempting to transmit in each time slot. The network is stable if the number of NDs attempting to transmit in each time slot is less than or equal the number attempting to transmit in the previous time slot. Based on the probability of random codebook selection by the NDs, they develop a PMF for the overloading ratio of each RB and then derive the maximum packet arrival rate that achieves stability; thus, robustness is measured through maximum supportable packet arrival rate. Similarly, Huang, Wang, and Zhu examine overloading stability in different downlink power-domain NOMA scenarios where stability is defined as the power allocation among NDs that is required for successful signal differentiation [67]. Their study investigates different cases with two NDs,

multiple NDs, and multiple channels, but stops short of analyzing network behavior outside of stability conditions.

This research contributes a graph theoretical model and analytical framework that measures the effect of variable overloading on temporal connectedness between NDs.

## 2.4.2 NOMA Robustness - Stochastic Geometry

The robustness of wireless networks is often approached through SG analysis due to the interference-limited nature of OMA. Stochastic geometry is a well-studied approach that provides robustness metrics such as outage probability, ergodic rate, and bit error probability through the spatial averaging of random point processes that represent the deployment of NDs in a physical space [68]–[70]. Recent SG analysis also includes a temporal aspect through queuing theory to consider maximum traffic density and age of information [71], [72].

Several researchers have investigated NOMA robustness through interference-based stochastic geometry analysis that measures robustness through outage probability and ergodic rate [36], [73]–[77]. These studies often suggest methods to minimize the outage probability of paired NDs, thereby ensuring the requisite overloading ratio to achieve the desired sum rate or spectral efficiency. However, the concept of overloading reduces the primacy of interference-based analysis since NDs in NOMA networks interfere by design.

This research considers NOMA robustness from a complementary temporal random graph theory perspective that measures robustness to variable overloading through temporal component membership and probabilistic connectivity over time.

## 2.4.3 Wireless Network Robustness - Network Science

Recall from Section 2.3 that network connectivity is a common measure of robustness in static/temporal graph-based network science analysis [55], [60]; thus, many researchers use graph theory to investigate the probabilistic connectivity of wireless networks. Related to this research, the authors in [78] employ random graphs to derive the probability of multihop connectivity between NDs in ad hoc wireless sensor networks. The authors in [79] extend the work in [78] using algebraic graph theory to characterize the quality of multihop wireless

network connectivity. However, both works only consider static networks. In [80], Scellato et al. consider the robustness of mobile time-varying networks by measuring the temporal efficiency (which is a normalized measure of the average shortest temporal distance between each pair of nodes in the network) across a temporal snapshot representation of Erdös-Rényi random graphs (Erdös-Rényi graphs are discussed further in Chapter 3). They remove nodes from the network with a fixed probability and quantify the resulting effect on robustness through the change in temporal efficiency.

This work explores the robustness of NOMA wireless networks measured through temporal graph component membership and time-varying probabilistic connectivity, as a function of NOMA overloading.

# CHAPTER 3:
# System Model

This chapter presents a temporal graph model that represents the NOMA wireless network, develops a corresponding temporal network ensemble, and discusses the relationship of the proposed model to existing graph models.

The text includes expanded versions of previously published material in

> B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal multiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

and

> B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917. © 2022 IEEE.

All previously published figures from these two publications are credited with a citation in the caption.

## 3.1   NOMA Graph Model

The relationship between NDs, RBs, and a single BS in a NOMA wireless network can be represented by a graph with ND nodes, RB nodes, one BS node. The graph includes two types of links: connectivity, and dependency. Connectivity links represent the transmission of information signals while dependency links represent the allocation of RBs on which the NDs are dependent for connectivity with the BS. The graph model in Figure 3.1 represents the combination of dependency and directed connectivity relationships over a single MAC frame in a general time division duplex (TDD) NOMA wireless network.

A TDD scheme employs a single carrier frequency for both uplink and downlink transmissions. As a result, uplink and downlink transmission subframes occur sequentially rather than simultaneously.[4] In this NOMA model, each TDD frame is composed of one uplink subframe and one downlink subframe. Network devices transmit to the BS in the uplink subframe, receive from the BS in the downlink subframe.

The NOMA temporal graph model in Figure 3.1 is a snapshot representation of a temporal network where each subframe corresponds to a separate discrete-time realization of the network. Each mixed dependency-connectivity graph realization depicts the connectivity between the NDs and BS (represented by blue lines), and the dependence of connectivity on the allocation of RBs to the NDs by the BS (represented by red lines). The nodes of the graph are the same in each subframe, but the connectivity and dependency links change as RBs are allocated to NDs by the BS. The temporal lengths of the uplink and downlink subframes are denoted by $t_U$ and $t_D$, respectively, and the total length of the TDD frame is $t = t_U + t_D$.

Consistent with the TDD subframe structure, the downlink subframe graph realization only contains *directed* connectivity links *from* the BS *to* the NDs, and the uplink subframe only contains *directed* connectivity links *from* the NDs *to* the BS. Thus, a TDD frame of length $t$ is the smallest time unit in which two different NDs can be connected to each other (unidirectionally or bidirectionally), and the values of $t_U$, $t_D$ and $t$ depend on the chosen NOMA network implementation.

### 3.1.1   Overloading Representation

Overloading, the critical parameter of this inquiry, is represented by the number of links between the each RB node and the NDs. A focused view of the dependency links is provided in Figure 3.2. The dependency link from the RBs to the BS has been removed since the model only considers one BS, so there is no ambiguity regarding the BS from which the RBs are allocated. The dependency subgraph in Figure 3.2 represents the overloading ratio in the degree of each RB node, where the degree of a node, denoted deg($\cdot$), is defined as the number of links connected to it.

---

[4]This assume that the NDs are not capable of full duplex operation (i.e. able to simultaneously transmit and receive RF signals).
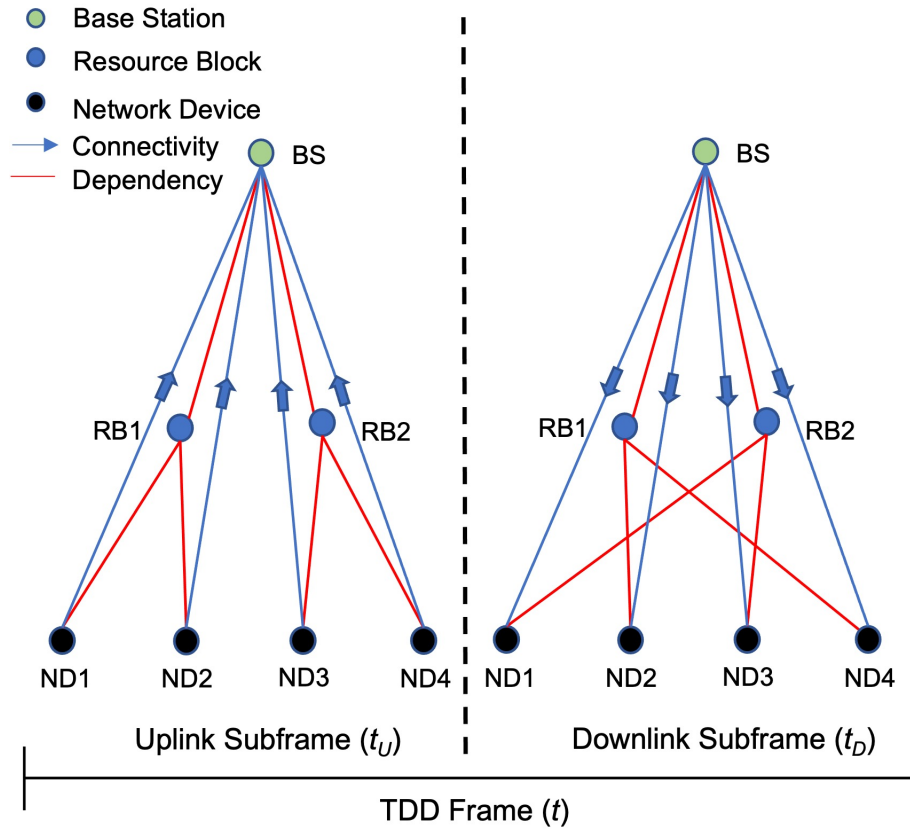
Figure 3.1. Mixed Dependency-Connectivity NOMA graph model in which the overloading ratio is represented by the number of links between the each RB node and the NDs. Adapted from [17] and [65].

Note that the directed connectivity between NDs and the BS can be implied from a combination of the dependency links and the TDD subframe in which they occur. For example, a dependency link that occurs between a RB and a ND in a downlink subframe enables a directed connectivity link *from* the BS *to* the ND. Similarly, a dependency link that occurs between a RB and a ND in an uplink subframe enables a directed connectivity link *to* the BS *from* the ND. The inference of directed connectivity from the TDD frame structure permits a simplification of the mixed dependency-connectivity graph to the dependency subgraph, without loss of connectivity information, in each subframe. Formally, each graph in Figure 3.2 can be specified as a bipartite graph $G(U, V, E)$ where $U$ is the node set of RBs, $V$ is the node set of NDs, and $E$ is the link set representing the allocation of RBs to NDs.
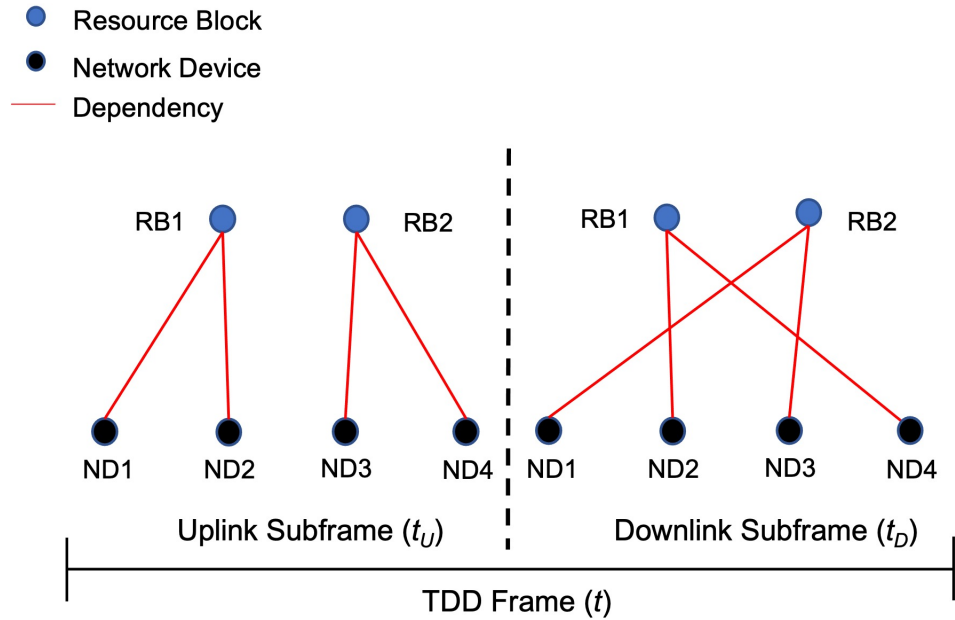
Figure 3.2. Dependency subgraph of the mixed connectivity-dependency NOMA graph model.

## 3.1.2 Model Generality

Is is important to understand that the NOMA graph model in Figure 3.1 is intended to apply broadly across different network implementations. As noted in Section 3.1, the TDD frame represents the smallest time unit in which two different NDs can be connected (unidirectionally or bidirectionally) through the BS. However, the terminology may be different depending on the wireless network standard. For example, in the context of 5G NR, this model could correspond to two consecutive slots with 15 kHz subcarrier spacing (i.e., the subframe length and slot length are both equal to 1 ms) in which all symbols are configured for uplink in the first slot (Slot Format 1) and downlink in the second (Slot Format 0) [12]. However, an NR frame is composed of 10 subframes, thereby allowing five opportunities for connectivity within the length of one frame. Thus, it is imperative that the terms "frame" and "subframe" in this model are interpreted as general constructs that are intended for adaptation to the wireless network transmission structure of the specific NOMA implementation. Those implementations could be different 5G NR subcarrier numerologies, or an entirely separate network standard such as an IEEE 802.11 wireless local area network [81].

The following sections refine this general model into a temporal network ensemble parameterized by the number of NDs, RBs, and the overloading ratio.

## 3.2 Temporal Network Ensemble

In this section, the NOMA network assumptions and representation are defined, and the mathematical formalism of the temporal network ensemble is developed.

### 3.2.1 NOMA Network Assumptions and Representation

Consider a NOMA network composed of one BS, $m = |V|$ NDs, $n = |U|$ RBs, and an overloading ratio $z$, where the parameters $m, n,$ and $z$ are all positive integers specified by the physical NOMA implementation, and $|\cdot|$ denotes set cardinality. Further, the number of NDs exceeds the number of overloaded RBs available, $|V| > z|U|$. Consistent with the motivating examples of massive connectivity and federated learning over a NTN, all NDs have a requirement to transmit in each frame, so all RBs are allocated to a subset of the NDs in each subframe (to send and receive information). This assumption, known as the dominant system model, is often employed in queuing theory network analysis [71], [82], [83].

The overloading ratio is constant across all RBs and each ND is allocated a single RB. Thus,

$$\deg(u) = z, \ \forall u \in U, \tag{3.1}$$

and

$$\deg(v) \in \{0, 1\}, \ \forall v \in V. \tag{3.2}$$

This representation allows each ND to potentially form a link with no more than a single RB, and models the overloading ratio uniformly in the degree of each RB node. Note that setting the maximum degree of ND nodes equal to the value 1, abstracts the implementation of some code-domain NOMA approaches (e.g. SCMA maps six NDs across four RBs [47]). However, a non-uniform RB degree can be introduced to accurately characterize the aggregate allocation of RB to NDs while still restricting the ND degree to $\{0, 1\}$. This will become apparent in the subsequent discussion of the degree vector.

**Degree Vector**

Restricting the degree of each ND, as specified in Equation (3.2), facilitates representation of the connectivity in each subframe graph realization by a labeled binary degree sequence, or a *degree vector* **d**, given by

$$\mathbf{d} = [\deg(v_1), \deg(v_2), \dots, \deg(v_m)]. \tag{3.3}$$

The degree vector indicates which NDs are granted a RB allocation in each subframe (i.e, $\deg(v) = 1$), and which are not (i.e, $\deg(v) = 0$). Note that the *specific* RB node with which each ND forms a link has no bearing on the connectivity that occurs in the subframe. The critical point is to understand which NDs are granted an allocation to *any* RB.

## 3.2.2   Network Ensemble

When paired with subframe type (uplink or downlink) to which it corresponds, the degree vector **d** fully specifies the directed connectivity occurring in that subframe. Thus, the ensemble of all possible network configurations is defined as the set of all unique degree vectors $\mathbf{d} \in D$ of length $m$ where the number of NDs in each vector, $v \in V$, with $\deg(v) = 1$ is equal to $zn \leq m$. That is,

$$\sum_{i=1}^{m} \deg(v_i) = zn < m, \ \forall \mathbf{d} \in D. \tag{3.4}$$

The cardinality (i.e., size) of the network ensemble $D$ is given by the binomial coefficient [84], [85],

$$|D| = \binom{m}{zn} = \frac{m!}{zn!(m - zn)!}, \tag{3.5}$$

where $m = |V|$ is the number of NDs, $n = |U|$ is the number of RBs, and $z$ is the overloading ratio.

A subframe graph realization of the NOMA network is generated by the random selection of a degree vector, $\mathbf{d} \in D$, from the entire ensemble. Each degree vector is only represented once in the ensemble, so random selection from the entire ensemble affords each vector an equal probability of selection, which is the reciprocal of the ensemble size, $|D|^{-1}$. This network ensemble sampling process is repeated, with replacement (i.e. the same degree

vector may be selected more than once during sampling), to generate an arbitrarily long *temporal network* of $T$ frames, denoted $\mathcal{G}(U, V, E, t)$ for $t = 1, 2, \ldots, 2T$. Recall that each frame is composed of one uplink subframe and one downlink subframe; therefore, $2T$ network ensemble samples are required to generate a temporal network of $T$ frames. The implications for RB allocation that result from defining the temporal network ensemble through a random graph generation process are discussed in Chapter 4.

### 3.2.3 Relationship to Existing Models

The NOMA temporal graph model and ensemble incorporate ideas from existing random graph models, interdependent network theory, and temporal network models. Specifically, the model and ensemble are a hybrid combination and adaptation of the well-known Configuration Model for random static networks [86], and the Activity-Driven Model for temporal networks [87].

**Interdependent Network Theory**

Stemming largely from the seminal work of Buldyrev et. al. [88], interdependent network theory studies the effects of dependency relationships on the behavior of interacting networks [89]–[91]. These ideas provide the basis for conceptualizing the role of RBs in NOMA through dependency links. Interdependent networks generally represent the level of dependence between networks through coupling. Similar to the work in [88], the NOMA model represents the dependence of connectivity between the NDs and the BS via a one-to-one coupling between NDs and RBs. This means that without the RB, the ND cannot send or receive information. However, unlike many interdependent network studies, the coupling relationship between the NDs and RBs is not bidirectional. This is because the RBs do not represent a functional network (i.e., a network that is performing a function such as information delivery, or electric power delivery), so they are not dependent on NDs to perform a function. Thus, the dependence coupling is directed from the NDs to the RBs since the allocation of RBs enables functional connectivity. This key difference contributes to the ability to simplify the mixed dependency-connectivity network into the bipartite representation without loss of information.

**The Configuration Model and Related Random Bipartite Models**

The Configuration Model is generative random graph model that is related to the classical random graph models originally developed by Paul Erdös and Aldfréd Rényi. The Erdös-Rényi Model (often referred to as a Poisson Random Graph), denoted $G(n, p)$, specifies an ensemble of random graphs with $n$ nodes and a link placed between each pair of nodes with probability $p$ [92]. This model only allows control over the degree distribution (via $p$) of the network rather than the exact degree composition in the network. Similarly, the random graph model $G(n, m)$ specifies an ensemble of graphs with $n$ nodes and $m$ links but, like $G(n, p)$, it does not provide the ability to specify the degrees of the nodes [55]. Control over node degree is required in the NOMA model to ensure the overloading ratio is accurately represented.

The Configuration Model allows control over the network degree composition by employing a degree sequence for graph generation. A degree sequence is a set of all node degrees in the network. The degree sequence is instantiated in the nodes by assigning a number of "stubs" to each labeled node that is equal to the corresponding degree from the degree sequence. Two stubs are then chosen uniformly at random to connect and form a link. The network ensemble is comprised of all possible "matchings" of stubs given by the degree sequence. That is, the degree of each node (and the corresponding number of stubs) is specified once, and the ensemble is comprised of all the different ways in which these stubs can be connected, including parallel edges and loops that the model removes before presenting the output graph [55], [56], [86].

The use of a degree sequence for graph generation is an important attribute that distinguishes the Configuration Model. However, despite providing control over the node degree, the Configuration Model does not fully meet the requirements of the NOMA temporal graph model. Specifically, once the stubs are assigned to ND nodes and RB nodes, it is clear which NDs will have connectivity during that subframe. The number of different ways in which the ND stubs can be mapped to the RB stubs does not provide any additional information about the connectivity. The NOMA model considers whether ND nodes are able to map to *any* RB by defining the ensemble through all possible allocations of stubs to nodes rather than the ways of connecting stubs that have already been allocated to nodes. Additionally, the Configuration Model degree sequence is defined as a non-increasing order of the labeled node degrees (i.e, $\deg(v_1) \geq \deg(v_2) \geq \ldots \geq \deg(v_m)$) [86]. For a binary degree sequence,

this results in NDs 1 through $zn$ always receiving RB allocations and NDs $zn + 1$ through $m$ receiving none. Thus, the NOMA temporal graph model requires a degree vector that reflects the allocation of RBs to specific NDs, rather than a degree sequence that always arbitrarily assigns the largest degrees to the nodes with the lowest integer labels in the network.

Similar models of random bipartite graphs are also considered in [93] and [94]. In [93], the authors study ensembles of directed labeled bipartite graphs, but with the stipulation that the nodes sets are bigregular. The biregular condition requires that the degree of all nodes in each bipartite node set be equal (but not that the degrees of both node sets are equal since that would make a regular graph). This condition does not fit within the requirements of the NOMA model since the investigation of evolving connectivity in the ND node set considers which NDs are granted a RB allocation in each subframe and which are not. Thus, each degree vector in the NOMA temporal graph ensemble is binary, and not regular.

The work in [94] also considers families of random bipartite graphs generated by a degree sequence. Of those considered, the authors define a case in which an ensemble of binary matrices (corresponding to the adjacency matrix between the two sets of the bipartite graph) is specified by an integer list of column sums. This ensemble of adjacency matrices is similar to the ensemble of degree vectors in the NOMA temporal graph model, but is not quite the same. Only specifying the column sums does not ensure that one ND does not form a link with more than one RB, or that the overloading ratio is equal across all RBs (depending on the organization of rows/columns in the matrix). The NOMA temporal graph model requires specification of both row and column sums in an adjacency matrix representation of the dependency subgraph in Figure 3.2.

**The Activity-Driven Model**

The activity-driven model is a random generative snapshot representation of a temporal network that assigns an activity potential drawn from a probability distribution to each node. The activity potential is a random variable assigned to each node that defines the probability that each node forms $m$ undirected links with $m$ uniformly and randomly selected nodes in the network in each snapshot [60], [87]. Nodes that are inactive in a snapshot may still form a link with another node if they are randomly selected for link creation by an active node. This random generative temporal network model has many similarities to the NOMA

45

temporal graph model. Specifically, the probability of RB allocation to each ND (which will be discussed further in Section 4.1) is similar to the activity potential. Additionally, like the activity potential, the RB allocation probability is fixed for all subframe snapshots in the NOMA temporal network. However, in contrast to the activity-driven model, the degree vectors (once sampled) specify the connectivity that will occur in each subframe rather than assigning a probability of connectivity to each ND. Finally, each ND node that is granted a RB allocation can form a link with any RB node, but this is a dependency relationship, not a connectivity relationship. All links that represent connectivity from the NDs connect through the BS.

# CHAPTER 4:
# Stochastic Temporal Component Framework

This chapter considers the NOMA temporal graph model as a random process and develops the implications for RB allocation probability and temporal component membership.

The text includes expanded versions of previously published material in

> B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal multiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

and

> B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917. © 2022 IEEE.

All previously published figures from these two publications are credited with a citation in the caption.

## 4.1   Network Sequence as a Bernoulli Random Process

Recall from Equations (3.2), (3.3), and (3.4) in Chapter 3 that, given a set of NOMA network parameters ($m$, $n$, and $z$), the allocation of RBs to NDs in each subframe graph realization is represented by a $1 \times m$ binary degree vector, $\mathbf{d}$. Each unique degree vector $\mathbf{d} \in D$ contains $zn$ ones and $m - zn$ zeros, where ones represent the allocation of a RB to a ND; thus, the probability of RB allocation to an individual ND in a graph realization generated from any degree vector $\mathbf{d} \in D$ is a Bernoulli random variable $X$ with probability mass function

(PMF) given by

$$f_X[x] = \begin{cases} p & x = 1 \\ 1 - p & x = 0 \\ 0 & \text{otherwise} \end{cases}, \tag{4.1}$$

where

$$p = \frac{zn}{m}, \tag{4.2}$$

and the expectation and variance of $X$ are $\mathbb{E}\{X\} = p$ and $\mathbb{V}\{X\} = p(1-p)$, respectively [84], [85].

As discussed in Chapter 3.2.2, the NOMA temporal graph model, $\mathcal{G}(U, V, E, t)$, is generated from the random sampling of degree vectors from the network ensemble, each of which has an equal probability of selection. Since each unique binary degree vector has $zn$ ones and $m - zn$ zeros, and the selection of each degree vector is equally likely, the probability of RB allocation to a single ND over successive subframe graph realizations are independent and identically distributed (IID) Bernoulli trials. Thus, a sequence of $N$ degree vectors randomly sampled from the network ensemble constitute a Bernoulli random process with $N$ realizations occurring across all NDs (i.e., all $v \in V$) with a PMF for RB allocation in each realization given by Equation 4.1. In matrix form, this is an $N \times m$ matrix, where the rows are the randomly selected degree vectors in ascending temporal order from $t_1$ to $t_N$, the columns are the NDs $v_1$ to $v_m$, and the $(i, j)$ elements are the degree of node $v_j$ in network realization $t_i$. Thus, each column represents the Bernoulli random process of each ND. A visual representation is shown in Figure 4.1. The green rectangle highlights a single network realization occurring at $t_i$, and the blue rectangle highlights a Bernoulli process of RB allocation for a single ND, $v_j$.

### 4.1.1 Probability of Resource Block Allocation Over time

Characterizing the network sequence of each ND as a Bernoulli random process provides a basis to consider the probability that a ND receives $k$ RB allocations in a temporal network of $N$ realizations. The statistical independence of each RB allocation event means the joint probability of any specific RB allocation sequence is the product all independent probabilities. If the specific order of the RB allocations is not relevant, then the joint probability is multiplied by the number of combinations that could result from a specific number of RB

Figure 4.1. Matrix representation of $N$ randomly degree vectors (rows) across $m$ NDs (columns). Adapted from [17].

allocations and "nulls." This combined product is the well-known binomial distribution. Thus, the probability of $k$ RB allocations in a temporal network of $N$ realizations is a binomial random variable with probability mass function

$$
f_K[k] = \begin{cases} \binom{N}{k} p^k (1-p)^{N-k} & 0 \le k \le N \\ 0 & \text{otherwise} \end{cases}, \tag{4.3}
$$

where $p$ is given by Equation 4.2 [84], [85]. This distribution can also be used to determine the probability that the number of RB allocations falls within a range that is bounded by the temporal network length. However, this calculation may become burdensome since it will require the sum of the binomial random variable evaluated at many values.

Similar to an example from [84], consider the probability that one ND is granted more than

70 RB allocations in a temporal network of 100 realizations, given by

$$\mathbb{P}\{K > 70\} = \sum_{k=71}^{100} \binom{100}{k} p^k (1-p)^{100-k}, \tag{4.4}$$

where $K$ is the random variable representing sum of RB allocations in the temporal network for a single ND,

$$K = \sum_{i=1}^{100} X_i, \tag{4.5}$$

each $X_i$ is a Bernoulli random variable with PMF given by (4.1), and $\mathbb{P}\{\cdot\}$ denotes probability. Notice that this is a sum of IID Bernoulli random variables and, by the Central Limit Theorem, a sum of IID random variables converges to a Gaussian (normal) distribution with mean $n\mu_K$ and variance $n\sigma_K^2$ (where $n$ is the number of IID random variable realizations in the sum) [84], [85]. Thus, a Gaussian random variable normalized by $n\mu_K$ and $n\sigma_K^2$ is given by

$$Z = \frac{K - n\mu_U}{\sqrt{n}\sigma_K} \tag{4.6}$$

and, using the expectation and variance of a Bernoulli random variable defined in Section 4.1, approximates the probability in Equation (4.4) by

$$\mathbb{P}\{K > 70\} = \mathbb{P}\left\{Z > \frac{70 - np}{\sqrt{np(1-p)}}\right\}, \tag{4.7}$$

which is a less involved calculation that can be accomplished using a Q-function lookup table [84], [85].

### 4.1.2 Resource Block Allocation Model Applicability

Most modern mobile wireless communications networks allocate RBs based on the current channel conditions between the NDs and the BS [6], [7], [12]. For example, during a certain time period, ND1 might have a high SNR with the BS on frequency $f_1$, and a low SNR with the BS on frequency $f_2$. During the same time period, ND2 might have a high SNR with the BS on frequency $f_2$, and a low SNR on frequency $f_1$. Assuming the BS has data to transmit to both NDs, ND1 will be scheduled on frequency $f_1$, and ND2 will be scheduled on frequency $f_2$. This scheduling decision maximizes the likelihood of error free transmission

50

to both NDs at the highest possible throughput (see Equation 2.1 in Chapter 2). However, if a subset of NDs experiences poor channel conditions across large portions of the available frequency bandwidth for an extended period of time, they may be underserved by the BS resulting in an uneven (or unfair) allocation of RBs.

The random RB allocation model resulting from the random sampling process that generates the NOMA temporal network is a departure from channel-based dynamic scheduling. This model assumes that the BS and NDs can achieve overloading (defined by the underlying physical NOMA implementation) across a random subset of NDs in each subframe. Though the random graph model was selected to facilitate tractable analysis that provides initial insights and forms the foundation for increasingly nuanced investigations, the Bernoulli random process also provides a maximum fairness approach to RB allocation. This is because each ND receives a nearly equal number of RB allocations as the temporal network length becomes large [85].

While channel-based dynamic scheduling will continue to play a prominent role in the eMBB 5G use case, the maximum fairness RB allocation model is applicable to mMTC and URLLC use cases such as massive connectivity and federated learning applications [31], [42], [95]. In these cases, lower data rate requirements may be more easily met across the available frequency spectrum (due to lower required SNR), thereby facilitating the use of a random RB allocation model.

## 4.2   Temporal Component Membership

Network device membership in a temporal component can be defined in terms of connectivity directly with the BS, or connectivity with the other NDs in the network through the BS. This research focuses on the latter definition, and treats the BS as a relay rather than a gateway, as might be the case in 5G NTNs [26], [27]. Additionally, as discussed in Chapter 2, this research adopts the temporal path definition in which only one link can be traversed during each snapshot of the temporal network. This definition is consistent with the TDD MAC frame structure and centralized network topology of the NOMA system model defined in Chapter 3. Each snapshot (i.e., graph realization) corresponds to an uplink or downlink subframe in which each ND only has an opportunity (contingent upon RB allocation) for directed connectivity with the BS. Thus, only one link is traversed during

each snapshot.

The decisions to define temporal component membership based on ND-to-ND connectivity through the BS and restrict temporal paths to single-link traversal in each snapshot render the TWCC and TSCC indistinguishable over the course of one frame as defined in [58] and [64]. Thus, we *amend* the definition of weak temporal connectedness as follows:

- *Weak Temporal Connectedness*: two nodes $i$ and $j$ of a time-varying graph are weakly connected if, either $i$ is temporally connected to $j$, *or* $j$ is temporally connected to $i$, in the underlying undirected time-varying graph.

The TWCC definition remains as stated in Section 2.3.2, and the temporal components correspond to three types of connectivity:

- TSCC: Bidirectional Connectivity
- TWCC (only): Unidirectional Connectivity
- Isolated: No Connectivity.

The following sections discuss the corresponding probabilities for each of these components.

## 4.3   Individual Probability of Temporal Component Membership

Defining temporal component membership based on connectivity between NDs through the BS requires two subframes to evaluate which NDs belong to which temporal components. Given a NOMA wireless network of $m$ NDs, $n$ RBs, and an overloading ratio $z$, the probability of temporal component membership is described by the temporal component event tree in Figure 4.2.

This event tree shows the RB allocation sample space for a single ND over two subframes, $t_U$ and $t_D$ (uplink and downlink). Resource block allocation is designated by the value 1 with probability $p$, and a lack of RB allocation (null) is designated by a 0 with probability $1 - p$. Two subframes yield a binary sequence of length two with four possible outcomes, $\{11, 10, 01, 00\}$. Strong temporal connectedness requires bidirectional temporal paths between two NDs in the directed graph, which corresponds to a RB allocation in both
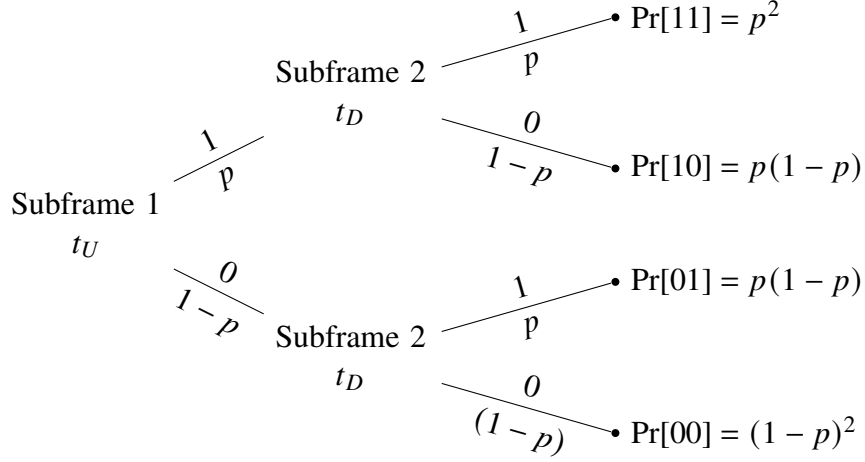
Figure 4.2. Temporal component event tree illustrating the mutually exclusive and exhaustive RB allocation probability space over the course of one frame. Adapted from [17].

subframes, e.g. {11}. Weak temporal connectedness requires a unidirectional temporal path between two NDs in the underlying undirected time-varying graph, which corresponds to one RB allocation in either subframe, either {10} or {01}. Finally, an isolated ND corresponds to zero RB allocations in either subframe, {00}. These mutually exclusive and exhaustive event probabilities are the result of an IID random process; thus, each event probability and its complement are used to define a new Bernoulli random process that characterizes the individual ND membership probability for the corresponding temporal component.

Let the probability that any ND, $v \in V$, is a member of the TSCC, TWCC (only[5]), or isolated during any frame in a network sequence be denoted by $p_s$, $p_w$, and $p_i$, respectively, where

$$p_s = p^2, \; p_w = 2p(1-p), \; p_i = (1-p)^2, \text{ and}$$

$$p_s + p_w + p_i = 1.$$

The probability that a ND is a member of the TSCC in any arbitrary frame is given by substituting $p_s$ for $p$ in Equation 4.1. Similarly, the probability that a ND is a member of

---

[5]All NDs in the TSCC are also part of the TWCC, but this characterization considers the NDs that are only part of the TWCC.

the TSCC for $k$ frames in a network sequence of $N$ frames is given by substituting $p_s$ for $p$ in Equation 4.3. This same approach also applies to the event probabilities for the TWCC and isolated events.

## 4.4 Joint Probability of Temporal Component Membership

The analysis of joint ND temporal component membership is different than for an individual ND, and can be characterized by the joint probability mass function defined in [84] as

$$f_{K_1, K_2}[k_1, k_2] \overset{\text{def}}{=} \mathbb{P}\{K_1 = k_1, K_2 = k_2\}. \tag{4.8}$$

Recall from Section 4.1 that the probability of RB allocation for one ND can be analyzed over an arbitrary number of network realizations from $(1, ..., N)$, where each realization is an IID Bernoulli random variable, and the joint probability of any specific sequence of events is the product of the individual event probabilities. However, each network realization is the result of a randomly selected degree vector with fixed values that are dependent on the network parameters, so the probability of RB allocation for two or more NDs are not independent within a *single* network realization. Thus, the joint probabilities cannot be found by the product of the marginal probabilities. Rather, the joint PMF must be calculated from the product of the conditional and marginal PMFs

$$f_{K_1, K_2}[k_1, k_2] = f_{K_1 \mid K_2}[k_1 \mid k_2]\, f_{K_2}[k_2], \tag{4.9}$$

where the conditional PMF is defined as

$$f_{K_1 \mid K_2}[k_1 \mid k_2] \overset{\text{def}}{=} \mathbb{P}\{K_1 = k_1 \mid K_2 = k_2\}, \tag{4.10}$$

and $[k_i \mid k_j]$ denotes that event $k_i$ is conditioned on the occurrence of event $k_j$ [84].

### 4.4.1 Example Joint Probability Calculation

For example, consider the joint probability that two NDs, $u_i$ and $u_j$ (for $i \neq j$), receive exactly one RB allocation between them in a NOMA wireless network with two RBs, six NDs, and an overloading ratio $z = 2$. The allocation of one RB can occur as $\deg u_i \rightarrow 1$ and $\deg u_j \rightarrow 0$ or $\deg u_i \rightarrow 0$ and $\deg u_j \rightarrow 1$, so the joint probability of receiving one RB allocation between them is

$$\mathbb{P}\left\{\deg u_i + \deg u_j = 1 \mid i \neq j\right\} = f_{U_i,U_j}[1,0] + f_{U_i,U_j}[0,1]. \tag{4.11}$$

Substituting the two terms on the right side of Equation (4.11) into Equation (4.9) and using the NOMA network parameters to determine the conditional and marginal probabilities gives

$$f_{U_1,U_2}[1,0] = f_{U_1 \mid U_2}[1 \mid 0]\, f_{U_2}[0] = \left(\frac{4}{5}\right)\left(\frac{1}{3}\right) = \frac{4}{15}, \tag{4.12}$$

and

$$f_{U_1,U_2}[0,1] = f_{U_1 \mid U_2}[0 \mid 1]\, f_{U_2}[1] = \left(\frac{2}{5}\right)\left(\frac{2}{3}\right) = \frac{4}{15}. \tag{4.13}$$

Substituting the results of Equations (4.12) and (4.13) back into Equation (4.11) gives the total probability of any two randomly selected NDs receiving a single RB in one network realization,

$$\mathbb{P}\left\{\deg u_i + \deg u_j = 1 \mid i \neq j\right\} = \frac{4}{15} + \frac{4}{15} = \frac{8}{15}. \tag{4.14}$$

Notice the two joint probabilities $f_{U_1,U_2}[1,0]$ and $f_{U_1,U_2}[0,1]$ are not equal to the product of the marginal probabilities where

$$f_{U_i}[1] = \frac{2}{3} \text{ and } f_{U_i}[0] = \frac{1}{3},$$

which confirms they are not independent.

The joint probability of a single RB allocation among two randomly selected NDs can also be calculated from a combinatorial perspective by selecting $\binom{4}{1}$ RB allocations, and $\binom{2}{1}$ nulls, from a total of $\binom{6}{2}$ NDs. Employing the product rule to determine the total number ways to select one RB allocation and one null, and dividing by the total number of ways to

select two NDs gives

$$\frac{\binom{4}{1}\binom{2}{1}}{\binom{6}{2}} = \frac{8}{15}. \tag{4.15}$$

### 4.4.2 Hypergeometric PMF

Notice that the combinatorial method of calculating the joint probability in Equation (4.15) is the same method used to determine the probability of drawing a sample of $n$ items containing $k$ successes, without replacement, from a population $N$ with a known number of successes $M$ [85]. This is the well-known hypergeometric distribution with PMF given in [85] by

$$f_K[k] = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}. \tag{4.16}$$

In the NOMA temporal network model, the successes are those NDs in the sample that receive a RB allocation. Substituting the NOMA parameters into Equation 4.16 gives

$$f_K[k] = \frac{\binom{nz}{k}\binom{m-nz}{m_s-k}}{\binom{m}{m_s}}, \tag{4.17}$$

where $m_s$ is the number of NDs sampled, $m$ is the total number of NDs in the network, the quantity of known successes is the number of RBs multiplied by the overloading ratio, $nz$, and $k$ is the number of RBs allocated to the unlabeled NDs in the sample.

**Unlabeled NDs**

The hypergeometric distribution provides the joint probability of any RB allocation among randomly selected unlabeled distinct NDs from a single network realization. Continuing with the example NOMA network paramaterized by two RBs, six NDs, and an overloading ratio $z = 2$, the joint probability of RB allocation among two NDs is

$$f_K[k] = \frac{\binom{4}{k}\binom{2}{2-k}}{\binom{6}{2}}, \tag{4.18}$$

which results in probabilities of $\{\frac{1}{15}, \frac{8}{15}, \frac{6}{15}\}$ for $k \in \{0, 1, 2\}$, respectively. These values correspond to the probability that any two randomly selected unlabeled distinct NDs have

0, 1, or 2 RBs allocated to them in a single network realization (i.e. subframe).

**Labeled NDs**

Determining the joint probability of RB allocation among any *labeled* combination of NDs requires equally distributing the unlabeled joint probability over the number of possible unique RB allocations to the labeled NDs, where the number of possible RB combinations is given by the binomial coefficient $\binom{m_s}{k}$. Thus, the labeled probabilities are given by

$$f_K[k]_{\text{labeled}} = \frac{f_K[k]}{\binom{m_s}{k}}. \tag{4.19}$$

**Temporal Component Membership**

Given the relationship between RB allocation and temporal component membership, the hypergeometric distribution can be used to stochastically characterize the temporal component membership of specific or randomly selected groups of NDs (i.e., labeled or unlabeled) ranging in size from two to the total number of NDs in the network (however, as the number of NDs becomes large, this may converge to the binomial distribution if $nz$ does not increase proportionally [96]).

Once the joint probability has been determined, it is fixed over all frames in the network sequence and each frame is an IID realization. Thus, after defining an exhaustive and mutually exclusive joint probability space (parameterized by the number of NDs, RBs, and overloading ratio), the joint temporal component membership probability can be determined over an arbitrary number of frames using the same approach discussed in Section 4.3.

## 4.5   Summary

In this chapter, the conceptualization of the NOMA temporal graph model as a Bernoulli random process was introduced. The implications of this idea were extended to the probability of RB allocation to NDs in each subframe, and formed the basis of the stochastic temporal component framework. This framework provides probabilistic tools to understand temporal component membership in terms of NOMA network parameters for individual NDs and groups of NDs. Thus, for a constant number of NDs and RBs, the probabilities

of membership in the TSCC and TWCC are measures of NOMA network robustness as a function of overloading. However, the framework does not consider the impact of messages that are sent in one frame, but not received until a later frame. The impact of "queued" messages on NOMA network robustness is considered through temporal connectivity in Chapter 5.

# CHAPTER 5:
# Temporal Connectivity

This chapter introduces the mathematical model of temporal connectivity that corresponds to the NOMA temporal graph ensemble, and develops expressions for the following measures of NOMA robustness:

- Probability of Temporal Connectivity in Each Frame
- Probability of Time to Initial Unidirectional Connectivity
- Probability of Time Window Between Unidirectional Connectivity
- Probability of Time to Initial Bidirectional Connectivity
- Probability of Minimum Time to Complete Bidirectional Connectivity.

The text includes expanded versions of previously published material in

> B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal multiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

and

> B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917. © 2022 IEEE.

All previously published figures from these two publications are credited with a citation in the caption.

## 5.1   Mathematical Model of Connectivity

The mathematical model of connectivity considers temporal connectivity between NDs through potential connectivity matrices, a frame connectivity matrix, and a queue matrix.

**Potential Connectivity Matrices**

Potential connectivity matrices are directed adjacency matrices that represent the possible connections between NDs enabled by the allocation of RBs in the uplink or downlink subframes. Each potential connectivity matrix is a function of RB allocation to the NDs resulting from the random sampling of degree vectors from the network ensemble. Specifically, given a binary $1 \times m$ degree vector $\mathbf{d}$, the $m \times m$ uplink potential connectivity matrix $\mathbf{U}$ is given by

$$\mathbf{U} = \mathrm{diag}(\mathbf{d})(\mathbf{1} - \mathbf{I}), \tag{5.1}$$

where $\mathbf{1}$ is $m \times m$ uniform matrix, $\mathbf{I}$ is $m \times m$ identity matrix, and $\mathrm{diag}(\cdot)$ is the diagonal operator.[6] Since $\mathbf{U}$ is a directed adjacency matrix, the binary values of the $(i, j)$ elements represent whether ND $i$ has the potential to *send* a message to ND $j$. For example, the degree vector $\mathbf{d} = [0\ 0\ 1\ 1]$ has an uplink potential connectivity matrix,

$$\mathbf{U} = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{Bmatrix}. \tag{5.2}$$

The matrix in Equation (5.2) shows that, in a network of four NDs, the third and fourth NDs have the *potential* to send a message to all other NDs in the network during that frame (depending on which NDs are granted a RB allocation in the downlink subframe). Note that this representation excludes potential connectivity from a ND to itself (i.e., no self-links) because a ND does not send information to itself via the BS.

Similarly, $\mathbf{D}$ is the $m \times m$ downlink potential connectivity matrix given by

$$\mathbf{D} = (\mathbf{1} - \mathbf{I})\mathrm{diag}(\mathbf{d}). \tag{5.3}$$

Like $\mathbf{U}$, the binary values of the $(i, j)$ elements of $\mathbf{D}$ represent whether ND $j$ has the potential to *receive* a message from ND $i$. Extending the same example, the degree vector

---

[6]The diagonal operator places a $1 \times m$ vector along the main diagonal of an $m \times m$ matrix.

$\mathbf{d} = [0\ 0\ 1\ 1]$ has a downlink potential connectivity matrix,

$$\mathbf{D} = \begin{Bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{Bmatrix}. \tag{5.4}$$

The matrix in Equation (5.4) shows that, in a network of four NDs, the third and fourth NDs have the *potential* to receive a message from all other NDs in the network during that frame (depending on which NDs are granted a RB allocation in the uplink subframe).

**Frame Connectivity and Queue Matrices**

The frame connectivity matrix $\mathbf{F}^t$ is a directed adjacency matrix that shows the directed connectivity achieved between NDs during frame $t$. Two NDs, $i$ and $j$, may establish directed connectivity in an arbitrary frame, $t$, in two ways. First, ND $i$ may be granted an uplink RB allocation and ND $j$ may be granted a downlink RB allocation, both in frame $t$. Second, ND $i$ may be granted an uplink RB allocation in frame $\tau < t$, and ND $j$ is not granted a downlink RB allocation until frame $t$. In this case, the message from ND $i$ is queued at the BS until ND $j$ is granted a downlink RB allocation in frame $t$. The frame connectivity matrix, $\mathbf{F}^t$, is an $m \times m$ directed adjacency matrix that describes the connectivity in frame $t$ resulting from either of these possibilities. The $(i, j)$ element of $\mathbf{F}^t$, denoted $F_{ij}^t$, is equal to one if ND $i$ is connected to ND $j$ in frame $t$, and zero otherwise.

The queue matrix, $\mathbf{Q}^t$, is an $m \times m$ matrix that accounts for "undelivered" messages. The $(i, j)$ element of $\mathbf{Q}^t$, denoted $Q_{ij}^t$, represents the number of uplink transmissions from ND $i$ that have not been received via downlink receptions at ND $j$. Thus, for $t \geq 1$, $\mathbf{F}^t$ and $\mathbf{Q}^t$ are defined recursively as

$$\mathbf{F}^t = \mathbb{1}_{\mathbb{Z}^+} \left( \mathbf{U} + \mathbf{Q}^{t-1} \right) \odot \mathbf{D} \tag{5.5}$$

and

$$\mathbf{Q}^t = \mathbf{U} + \mathbf{Q}^{t-1} - \mathbf{F}^t \tag{5.6}$$

where $\odot$ denotes the Hadamard (element-wise) product defined in [97], $\mathbf{Q}^0$ is an $m \times m$ null (or zero) matrix $\mathbf{0}$ from [98] (representing empty queues when the network sequence begins), and $\mathbb{1}_{\mathbb{Z}^+}(\cdot)$ is the indicator function (or characteristic function on a set) from [99]

for the set of positive integers, defined as

$$\mathbb{1}_{\mathbb{Z}^+}(\mathbf{X}) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } x_{ij} \in \mathbb{Z}^+ \\ 0 & \text{otherwise} \end{cases}. \tag{5.7}$$

This mathematical conceptualization captures the directed connectivity between NDs that results from RB allocation in that frame as well as queued messages from previous frames. The representation also makes the physical assumptions that the message length, transmission rate, and transmission time support the transmission/reception of up to $m-1$ messages in the uplink and downlink subframes.

**Example**

Consider an example frame $t = 1$ in a network with four NDs. Let the degree vectors $\mathbf{d}_{UL} = [0\ 0\ 1\ 1]$ and $\mathbf{d}_{DL} = [1\ 0\ 1\ 0]$ represent the uplink and downlink RB allocations in frame $t = 1$, respectively. Employing Equations (5.1), (5.3), and (5.5) together with the definition that $\mathbf{Q}^0 = \mathbf{0}$ gives

$$\mathbf{F}^1 = \mathbb{1}_{\mathbb{Z}^+} \left\{ \left( \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right) \odot \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \right\} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}. \tag{5.8}$$

In this frame, NDs three and four were able to send a message in the uplink, and NDs one and three were able to receive a message in the downlink. The resulting frame connectivity matrix in Equation (5.8) reflects the unidirectional connectivity from NDs three and four to ND one, and from ND four to ND three. Notice that there is an odd number of non-zero elements in $\mathbf{F}^1$ because ND three received both an uplink and downlink allocation, but self-links are not permitted because the ND does not send a message to itself. Extending this example to the queue matrix defined in Equation (5.6) gives

$$\mathbf{Q}^1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{5.9}$$

The resulting matrix in Equation (5.9) shows the queued messages from ND three to NDs two and four, and from ND four to ND two. This matrix will be included in the connectivity for frame $t = 2$ through addition with the uplink potential connectivity matrix that results from the uplink RB allocation in frame $t = 2$.

## 5.2 Probability of Temporal Connectivity in Each Frame

The frame connectivity matrix defined in Equation (5.5) describes connectivity between ND $i$ and ND $j$ in frame $t$ through a logical OR operation followed by a logical AND operation. In this section, these logical operations are represented stochastically by considering the probability of temporal connectivity from ND $i$ to ND $j$ in frame $t$ as a function of the random processes that underlie each realization of $\mathbf{U}$, $\mathbf{D}$, and $\mathbf{Q}^{t-1}$.

### 5.2.1 Potential Connectivity Matrices as Bernoulli Trials

The random allocation of RBs in the uplink and downlink subframes described in Chapter 4.1 characterizes the $(i, j)$ elements of $\mathbf{U}$ and $\mathbf{D}$, denoted $U_{ij}$ and $D_{ij}$, as Bernoulli random variables with $p$ defined in Equation (4.2), for $i \neq j$. Thus,

$$\mathbb{P}\left\{U_{ij} = 1\right\} = \mathbb{P}\left\{D_{ij} = 1\right\} = p \tag{5.10}$$

and

$$\mathbb{P}\left\{U_{ij} = 0\right\} = \mathbb{P}\left\{D_{ij} = 0\right\} = q \tag{5.11}$$

where $q = 1 - p$. Note that $\mathbf{U}$ and $\mathbf{D}$ have no reference to the frame in which they occur. Each realization of both matrices is statistically independent due to their construction from the random selection of degree vectors from the network ensemble.

### 5.2.2 Queue Dynamics as a Markov Process

As briefly discussed in Section 5.1, the value of $Q_{ij}^t$ represents the number of transmissions from ND $i$ that have not yet been received by ND $j$. The network sequence begins with empty queues between all NDs, so the value for each queue must be a non-negative integer, $Q_{ij}^t \geq 0$, for each frame $t \geq 0$. As the network sequence progresses through time, $Q_{ij}^t$ changes based on the uplink RB allocations to ND $i$, and the downlink RB allocations to

ND $j$. For example, if ND $i$ is granted an uplink RB allocation in frame $t$, and ND $j$ is not granted a downlink RB allocation in frame $t$, then $Q_{ij}^t = Q_{ij}^{t-1} + 1$. Similarly, if ND $i$ is not granted an uplink RB allocation in frame $t$, and ND $j$ is granted a downlink RB allocation in frame $t$, then $Q_{ij}^t = Q_{ij}^{t-1} - 1$ for $Q_{ij}^{t-1} > 0$. If $Q_{ij}^{t-1} = 0$, then $Q_{ij}^t$ remains zero. Thus, $Q_{ij}^t$ (for $i \neq j$) evolves as a discrete Markov process. Specifically, the queue dynamics can be modeled as a one-dimensional random walk with one reflecting barrier [85]. A random walk of this type is depicted in Figure 5.1.

The states of the walk, $(e_0, e_1, \ldots, e_l, \ldots)$, are the values of $Q_{ij}^t$. State $e_0$ is the state in which $Q_{ij}^t = 0$, and the queue from ND $i$ to ND $j$ is empty. Similarly, state $e_l$ is the state in which $Q_{ij}^t = l$, and the queue from ND $i$ to ND $j$ contains $l$ messages.
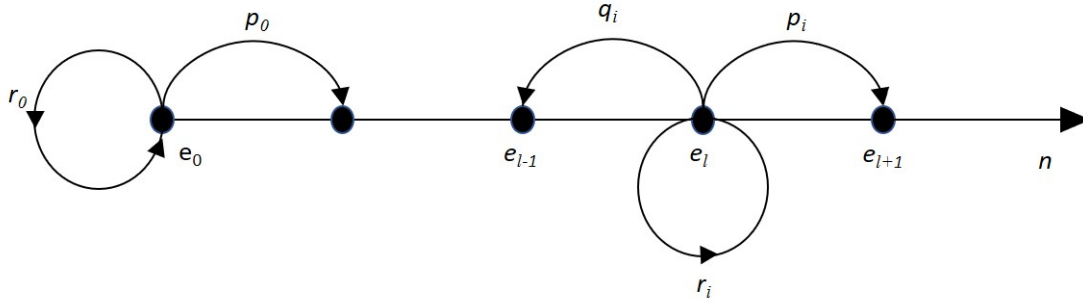


Figure 5.1. One-dimensional random walk with one reflecting barrier. Adapted from [85].

The walk dynamics are governed by state transition probabilities. These probabilities describe the likelihood that the "walker" (i.e., the queue) will transition from the current state to a different state, or remain in the same state during a transition opportunity. The probabilities are represented in a state transition matrix $\mathbf{P}$, defined in [84] as

$$\mathbf{P} = \left\{ \begin{matrix} P_{00} & P_{01} & P_{02} & P_{03} & \ldots \\ P_{10} & P_{11} & P_{12} & P_{13} & \ldots \\ P_{20} & P_{21} & P_{22} & P_{23} & \ldots \\ P_{30} & P_{31} & P_{32} & P_{33} & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{matrix} \right\}.$$

64

Each $(i, j)$ element, denoted $P_{ij}$, is the probability of transition from state $i$ to state $j$ [84]. Notice that all states have three state transition probabilities except for state $e_0$. The message queue cannot be negative, so there are only two state transition probabilities if the current state is an empty queue. This is the "reflecting barrier" of the one-dimensional random walk [85]. Additionally, the sum of the state transition probabilities from each state are always equal to 1 (i.e., each row in **P** sums to 1).

In the temporal NOMA network model, state transitions occur in each frame. The next section discusses the transition probabilities.

**State Transition Probabilities**

The random walk is governed by state transition probabilities that are defined in terms of the RB allocation probability from Equation (4.2). When $Q_{ij}^{t-1} = 0$ and the queue is in state $e_0$, the queue will only increase by one message if ND $i$ is granted an uplink RB allocation with probability $p$, and RB $j$ is not granted a downlink RB allocation with probability $q = 1 - p$. All other possibilities return the empty queue to the same state, giving

$$p_0 = pq \tag{5.12}$$

and

$$r_0 = 1 - p_0, \tag{5.13}$$

where, from [85], $p_0$ is the transition probability to state $e_1$, $r_0$ is the transition probability to state $e_0$, and

$$p_0 + r_0 = 1. \tag{5.14}$$

When $Q_{ij}^{t-1} > 0$ and the queue is in state $e_l$ (for $l > 0$), $Q_{ij}^t$ can transition to state $e_{l-1}$, $e_l$, or , $e_{l+1}$. In this case, the queue is not empty, so it can increase by one message with probability $p_l$, remain the same with probability $r_l$, or decrease by one message with probability $q_l$, where, from [85]

$$p_l + r_l + q_l = 1. \tag{5.15}$$

An increase in the queue occurs when ND $i$ is granted an uplink RB allocation and ND $j$ is not granted a downlink RB allocation. The queue remains the same when ND $i$ and ND $j$ are both granted uplink and downlink RB allocations, respectively, or when neither do. The

queue decreases when ND $i$ is not granted an uplink RB allocation and RB $j$ is granted a downlink RB allocation. The probabilities for these events are given by

$$p_l = pq, \tag{5.16}$$

$$r_l = p^2 + q^2, \tag{5.17}$$

and

$$q_l = qp, \tag{5.18}$$

respectively. Thus, the state transition matrix, $\mathbf{S}$, is defined as

$$\mathbf{S} = \begin{Bmatrix} r_0 & p_0 & 0 & 0 & 0 & \dots \\ q_1 & r_1 & p_1 & 0 & 0 & \dots \\ 0 & q_2 & r_2 & p_2 & 0 & \dots \\ 0 & 0 & q_3 & r_3 & p_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{Bmatrix} \tag{5.19}$$

where the $(i, j)$ elements of $\mathbf{S}$, denoted $S_{ij}$, are the transition probabilities from state $i$ to state $j$ [85].

**Higher Order Transition Probabilities**

The state transition matrix in Equation (5.19) only provides the queue transition probabilities for one frame. However, the Chapman-Kolmogorov equation states that the probability of transitioning from state $i$ to state $j$ in $\ell$ transitions is found by raising the state transition matrix to the $\ell^{th}$ power [84], [85]. This result is applied by raising $\mathbf{S}$ to the $\ell^{th}$ power, denoted $\mathbf{S}^{(\ell)}$, to determine the transition probabilities from state $i$ to state $j$ in $\ell$ frames. Thus, given a vector of queue state probabilities in frame $t$, denoted $\mathbf{s}^t = [s_0^t,\ s_1^t,\ s_2^t, \dots, s_l^t, \dots]$, the probability that $Q_{ij}^t$ transitions to state $e_l$ in frame $t + \ell$ is given by

$$\mathbf{s}^{t+\ell} = \mathbf{s}^t \mathbf{S}^{(\ell)} \tag{5.20}$$

where

$$s_l^{t+\ell} = \mathbb{P}\left\{Q_{ij}^{t+\ell} = e_l\right\}. \tag{5.21}$$

66

**Stochastic Queue State**

The critical contribution of the queue to the probability of connectivity in frame $t$ is whether or not there is a message in the queue at the conclusion of the previous frame, namely $\mathbb{P}\left\{Q_{ij}^{t-1} > 0\right\}$. The network sequence begins at $t = 0$ with empty queues; therefore, the vector of queue state probabilities is given by $\mathbf{s}^0 = [1, 0, 0, \ldots]$ since the probability of an empty queue is certain. Substituting these queue state probabilities into Equation (5.20) gives

$$\mathbf{s}^\ell = \mathbf{s}^0 \mathbf{S}^{(\ell)}, \tag{5.22}$$

and the queue state transition probabilities in frame $\ell = t - 1$ are given by

$$\mathbf{s}^{t-1} = \mathbf{s}^0 \mathbf{S}^{(t-1)}, \tag{5.23}$$

for $t \geq 2$. The probability that an individual queue is empty at the conclusion of frame $t - 1$ is equal to the first element of the queue state probability vector at frame $t - 1$, denoted $s_0^{t-1}$. Thus, the probability of at least one queued message from ND $i$ to ND $j$ in frame $t - 1$ is given by

$$\mathbb{P}\left\{Q_{ij}^{t-1} > 0\right\} = 1 - s_0^{t-1}. \tag{5.24}$$

Equation (5.24) provides the stochastic queue state in frame $t-1$ required to define stochastic frame connectivity in frame $t$.


## 5.2.3 Stochastic Frame Connectivity

Recall that each element of $\mathbf{F}^t$ in Equation (5.5) is the result of a logical OR operation followed by a logical AND operation. The OR operation corresponds to a transmission from ND $i$ in some frame $1 \leq \tau \leq t$, and is represented by the sum of $\mathbf{U}$ and $\mathbf{Q}^{t-1}$. The AND operation corresponds to reception by ND $j$ in frame $t$, and is represented by the Hadamard product of $\mathbf{D}$ with the result of the OR operation. These logical operations are considered stochastically through the probability of events.

**Probability of Events**

The following brief review of the probability of events is from [84] and [85]. The probability of event $A_1$ *and* $A_2$ occurring is their intersection, or joint probability, denoted

$$\mathbb{P}\{A_1 \cap A_2\} = \mathbb{P}\{A_1, A_2\}.$$

When $A_1$ and $A_2$ are statistically independent, their joint probability is the product of the individual probabilities, denoted

$$\mathbb{P}\{A_1, A_2\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_2\}.$$

The probability of event $A_1$ *or* $A_2$ occurring is their union minus their intersection,

$$\mathbb{P}\{A_1 \cup A_2\} = \mathbb{P}\{A_1\} + \mathbb{P}\{A_2\} - \mathbb{P}\{A_1, A_2\}.$$

**Stochastic Representation of Frame Connectivity**

Characterizing frame connectivity through the probability of events stochastically recasts Equation (5.5) as

$$\mathbb{P}\left\{F_{ij}^t = 1\right\} = \mathbb{P}\left\{(U_{ij} = 1 \ \cup \ Q_{ij}^{t-1} > 0), (D_{ij} = 1)\right\}. \tag{5.25}$$

Consider the first frame of a network sequence, $t = 1$, as an example. The lack of queued messages at the beginning of the network sequence (since $\mathbf{Q}^0 = \mathbf{0}$) means that a link from ND $i$ to ND $j$ requires an uplink RB allocation to ND $i$ and a downlink RB allocation to ND $j$, with probabilities given in equation (5.10). Given the independence of RB allocation events, the probability of a link from ND $i$ to ND $j$ in frame $t = 1$ is

$$\mathbb{P}\left\{F_{ij}^1 = 1\right\} = \mathbb{P}\left\{U_{ij} = 1\right\}\mathbb{P}\left\{D_{ij} = 1\right\} = p^2. \tag{5.26}$$

However, the impact of queued messages must be considered in all subsequent frames, $t > 1$.

The probability that, for $t > 1$, ND $i$ is either granted an uplink RB allocation to transmit to

68

ND $j$, or has a previously queued message for ND $j$ is given by

$$\mathbb{P}\left\{U_{ij} = 1 \cup Q_{ij}^{t-1} > 0\right\} =$$
$$\mathbb{P}\left\{U_{ij} = 1\right\} + \mathbb{P}\left\{Q_{ij}^{t-1} > 0\right\} - \mathbb{P}\left\{U_{ij} = 1\right\}\mathbb{P}\left\{Q_{ij}^{t-1} > 0\right\}. \quad (5.27)$$

Substituting the probabilities from Equations (5.10) and (5.24) into Equation (5.27), after simplification, gives

$$\mathbb{P}\left\{U_{ij} = 1 \cup Q_{ij}^{t-1} > 0\right\} = 1 - s_0^{t-1} + p(s_0^{t-1}). \quad (5.28)$$

Finally, substituting Equation (5.10) and the result of Equation (5.28) into Equation (5.25) gives

$$\mathbb{P}\left\{F_{ij}^{t} = 1\right\} = p(1 - s_0^{t-1} + p s_0^{t-1}). \quad (5.29)$$

### 5.2.4 Discussion

Equation (5.29) describes the probability of directed connectivity from ND $i$ to ND $j$ in any frame $t > 0$ based on the underlying random processes of $\mathbf{U}$, $\mathbf{D}$, and $\mathbf{Q}^{t-1}$. The Bernoulli process that characterizes RB allocation and the Markov process that describes the queue dynamics are both parameterized in terms of $p$ from Equation (4.2). Thus, for a fixed number of NDs and RBs, this result expresses the probability of temporal network connectivity, or network *robustness*, in frame $t$ as a function of the overloading ratio, $z$. A comparison of the probabilities of directed connectivity from ND $i$ to ND for frames $1 \leq k \leq 30$ in a network with 1,000 NDs, 100 RBs, and overloading ratios $z \in \{2, 3, 4\}$ is depicted in Figure 5.2. As one might expect, the probability of directed connectivity increases with $z$ (i.e., RBs are more available). This provides a means to measure the robustness of NOMA networks to variable overloading by the effect on probabilistic connectivity.

Employing probabilistic temporal connectivity as a robustness measure is particularly beneficial for designing NOMA networks that support massive connectivity or federated learning applications, whose performance depends on the ability to provide timely updates [16], [31]. The next section explored the probability of *initial* connectivity between ND $i$ and ND $j$ in frame $t$.
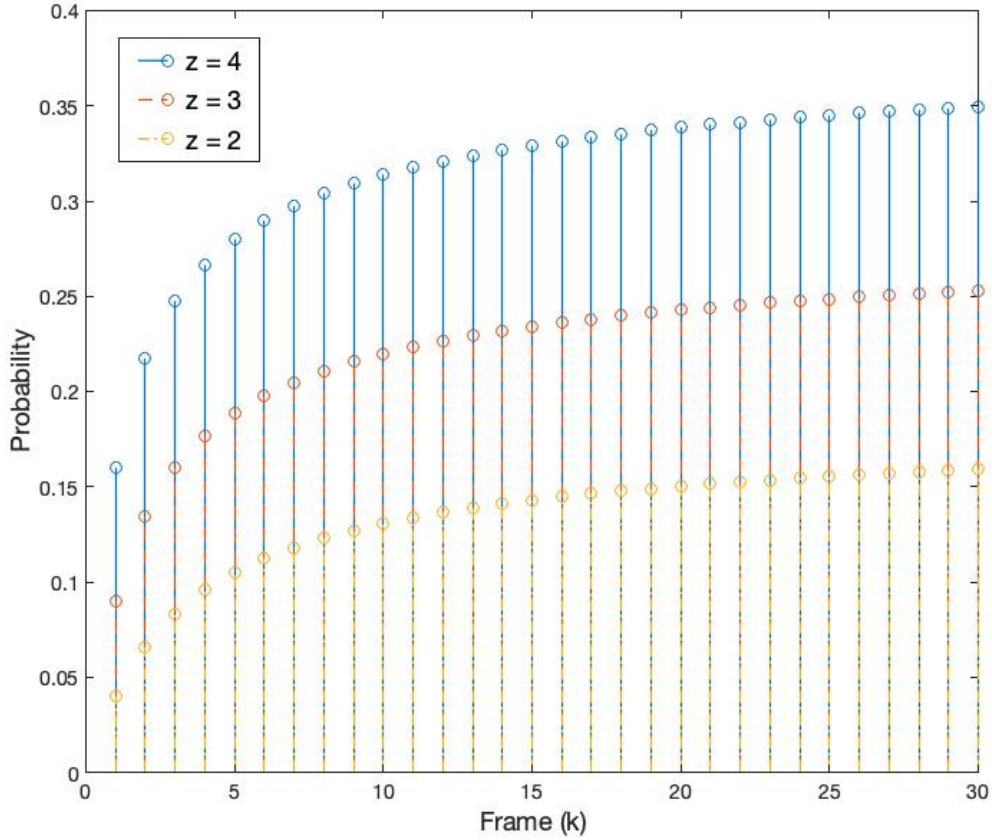
Figure 5.2. Probability of directed connectivity from ND $i$ to ND $j$ for frame $1 \leq k \leq 30$ in a network with 1,000 NDs, 100 RBs, and overloading ratios $z \in \{2, 3, 4\}$.

## 5.3 Probability of Time to Initial Unidirectional Connectivity

In addition to the probability of directed connectivity in frame $t$, NOMA network designers may need to consider the time required for ND $i$ and ND $j$ to first connect. Mathematically, the probability that the first directed link from ND $i$ to ND $j$ occurs in frame $T$ is the joint probability that $F_{ij}^T$ is equal to one, and equal to zero in all previous frames, expressed as $\mathbb{P}\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\}$.

In each frame $t$, a directed link from ND $i$ to ND $j$ either occurs, or does not. Thus, the

70

temporal evolution of $F_{ij}^t$ can be conceptualized as a decision tree rooted at frame $t = 0$ when the network sequence begins. The tree bifurcates into two branches at frame $t = 1$, where $F_{ij}^1 \in \{0, 1\}$ represents a link (value of 1) or no link (value of 0), and the probability that $F_{ij}^1 = 1$ is given by Equation (5.26). However, for all subsequent frames $t > 1$, the set of directed $i$-$j$ links that have not previously been established is limited to only those for which $F_{ij}^1 = 0$. Thus, by following the decision tree down the branch where $\sum_{t=0}^{T-1} F_{ij}^t = 0$ in each frame, the complement of the probability of the zero branch in frame $T$ is the probability that the first directed link from ND $i$ to ND $j$ occurs in frame $T$. An example decision tree for the first two frames of a network sequence is shown in Figure 5.3. The following sections formally develop this idea through the use of conditional probabilities.
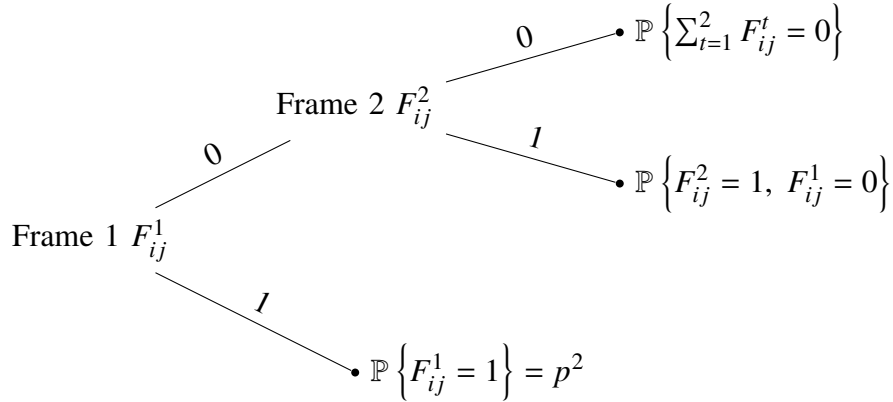


Figure 5.3. Decision tree depicting the branch of no directed connectivity from ND $i$ to ND $j$ in frame $t$. The complement of this branch in each frame is the probability of initial directed connectivity from ND $i$ to ND $j$ in frame $t$.

## 5.3.1 Joint Probability of Initial Connectivity

The joint probability of event $A_1$ and $A_2$ is the product of the probability of event $A_1$ conditioned on the occurrence of event $A_2$ and the individual probability of $A_2$ [84],

$$\mathbb{P}\{A_1, A_2\} = \mathbb{P}\{A_1|A_2\}\,\mathbb{P}\{A_2\}. \tag{5.30}$$

Thus, the joint probability that the first directed link from ND $i$ to ND $j$ occurs in frame $T$ can be expressed as the conditional probability

$$\mathbb{P}\left\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} = \mathbb{P}\left\{F_{ij}^T = 1 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} \mathbb{P}\left\{\sum_{t=0}^{T-1} F_{ij}^t = 0\right\}. \qquad (5.31)$$

Notice that the first term on the right side of Equation (5.31), $\mathbb{P}\left\{F_{ij}^T = 1 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}$, is similar to the term on the left side of Equation (5.29). Both equations consider the probability of directed connectivity from ND $i$ to ND $j$ in frame $T$, but Equation (5.31) conditions that probability on the event that no directed connectivity from ND $i$ to ND $j$ has occurred up to frame $T - 1$. This has no effect on the $p$ terms in Equation (5.29), since those terms result from IID realizations of the subframe sampling process. However, the $s_0^{t-1}$ term (which is the probability that the queue is empty in frame $t - 1$) is dependent on previous queue states and must be modified to express the probability that queue is empty in frame $T - 1$, conditioned on $\sum_{t=0}^{T-1} F_{ij}^t = 0$, given by

$$\mathbb{P}\left\{F_{ij}^T = 1 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} =$$
$$p\left(1 - \mathbb{P}\left\{Q_{ij}^{T-1} = 0 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} + p\mathbb{P}\left\{Q_{ij}^{T-1} = 0 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}\right). \quad (5.32)$$

Equation (5.32) includes two terms with the conditional probability that the queue is empty in frame $T - 1$ when no directed connectivity from ND $i$ to ND $j$ has occurred up to frame $T - 1$, $\mathbb{P}\left\{Q_{ij}^{T-1} = 0 \ \middle| \ \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}$. Substituting the result of Equation (5.32) back into Equation (5.31) will result in both of these terms being multiplied by the probability that no directed connectivity from ND $i$ to ND $j$ has occurred up to frame $T - 1$, $\mathbb{P}\left\{\sum_{t=0}^{T-1} F_{ij}^t = 0\right\}$. By the joint probability relation in Equation (5.30), both of these terms will become the joint probability of the two events, $\mathbb{P}\left\{Q_{ij}^{T-1} = 0, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}$. Thus, substituting the result of equation (5.32) into equation (5.31) and simplifying with the joint probability relation

gives

$$\mathbb{P}\left\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} =$$

$$p\left(\mathbb{P}\left\{\sum_{t=0}^{T-1} F_{ij}^t = 0\right\} - \mathbb{P}\left\{Q_{ij}^{T-1} = 0, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} + p\mathbb{P}\left\{Q_{ij}^{T-1} = 0, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}\right). \quad (5.33)$$

It is clear from Equation (5.33) that the probability that no directed link from ND $i$ to ND $j$ has occurred up to frame $T - 1$ and the joint probability of an empty queue in frame $T - 1$ when $\sum_{t=0}^{T-1} F_{ij}^t = 0$ are both critical to understanding the probability of initial directed connectivity from ND $i$ to ND $j$ in frame $T$. The following sections develop each of these ideas.

### 5.3.2 Joint Queue Transition Probabilities

This section explores the impact of the knowledge that $\sum_{t=0}^{T-1} F_{ij}^t = 0$, on the probability that the queue from ND $i$ to ND $j$ is empty in frame $T - 1$. Consider the queue from ND $i$ to ND $j$ in frame $t = 1$ when no directed link occurred in frame $t = 1$ (i.e., $F_{ij}^1 = 0$). The knowledge that $F_{ij}^1 = 0$ limits the mutually exclusive probability space of uplink RB allocation to ND $i$ and downlink RB allocation to ND $j$ that could have occurred in frame $t = 1$ to the following subspace:

1. ND $i$ is granted an uplink RB allocation with probability $p$, and ND $j$ is not granted a downlink RB allocation with probability $q = 1 - p$.
2. ND $i$ is not granted an uplink RB allocation with probability $q$, and ND $j$ is granted a downlink RB allocation with probability $p$.
3. Neither ND $i$ or ND $j$ are granted a RB allocation, each with probability $q$.

Of these three possibilities, only the first results in a transition of $Q_{ij}^1$ from a 0 to a 1 (recall that all queues are empty at $t = 0$). This means the joint probability that $Q_{ij}^1 = 0$ and $F_{ij}^1 = 0$ is equal to $qp + q^2$. Similarly, the joint probability that $Q_{ij}^1 = 1$ and $F_{ij}^1 = 0$ is equal to $pq$.

The knowledge that $F_{ij}^1 = 0$ excludes the possibility that ND $i$ and ND $j$ are both granted a RB allocation since, from Equation (5.26), this results in $F_{ij}^1 = 1$. This restriction on the

RB allocations that could have occurred in frame $t = 1$ directly impacts the queue state transition probabilities in Equation (5.19). When the queue is in state $e_0$ (i.e., the queue is empty), the probability that the queue remains empty after one transition, $r_0$, excludes $p^2$ (i.e., when ND $i$ and ND $j$ are both granted a RB allocation). Thus, the *restricted* transition probabilities when the queue is in state $e_0$, denoted by $(\cdot)'$, are

$$r'_0 = qp + q^2 = q(p + q) = q(p + 1 - p) = q \tag{5.34}$$

and

$$p'_0 = pq. \tag{5.35}$$

Similarly, when the queue is in state $e_l$ (i.e., the queue is not empty), the probability subspace also excludes $qp$ from transition probability $q_l$ (i.e., recall that $q_l$ is the probability that the queue reduces by one message). This is because a queue reduction implies a link from ND $i$ to ND $j$ occurred in some frame $2 \le t \le T - 1$, which violates the condition that $\sum_{t=0}^{T-1} F_{ij}^t = 0$. Thus, the *restricted* transition probabilities when the queue is in state $e_l$ are

$$r'_l = q^2, \tag{5.36}$$

$$p'_l = p_l = pq, \tag{5.37}$$

and

$$q'_l = 0. \tag{5.38}$$

**Restricted State Transition Matrix**

Updating the state transition matrix in Equation (5.19) with the joint state transition probabilities in Equations (5.34)–(5.38) gives the *restricted* state transition matrix, **H**, defined as

$$\mathbf{H} = \left\{ \begin{matrix} r'_0 & p'_0 & 0 & 0 & 0 & \cdots \\ 0 & r'_1 & p'_1 & 0 & 0 & \cdots \\ 0 & 0 & r'_2 & p'_2 & 0 & \cdots \\ 0 & 0 & 0 & r'_3 & p'_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{matrix} \right\}. \tag{5.39}$$

Note that referring to $\mathbf{H}$ as a state transition matrix is possibly an abuse of terminology since the joint transition probabilities in each row sum to less than one. The first row sums to $1 - p^2$, and all others sum to $1 - p^2 - qp$. However, each row can be normalized to conditional probabilities that sum to one by dividing by the probability that $F_{ij}^1 = 0$. The first row corresponds to the queue in state $e_0$, so the probability that $F_{ij}^1 = 0$ is equal to $1 - p^2$. All others rows correspond to the queue in state $e_l$, so probability that $F_{ij}^1 = 0$ is equal to $1 - p^2 - qp$. Higher order transition probabilities can be similarly normalized by the probability that the $\sum_{t=0}^{T-1} F_{ij}^t = 0$ for states $e_0$ and $e_l$. However, recall that the conditional queue probabilities in Equation (5.32) are converted to joint probabilities in Equation (5.33) due to multiplication by $\mathbb{P}\left\{\sum_{t=0}^{T-1} F_{ij}^t = 0\right\}$. Thus, rather than normalizing the rows of $\mathbf{H}$ to conditional transition probabilities (only to convert them back to joint probabilities), the *restricted* state transition matrix intentionally incurs some terminological abuse.

**Restricted Stochastic Queue State**

Updating Equation (5.23) with the *restricted* queue state transition matrix from Equation (5.39) defines the *restricted* queue state probability vector, $\mathbf{h}$, as

$$\mathbf{h}^{t-1} = \mathbf{s}^0 \mathbf{H}^{(t-1)}. \tag{5.40}$$

As discussed in Section 5.2.2, the queue is empty when the network sequence begins, so the vector of queue state probabilities at $t = 0$ remains unchanged as $\mathbf{s}^0 = [1, 0, 0, \ldots]$. It follows that, similar to the previous argument in Section 5.2.2, the joint probability that the queue is empty and that no directed link from ND $i$ to ND $j$ has occurred up to frame $T - 1$ is equal to the first element of the *restricted* queue state probability vector, denoted $h_0^{t-1}$, at frame $T - 1$, given by

$$\mathbb{P}\left\{Q_{ij}^{T-1} = 0, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} = h_0^{T-1} \tag{5.41}$$

It is clear from Equations (5.39) and (5.40) that raising $\mathbf{H}$ to the $(t-1)$ power and multiplying by $\mathbf{s}^0$ gives

$$h_0^{t-1} = r_0'^{(t-1)} = q^{(t-1)}. \tag{5.42}$$

Thus, the joint probability of an empty queue in frame $T - 1$, subject to the constraint of no

75

prior directed connectivity from ND $i$ to ND $j$, is

$$\mathbb{P}\left\{Q_{ij}^{T-1} = 0, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} = q^{(T-1)}. \tag{5.43}$$

### 5.3.3 Probability Mass Function

Equipped with the result of the joint queue probability from Equation (5.43), this section derives the probability that no directed link from ND $i$ to ND $j$ has occurred up to frame $T - 1$, and develops the PMF for initial directed connectivity between two NDs.

Recall the branching conceptualization described at the beginning of Section 5.3. Since link establishment from ND $i$ to ND $j$ is a binary proposition (either it occurs or it does not), the joint probability that no directed link occurs from ND $i$ to ND $j$ up to frame $T - 1$ can be defined as the complement of equation Equation (5.33) with respect to the set of NDs between which no directed link has occurred up to frame $T - 2$ (i.e., the decision tree branch where $\sum_{t=0}^{T-2} F_{ij}^t = 0$).

Consider frame $t = 2$ in Figure 5.3. The probability of the branch of two consecutive zeros in Frame 2 (top branch) can be found by subtracting the probability of the 0–1 branch of Frame 2 from the zero branch of Frame 1. Expressed mathematically, the probability that no directed link occurs from ND $i$ to ND $j$ up to frame $t = 2$ is given by

$$\mathbb{P}\left\{\sum_{t=0}^{2} F_{ij}^t = 0\right\} = \mathbb{P}\left\{F_{ij}^1 = 0\right\} - \mathbb{P}\left\{F_{ij}^2 = 1, F_{ij}^1 = 0\right\} \tag{5.44}$$

Generalizing this to an arbitrary frame, $T - 1$, gives

$$\mathbb{P}\left\{\sum_{t=0}^{T-1} F_{ij}^t = 0\right\} = \mathbb{P}\left\{\sum_{t=0}^{T-2} F_{ij}^t = 0\right\} - \mathbb{P}\left\{F_{ij}^{T-1} = 1, \sum_{t=0}^{T-2} F_{ij}^t = 0\right\}. \tag{5.45}$$

Substituting the results of Equations (5.43) and (5.45) into Equation (5.33) yields a recursive expression for the probability that a directed link from ND $i$ to ND $j$ is first established in

frame $T$, given by

$$\mathbb{P}\left\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\} =$$

$$p\left(\mathbb{P}\left\{\sum_{t=0}^{T-2} F_{ij}^t = 0\right\} - \mathbb{P}\left\{F_{ij}^{T-1} = 1, \sum_{t=0}^{T-2} F_{ij}^t = 0\right\} - q^{T-1} + pq^{T-1}\right). \quad (5.46)$$

Thus, given a set of NOMA network parameters, Equations (5.26) and (5.46) provide the basis to define the PMF for the frame $k$ in which a directed link between ND $i$ and ND $j$ is first established,

$$\mathrm{f}_K[k] = \begin{cases} 0 & , k = 0 \\ p^2 & , k = 1 \\ p\left(\mathbb{P}\left\{\sum_{t=0}^{k-1} F_{ij}^t = 0\right\} \\ -q^{k-1} + pq^{k-1}\right) & , k \geq 2 \end{cases} . \quad (5.47)$$

One can imply from the PMF definition that the $\mathbb{P}\left\{F_{ij}^0 = 0\right\} = 1$, and the $\mathbb{P}\left\{F_{ij}^1 = 0\right\} = 1 - p^2$. These definitions are both consistent with the constraints that there is no connectivity between NDs at the beginning of the network sequence and, correspondingly, that all queues are empty at the beginning of the network sequence.

### 5.3.4 Discussion

Intuitively, increasing RB availability (via the overloading ratio) increases the connectivity occurring between NDs in each frame, and reduces the time to initial unidirectional connectivity. The result in Equation (5.47) confirms this intuition, as shown by comparison of the resulting PMFs for networks with 10,000 NDs, 1,000 RBs, and $z \in \{2, 3, 4\}$ in Figure 5.4. For example, the probability of initial directed connectivity from ND $i$ to ND $j$ in the third frame (i.e., $k = 3$) is 0.0768 when $z = 2$, but increases to 0.1728 when $z = 4$. Thus, similar to the analysis of directed connectivity in each frame in Section 5.2, the network demonstrates more robust temporal connectivity as the overloading ratio increases.

As referenced in the introduction, these results are applicable for NOMA networks supporting federated learning applications. Federated learning systems partition the training data or model parameters among NDs for local computation, and then exchange all results for
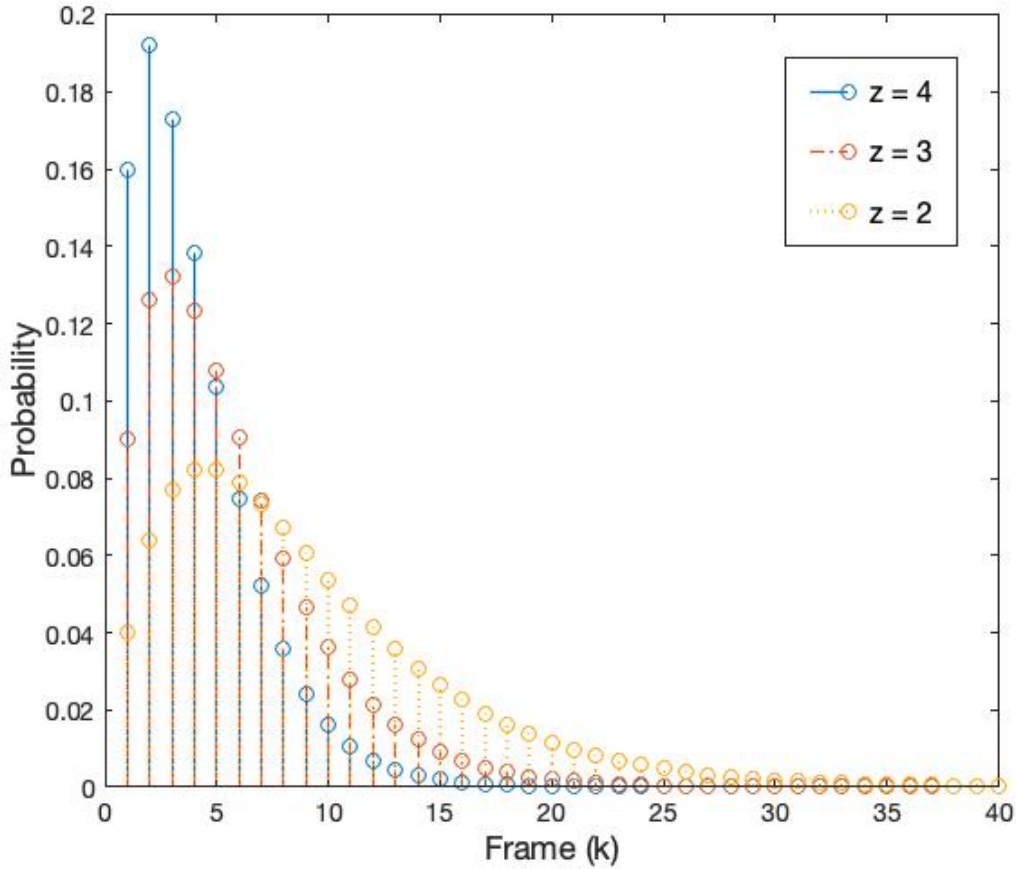
Figure 5.4. PMF comparison for initial directed connectivity in a network with 10,000 NDs, 1,000 RBs, and overloading ratios $z \in \{2, 3, 4\}$. Source: [65]. © 2022 IEEE.

model convergence [31]. Consequently, the time required to achieve pairwise communications between all NDs (i.e., a complete graph) is an important factor in evaluating the impact of network design on application performance. Additionally, some federated learning approaches only require a subset of NDs to exchange local computations for convergence [100]; hence, considering each computational iteration/epoch (leading to convergence) as the beginning of a network sequence, these results provide a stochastic understanding of when a sufficient number of NDs will have communicated to achieve algorithmic convergence.

## 5.4 Probability of Time Window between Unidirectional Connectivity

In the same way that understanding how soon to expect connectivity between two NDs is important, characterizing the amount of time between connectivity events is also important. The time "window" between directed connectivity from ND $i$ to ND $j$ is particularly important for mMTC and URLLC network use-cases. A mMTC use case might include large sensor deployments or IoT applications in which a large number of NDs have periodic reporting requirements to an automated information fusion center [101]. A use-case in which the mMTC deployment is supporting closed-loop control algorithms would add URLLC requirements as well.

The time window between directed connectivity events from ND $i$ to ND $j$ in the NOMA temporal network model is similar to the time between node activation in the activity-driven model, which obeys a geometric distribution [60]. This becomes evident when the behavior of the queue is considered as the network sequence becomes long.

Recall from Equations (5.23) and (5.24) that the probability of an empty queue is a function of the queue state transition matrix $\mathbf{S}$ raised to the $(t - 1)$ power. Since all state transition probabilities in $\mathbf{S}$ are positive real numbers less than one, each successive multiplication of $\mathbf{S}$ by itself reduces the value of $S_{1,1}$ toward zero which, by Equation (5.23), is equal to $s_0^{t-1}$. Expressed as a limit, $s_0^{t-1}$ approaches zero as the network sequence becomes long,

$$\lim_{t \to \infty} s_0^{t-1} = 0. \tag{5.48}$$

This means the probability that the queue is empty goes to zero as the network sequence becomes long. Recall from Equation (5.29) that the $\mathbb{P}\left\{F_{ij}^t = 1\right\} = p(1 - s_0^{t-1} + ps_0^{t-1})$. Incorporating the result from Equation (5.48) into Equation (5.29) gives,

$$\lim_{t \to \infty} \mathbb{P}\left\{F_{ij}^t = 1\right\} = \lim_{t \to \infty} p(1 - s_0^{t-1} + ps_0^{t-1}) = p. \tag{5.49}$$

Since the probability of directed connectivity from ND $i$ to ND $j$ converges to $p$ as the network sequence becomes long, the time window between directed connectivity from ND $i$ to ND $j$ is described by a geometric probability distribution parameterized by $p$. Specifically, the probability that the number of frames between directed connectivity from

ND $i$ to ND $j$ is equal to $k$ is given by the PMF,

$$f_K[k] = p(1-p)^k, \qquad \text{for} \quad 0 \le k < \infty. \tag{5.50}$$

However, since Equation (5.48) is a limit, the probability that the queue is empty approaches zero, but never reaches it. Therefore, the PMF in Equation (5.50) will become increasingly accurate as the length of the network sequence increases, but the $\mathbb{P}\left\{F_{ij}^t = 1\right\}$ never reaches $p$. The results in Chapter 6 further illustrate this.

## 5.5 Probability of Time to Initial Bidirectional Connectivity

Similar to the idea expressed at the beginning of Section 5.3, NOMA network designers may need to consider the time required for initial bidirectional connectivity between two NDs, as might be required for any application requiring a two-way handshake.

Initial bidirectional connectivity between ND $i$ and ND $j$ in frame $T$ can occur in two ways:

1. ND $i$ can establish directed connectivity to ND $j$ in some frame $0 \le t \le T$, and ND $j$ establishes directed connectivity to ND $i$ in frame $T$.
2. ND $j$ can establish directed connectivity to ND $i$ in some frame $0 \le t \le T$, and ND $i$ establishes directed connectivity to ND $j$ in frame $T$.

Establishing initial bidirectional connectivity by the first method requires that both NDs are granted RB allocations in a way that meets the following three conditions:

1. $\sum_{t=0}^{T} F_{ij}^t > 0$
2. $F_{ji}^T = 1$
3. $\sum_{t=0}^{T-1} F_{ji}^t = 0$.

The first condition ensures that at least one directional link from ND $i$ to ND $j$ occurs in the time interval $0 < t \le T$, while the second and third conditions ensure that the first link from ND $j$ to ND $i$ does not occur until frame $T$. Reversing the order of $i$ and $j$ in each condition provides the set of conditions required for initial bidirectional connectivity by the

second method.

Let $B_{ij}^T \in \{0, 1\}$ denote the event of initial bidirectional connectivity in frame $T$ by the first method, and define $B_{ij}^T$ as the the logical conjunction of the three conditions,

$$B_{ij}^T \stackrel{\text{def}}{=} \sum_{t=0}^{T} F_{ij}^t > 0 \ \wedge \ \left( F_{ji}^T = 1 \ \wedge \ \sum_{t=0}^{T-1} F_{ji}^t = 0 \right), \tag{5.51}$$

where $\wedge$ denotes the logical conjunction (AND) operation. Thus, $B_{ij}^T$ only assumes a value of one when all three conditions are true. Correspondingly, $B_{ji}^T$ represents the event of initial bidirectional connectivity in frame $T$ by the second method. The disjunction of $B_{ij}^T$ and $B_{ji}^T$ (i.e., $B_{ij}^T$ OR $B_{ji}^T$) represents the event of initial bidirectional connectivity between ND $i$ and ND $j$ in frame $T$.

### 5.5.1 Bidirectional Connectivity Matrix

Recall that the frame connectivity matrix, $\mathbf{F}^t$, in Equation (5.5) is a directed adjacency matrix that describes the connectivity achieved during frame $t$. The rows of $\mathbf{F}^t$ represent the directed connectivity from ND $i$ to ND $j$ and the columns represent the directed connectivity from ND $j$ to ND $i$. The Hadamard product of $\mathbf{F}^t$ with its transpose results in an undirected symmetric adjacency matrix, $\mathbf{A}^t$, with elements denoted $A_{ij}^t$, given by

$$\mathbf{A}^t = \mathbf{F}^t \odot (\mathbf{F}^t)^\top \tag{5.52}$$

where $\top$ denotes the transpose operation. Similar to $B_{ij}^T$, $A_{ij}^t \in \{0, 1\}$ where the value one represents the occurrence of bidirectional connectivity in frame $t$, and zero represents the converse. As a result of the matrix symmetry of $\mathbf{A}^t$ (i.e., $A_{ij}^t = A_{ji}^t$), establishing initial bidirectional connectivity between ND $i$ and ND $j$ in frame $T$ is represented by the conjunction of the events $A_{ij}^T = 1$ and $\sum_{t=0}^{T-1} A_{ij}^t = 0$.

### 5.5.2 Probability of Initial Bidirectional Connectivity

The logical representations of initial bidirectional connectivity between ND $i$ and ND $j$ developed in Sections 5.5 and 5.5.1 lead to the logical expression,

$$\left(A_{ij}^T = 1\right) \wedge \left(\sum_{t=0}^{T-1} A_{ij}^t = 0\right) = \left(B_{ij}^T = 1\right) \vee \left(B_{ji}^T = 1\right), \tag{5.53}$$

where $\vee$ denotes the logical disjunction (OR) operation. Considering the logical expression in Equation (5.53) stochastically gives the probability of initial bidirectional connectivity between ND $i$ and ND $j$ in frame $T$,

$$\mathbb{P}\left\{A_{ij}^T = 1, \sum_{t=0}^{T-1} A_{ij}^t = 0\right\} = \mathbb{P}\left\{B_{ij}^T = 1 \cup B_{ji}^T = 1\right\}. \tag{5.54}$$

Expanding the the right side of Equation (5.54) using the relation $\mathbb{P}\{A_1 \cup A_2\} = \mathbb{P}\{A_1\} + \mathbb{P}\{A_2\} - \mathbb{P}\{A_1, A_2\}$, introduced in Section 5.2.3 from [84] and [85], gives

$$\mathbb{P}\left\{A_{ij}^T = 1, \sum_{t=0}^{T-1} A_{ij}^t = 0\right\} = \mathbb{P}\left\{B_{ij}^T = 1\right\} + \mathbb{P}\left\{B_{ji}^T = 1\right\} - \mathbb{P}\left\{B_{ij}^T = 1, B_{ji}^T = 1\right\}, \tag{5.55}$$

which provides three separate terms for examination.

First, the $\mathbb{P}\left\{B_{ij}^T = 1\right\}$ is expressed by rewriting Equation (5.51) as a joint probability,

$$\mathbb{P}\left\{B_{ij}^T = 1\right\} = \mathbb{P}\left\{\sum_{t=0}^{T} F_{ij}^t > 0, \ F_{ji}^T = 1, \ \sum_{t=0}^{T-1} F_{ji}^t = 0\right\}. \tag{5.56}$$

Notice that the second are third terms on the right side of Equation (5.56) are in same form as the probability of initial directed connectivity from ND $i$ to ND $j$ in Equation (5.46). The only difference is that ND $i$ and ND $j$ are reversed (i.e., this expression considers the connectivity from ND $j$ to ND $i$). Additionally, the first term on the right side of Equation (5.56) considers directed connectivity from ND $i$ to ND $j$. It follows from the analysis in Sections 5.2 and 5.3 that the directed connectivity from ND $i$ to ND $j$ is

statistically independent from the directed connectivity from ND $j$ to ND $i$, resulting in

$$\mathbb{P}\left\{B_{ij}^T = 1\right\} = \left(1 - \mathbb{P}\left\{\sum_{t=0}^{T} F_{ij}^t = 0\right\}\right)\mathbb{P}\left\{F_{ji}^T = 1, \sum_{t=0}^{T-1} F_{ji}^t = 0\right\} \tag{5.57}$$

where $\mathbb{P}\left\{\sum_{t=0}^{T} F_{ij}^t > 0\right\} = 1 - \mathbb{P}\left\{\sum_{t=0}^{T} F_{ij}^t = 0\right\}$, and all probabilistic terms can be calculated from Equations (5.45) and (5.46).

Next, note that reversing $i$ and $j$ in Equation (5.57) does not change the result since the probabilities are equal $\forall F_{ij}^t$ (for $i \neq j$), which gives

$$\mathbb{P}\left\{B_{ij}^T = 1\right\} = \mathbb{P}\left\{B_{ji}^T = 1\right\}. \tag{5.58}$$

Next, from the definition in Equation (5.51), the joint event that $B_{ij}^T = 1$ and $B_{ji}^T = 1$ can only occur if the initial directed connectivity from ND $i$ to ND $j$ and the initial directed connectivity from ND $j$ to ND $i$ both happen in frame $T$. Both of these events are described by Equation (5.46). Given the independence of these events, and that reversing $i$ and $j$ in Equation (5.46) does not change the result, the probability is given by the square of Equation (5.46),

$$\mathbb{P}\left\{B_{ij}^T = 1, B_{ji}^T = 1\right\} = \mathbb{P}\left\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}^2. \tag{5.59}$$

Finally, substituting the results of Equations (5.57), (5.58), and (5.59) back into Equation (5.55) gives

$$\mathbb{P}\left\{A_{ij}^T = 1, \sum_{t=0}^{T-1} A_{ij}^t = 0\right\} =$$

$$2\left(1 - \mathbb{P}\left\{\sum_{t=0}^{T} F_{ij}^t = 0\right\}\right)\mathbb{P}\left\{F_{ji}^T = 1, \sum_{t=0}^{T-1} F_{ji}^t = 0\right\} - \mathbb{P}\left\{F_{ij}^T = 1, \sum_{t=0}^{T-1} F_{ij}^t = 0\right\}^2. \tag{5.60}$$

Thus, given a set of NOMA network parameters, Equation (5.60) provide the basis to define the PMF for the frame $k$ in which bidirectional connectivity between ND $i$ and ND $j$ is first

established,

$$
f_K[k] = \begin{cases} 0 & , k = 0 \\ 2\big(1 - \mathbb{P}\big\{ \sum_{t=0}^{k} F_{ij}^t = 0 \big\}\big) & \\ \mathbb{P}\big\{ F_{ji}^k = 1, \sum_{t=0}^{k-1} F_{ji}^t = 0 \big\} & \\ -\mathbb{P}\big\{ F_{ij}^k = 1, \sum_{t=0}^{k-1} F_{ij}^t = 0 \big\}^2 & , k \geq 1 \end{cases}
$$

(5.61)

### 5.5.3 Discussion

Similar to the results for initial unidirectional connectivity in Section 5.3, increasing RB availability via overloading reduces the time to initial bidirectional connectivity. A comparison of the PMFs resulting from Equation (5.61) for networks with 100 NDs, 10 RBs, and overloading ratios $z \in \{2, 3, 4\}$ is depicted in Figure 5.5. Notice that, when compared to Figure 5.4, the peak of each distribution is shifted to the right and the tail is slightly longer. This reflects the larger number of frames required to achieve directed connectivity in both directions.

The initial bidirectional connectivity results are consistent with the analytical development in Sections 5.2–5.4, and demonstrate the relationship between robust temporal connectivity and overloading variability. These results have general applicability for any network service or application that requires bidirectional communication, such as a Transmission Control Protocol (TCP) synchronization and acknowledgement, or key exchange for authentication and encryption.
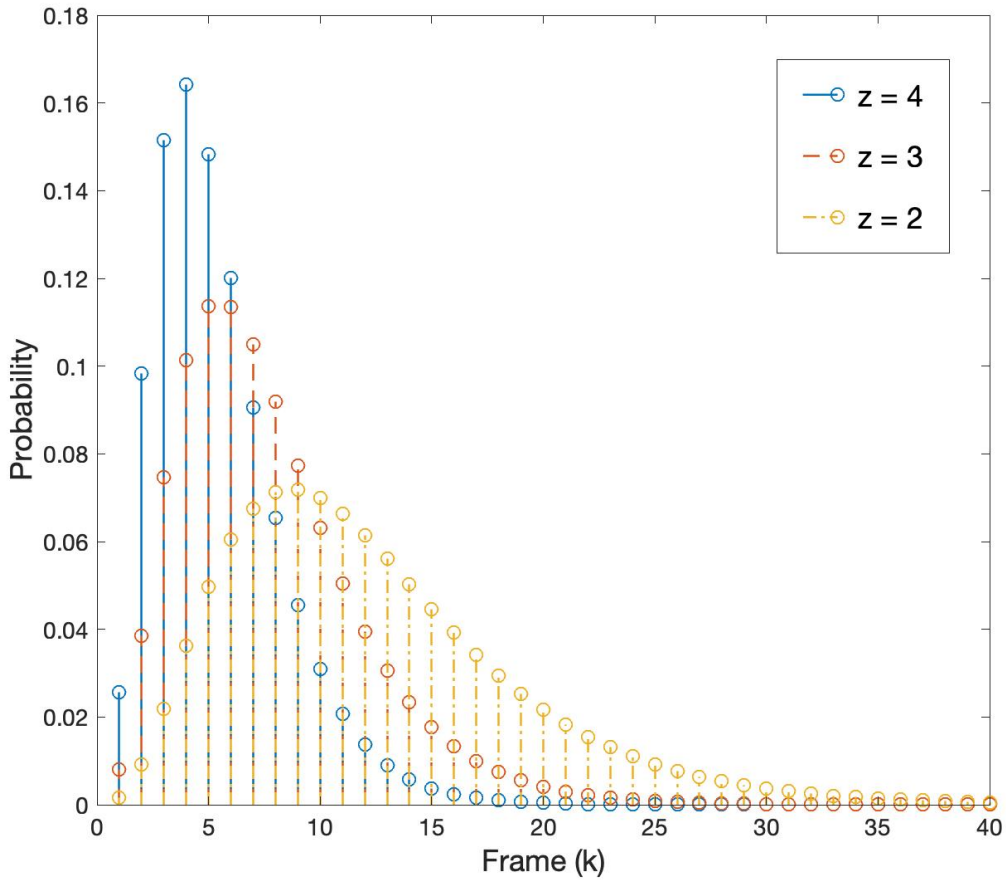
Figure 5.5. PMF comparison for initial bidirectional connectivity in a network with 100 NDs, 10 RBs, and overloading ratios $z \in \{2, 3, 4\}$.

## 5.6 Probability of Minimum Time to Complete Bidirectional Connectivity

This section examines the minimum number of frames required to achieve a complete affine graph and the probability of occurrence. Recall from Chapter 2.3.2 that that the affine graph provides a temporal measure of bidirectional connectivity between ND pairs in the network. This is because the affine graph is the static projection of all temporal strongly connected node pairs in a time interval, and temporal strongly connected nodes are those that each

have a temporal path to the other. Thus, the affine graph shows all pairwise connected nodes over a defined time interval. Pairwise ND communication is important in distributed computing; thus, similar to Sections 5.3 and 5.5, this analysis relates NOMA overloading to the supported convergence time of distributed computing applications that could be deployed on 5G or B5G NTNs [31], [102].

A complete affine graph occurs when the sum of all positive elements from all bidirectional connectivity adjacency matrices over a specified number of frames, $T$, is equal to $m(m-1)$,

$$\sum_{t=1}^{T} \mathbb{1}_{\mathbb{Z}^+} \left( \mathbf{A}^t \right) = m(m-1), \tag{5.62}$$

where $\mathbf{A}^t$ is defined in Equation (5.52). The minimum number of frames required to form a complete affine graph, denoted $F_{min}$, is the sum of the number of frames required for all NDs to be granted an uplink RB allocation, and the number of frames required for all NDs to be granted a downlink RB allocation *after* all have been granted an uplink RB allocation. The corresponding probability of this event depends on the NOMA network parameters ($m$ NDs, $n$ RBs, and overloading ratio $z$) and includes the following possible cases.

## 5.6.1   Case 1: $zn = m - 1$

This case occurs when all but one ND are granted a RB allocation in each subframe. A complete affine graph can be achieved in two frames with probability

$$\mathbb{P}\left\{F_{min} = 2\right\} = p^2(1-p), \tag{5.63}$$

where $p$ is the probability of RB allocation for an individual ND defined in Equation (5.10).

Consider the case in which a single ND, $i$, is not granted a RB allocation in the first uplink subframe, but is granted a RB allocation in the first downlink subframe. This means ND $i$ has received transmissions from all other NDs (since ND $i$ does not transmit to itself). If ND $i$ is granted an uplink RB allocation in the second uplink subframe, then all NDs have been granted an uplink RB allocation. Since ND $i$ has already received transmissions from all other NDs, it does not require an additional downlink RB allocation, so if all other $m - 1$ NDs are granted a downlink RB allocation in the second downlink subframe, then a

complete affine graph will have been achieved. The probability of this sequence of frames is characterized entirely by the RB allocation probability, $p$, to ND $i$.

## 5.6.2   Case 2: $\frac{m}{2} \leq zn < m - 1$

This case occurs when at least half of the NDs in the network are granted a RB allocation in each subframe, but less than $m - 1$. A complete affine graph can be achieved in three frames with probability

$$\mathbb{P}\left\{F_{min} = 3\right\} = \left[\frac{\binom{zn}{2zn-m}}{\binom{m}{zn}}\right]^2 . \tag{5.64}$$

In this case, a minimum of two uplink subframes are required for all NDs to be granted at least one uplink RB allocation. Similarly, a minimum of two downlink subframes are required for all NDs to be granted at least one downlink RB allocation. All NDs must be granted a downlink RB allocation *after* all NDs have been granted an uplink RB allocation in order to receive a transmission directly in the same frame, or indirectly from the message queue. Thus, the downlink RB allocation must begin in the second downlink subframe, and conclude in the third downlink subframe.

Consider a NOMA wireless network parameterized by $m = 6$, $n = 2$, and $z = 2$. In the first frame, the probability of selecting four NDs which have not yet been granted an uplink RB allocation is equal to one since the network sequence has just begun. In the second frame, the $m - zn = 2$ NDs that were not granted an uplink RB allocation in the first uplink subframe must be selected, as well as $zn - (m - zn) = 2$ of the $zn = 4$ NDs that were granted an uplink RB allocation in the first uplink subframe. The probability of this event is

$$\frac{\binom{m-zn}{m-zn}\binom{zn}{zn-(m-zn)}}{\binom{m}{zn}} = \frac{\binom{zn}{2zn-m}}{\binom{m}{zn}} = \frac{\binom{2}{2}\binom{4}{2}}{\binom{6}{4}} = \frac{\binom{4}{2}}{\binom{6}{4}} .$$

The second downlink subframe is the first opportunity for NDs to be granted a downlink RB allocation after all NDs have been granted an uplink RB allocation. Thus, the RB allocation sequence that is required for the uplink in the first and second frames must repeat for the

downlink in the second and third frames to achieve a complete affine graph, which gives

$$\mathbb{P}\left\{F_{min} = 3\right\} = \left[\frac{\binom{4}{2}}{\binom{6}{4}}\right]^2.$$

Generalizing this to any combination of NDs, RBs, and overloading ratio that meet these conditions results in Equation 5.64. Note that when $zn = \frac{m}{2}$, Equation (5.64) reduces to

$$\mathbb{P}\left\{F_{min} = 3\right\} = \binom{m}{m/2}^{-2}. \tag{5.65}$$

## 5.6.3 Case 3: $1 < zn < \frac{m}{2}$

This case occurs when more than one, but less than half, of the NDs are granted a RB allocation in each subframe. There are three sub-cases within Case 3.

**Subcase 3.1:** $m \mod zn = 0$

This occurs when the product of RBs and the overloading ratio divide evenly into the number of NDs. The number of frames required for all NDs to be granted a RB allocation is equal to $\frac{ND}{zn}$, and the downlink RB allocation can begin in the same frame that all NDs are granted an uplink allocation. Hence, the minimum number of frames required to achieve a complete affine graph is given by

$$F_{min} = \frac{2m}{zn} - 1. \tag{5.66}$$

The probability of this event is a product of the hypergeometric distribution evaluated with constant population ($m$), sample size ($zn$), and desired number of successes ($zn$), but a decreasing number of known successes in the population. The NDs that have not yet been granted an uplink or downlink RB allocation are the number of successes in the population, and this value is decremented by $zn$ after each subframe. The probability of selecting $zn$ NDs that have not been granted a RB allocation in the first uplink subframe is one since no NDs have been granted an RB allocation at the beginning of the network sequence. In the second uplink subframe, $zn$ of the $m - zn$ NDs that have not been granted an uplink RB allocation must be selected. This decrementing operation continues until all NDs have been granted an uplink RB allocation, and then repeats for the downlink, resulting in a

probability given by

$$\mathbb{P}\left\{F_{min} = \frac{2m}{zn} - 1\right\} = \prod_{i=1}^{\frac{m}{zn}-1} \frac{\binom{m-i(zn)}{zn}^2}{\binom{m}{zn}^2}.$$  (5.67)

Since no NDs that have been granted a RB allocation are selected again, the second term in the numerator from Equation (4.17) that accounts for the selection of failures reduces to one (since none are selected in each frame) and the square operation accounts for the repetition of the process in the uplink and downlink.

**Subcase 3.2:** $m \mod zn = 1$

This occurs when any multiple of the product of RBs and the overloading ratio is one less than the number of NDs. The logic for the minimum frames required for a complete affine graph is the same as Case 1, except there are more frames prior to reaching the frame in which only a single ND has not been granted an uplink RB allocation. Accounting for those additional frames, the minimum number of frames required to achieve a complete affine graph is given by

$$F_{min} = 2 \left\lfloor \frac{m}{zn} \right\rfloor.$$  (5.68)

The probability formulation for this case is similar to Subcase 3.1 in that a product of hypergeometric distributions is evaluated at decrementing values of the number of successes in the population. The primary differences occur in downlink subframe number $\left\lfloor \frac{m}{zn} \right\rfloor$ and in uplink subframe number $\left\lfloor \frac{m}{zn} \right\rfloor + 1$, which immediately follows. In downlink subframe number $\left\lfloor \frac{m}{zn} \right\rfloor$, all NDs have been granted an uplink RB allocation except ND $i$. Thus, ND $i$ must be granted a downlink RB allocation in this subframe to have received transmissions from all other NDs. This occurs with probability $\binom{m-1}{zn-1}/\binom{m}{zn}$. Similarly, ND $i$ must receive an uplink RB allocation in the following uplink subframe so all remaining $m - 1$ NDs can receive from ND $i$ in the following $\left\lfloor \frac{m}{zn} \right\rfloor$ subframes. This also occurs occurs with probability $\binom{m-1}{zn-1}/\binom{m}{zn}$. Thus, the total probability of achieving a complete affine graph in this subcase is given by

$$\mathbb{P}\left\{F_{min} = 2 \left\lfloor \frac{m}{zn} \right\rfloor\right\} = \frac{\binom{m-1}{zn-1}^2}{\binom{m}{zn}^{2\alpha+1}} \prod_{i=1}^{\alpha-1} \binom{m-izn}{zn} \prod_{j=0}^{\alpha-1} \binom{m-1-jzn}{zn},$$  (5.69)

where, on the right side of the equation, all denominators of the product are collected into the denominator of the first term, the second and third terms are the decrementing hypergeometric distributions for the uplink and downlink, respectively, and $\alpha = \left\lfloor \frac{m}{zn} \right\rfloor$.

**Subcase 3.3:** $m \mod zn > 1$

This occurs when any multiple of the product of RBs and the overloading ratio does not fall in either Subcase 3.1 or 3.2. The minimum number of frames required to achieve a complete affine graph is given by

$$F_{min} = 2 \left\lceil \frac{m}{zn} \right\rceil - 1. \tag{5.70}$$

The probability of achieving the affine graph in the minimum number of frames can be directly calculated, but does not have a concise expression. Rather, all possible ways in which all NDs are granted a RB allocation in the minimum number of frames must be enumerated, the probability of each of these possibilities calculated, and the pairwise multiplication of all probabilities must be summed to account for all possible uplink and downlink combinations of RB allocations.

## 5.7  Summary

In this chapter, the mathematical model of connectivity for the NOMA graph model and network ensemble were presented, and stochastic expressions for temporal connectivity were developed. The key results include the:

- Probability of Temporal Connectivity in Each Frame
- Probability of Time to Initial Unidirectional Connectivity
- Probability of Time Window Between Unidirectional Connectivity
- Probability of Time to Initial Bidirectional Connectivity
- Probability of Minimum Time to Complete Bidirectional Connectivity.

These results provide the basis for a stochastic characterization of robustness, measured through time-varying connectivity, in a NOMA wireless network with random and uniform RB allocation across all NDs.

# CHAPTER 6:
# Simulation Results

This chapter discusses the simulations conducted to test the internal validity of the analytical results developed in Chapter 5.

The text includes expanded versions of previously published material in

> B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal multiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

and

> B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917. © 2022 IEEE.

All previously published tables from these two publications are credited with a citation in the caption.

## 6.1   Simulation Environment and Analysis Methods

All simulations for this research are conducted in the MATLAB® technical computing environment.[7] Given a set of NOMA network parameters ($m$ NDs, $n$ RBs, and an overloading ratio $z$), the network ensemble is generated as the set of all unique binary degree vectors, and temporal network sequences are generated through random uniform sampling from the network ensemble, as described in Chapter 3.2. Connectivity between NDs resulting from the randomly sampled temporal network sequence is calculated according to the mathematical model of connectivity in Chapter 5.1.

---

[7]Information about obtaining the MATLAB® code and simulation data is available in the supplemental.

The error between the analytical and simulation results for probabilistic connectivity in each frame is measured by the mean absolute deviation (mAD) across all simulated frames, and the maximum absolute deviation (MAD) of any simulated frame [96]. The mAD and MAD error statistics are appropriate because they characterize the number of links that are incorrectly predicted by the analytical results. This includes links that are predicted but do not occur, and links that exceed the number predicted. The error between the analytical and simulation results for the time window between directed connectivity and minimum time to complete bidirectional connectivity are discussed in Sections 6.4 and 6.6, respectively.

Finally, for the analytical results from Chapter 5.3–5.5 that produce PMFs, the similarity between the analytical and simulated probability distributions is measured using the Jensen-Shannon divergence (JSD) [103], given by

$$\text{JSD}(p_1||p_2) = \frac{1}{2}\text{D}_{KL}(p_1||M) + \frac{1}{2}\text{D}_{KL}(p_2||M), \tag{6.1}$$

where $p_1$ and $p_1$ are discrete probability distributions, $M = (p_1 + p_2)/2$, and $\text{D}_{KL}$ is the Kullback-Leibler divergence defined as

$$\text{D}_{KL}(p_1||p_2) = \sum_{k \in \mathcal{K}} p_1(k) \log_2(p_1(k)/p_2(k)). \tag{6.2}$$

The JSD measure subtracts the individual Shannon entropy of each distribution from the entropy of the mixture of the distributions. Thus, the JSD is equal to zero if the two distributions are identical, and small JSD values indicate highly similar distributions.

## 6.2   Directed Connectivity in Each Frame

The probability of directed connectivity in each frame given in Equation (5.29) is tested across a range of network sizes and overloading ratios. The simulation networks are composed of $m \in \{100, 1000, 10000\}$ NDs, $n = 0.1m$ RBs, and overloading ratios $z \in \{2, 3, 4\}$, where each network size is simulated with each overloading ratio. We found the overloading ratio $z = 2$ as the most common in our literature review (with $z = 3$ as the highest observed), so this research considers $z \in \{2, 3, 4\}$. Network sequences of 100 frames are generated for each set of network parameters through uniform random sampling from the network ensemble. The number of directed links from ND $i$ to ND $j$ occurring in each frame are recorded,

and this process is repeated 1,000 times for each set of network parameters. Experimental probabilities for each frame are generated by averaging across all trials for each frame and normalizing by $m(m-1)$ possible links.

The error results for unidirectional connectivity in each frame are summarized in Table 6.1. The mAD and MAD error statistics both firmly support the accuracy of the analytical predictions with no simulation error exceeding .02%. The JSD is not included for these simulations because the results are not probability distributions, and the JSD is a measure of distribution similarity.[8]

In more practical terms, consider the number of incorrectly predicted links by multiplying the mAD and MAD by the number of potential $i$-$j$ links in each network, $m(m-1)$. Doing so shows that, on average, the analytical prediction errs by approximately 2 out of 9,900 links per frame for the network of 100 NDs, and no error is greater than 8 links. Similarly, for the network of 10,000 NDs, the largest mean error is 2,090 links, and the maximum error is less than 7,090 out of 99,990,000 links per frame.

Table 6.1. Simulation Error Results for the Probability of Directed Connectivity in Each Frame.

| Network Parameters | | Error | |
|---|---|---|---|
| NDs ($m$) | Overloading ($z$) | Mean AD | Max. AD |
| | 2 | $1.61 \times 10^{-4}$ | $5.69 \times 10^{-4}$ |
| 100 | 3 | $1.74 \times 10^{-4}$ | $4.99 \times 10^{-4}$ |
| | 4 | $1.95 \times 10^{-4}$ | $7.64 \times 10^{-4}$ |
| | 2 | $4.76 \times 10^{-5}$ | $1.90 \times 10^{-4}$ |
| 1,000 | 3 | $5.74 \times 10^{-5}$ | $2.90 \times 10^{-4}$ |
| | 4 | $5.24 \times 10^{-5}$ | $1.53 \times 10^{-4}$ |
| | 2 | $1.60 \times 10^{-5}$ | $6.40 \times 10^{-5}$ |
| 10,000 | 3 | $1.54 \times 10^{-5}$ | $5.83 \times 10^{-5}$ |
| | 4 | $2.09 \times 10^{-5}$ | $7.80 \times 10^{-5}$ |

---

[8]However, each individual frame can be considered a Bernoulli PMF parameterized by $p = \mathbb{P}\left\{F_{ij}^k = 1\right\}$.

A visual comparison of the analytical and simulation results are shown in Figures 6.1 and 6.2. The probabilities predicted by Equation (5.29) are denoted by blue circles, and the experimental probabilities recorded from the simulation are denoted by red "x's."
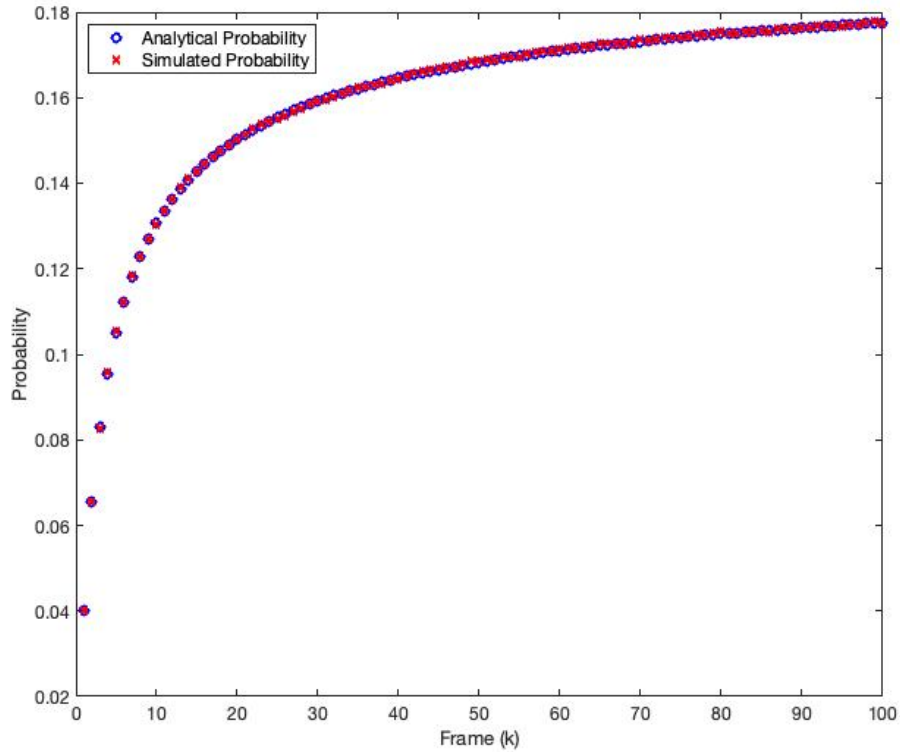


Figure 6.1. Probability of directed connectivity in the first 100 frames of a temporal network sequence. The network includes 100 NDs, 10 RBs, and an overloading ratio $z = 2$.
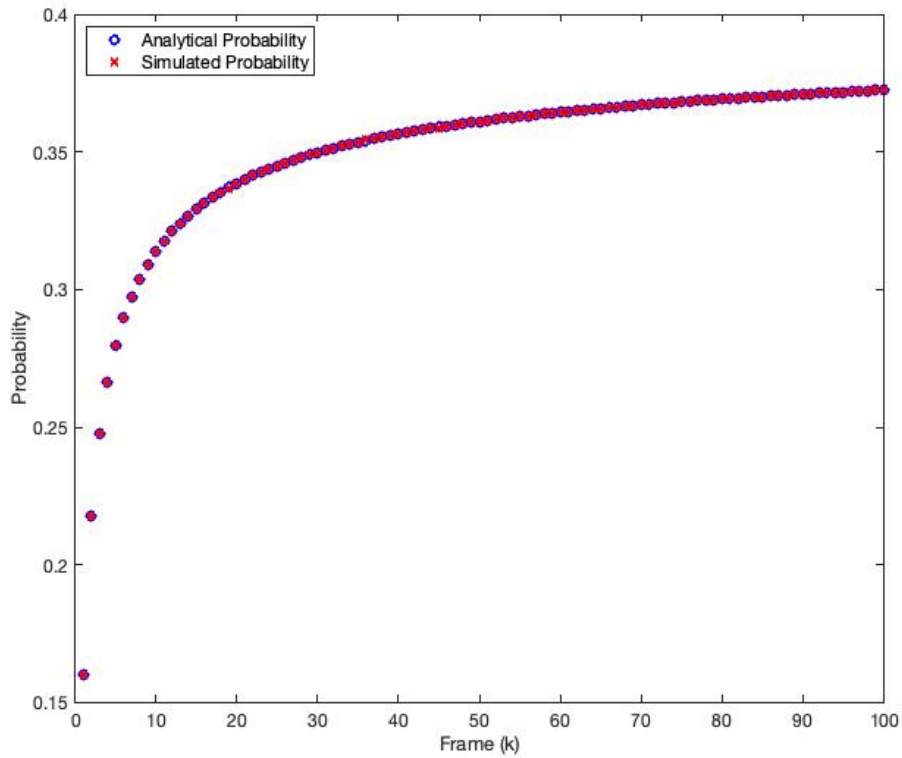
Figure 6.2. Probability of directed connectivity in the first 100 frames of a temporal network sequence. The network includes 10,000 NDs, 1,000 RBs, and an overloading ratio $z = 4$.

### 6.2.1 Error Analysis

As expected, all measures of error decrease as the network size increases. This is due the increased $i$-$j$ link sample size in each frame. The 100 ND network only has 9,900 potential links to measure in each frame, whereas the 10,000 ND network has 99,990,000 potential links per frame. By the law of large numbers, the sample mean converges to the distribution mean as the sample grows large which, in this case, is the probability of directed connectivity in each frame [84].

Perhaps counter-intuitively, the mAD and MAD both generally increase as $z$ increases within each fixed network size $m$. This is a result of the increased variability, or uncertainty, in

95

RB allocation caused by the composition of the network ensemble. The uncertainty can be quantified by Shannon entropy, $H$, defined as

$$H = -\sum_i p_i \log p_i, \tag{6.3}$$

where $p_i$ is the probability of each event [53].

Recall from Chapter 3.2 that the network ensemble, $D$, is composed of all unique binary degree vectors with $zn$ ones and $m-zn$ zeros, and the probability of random uniform selection of any degree vector from the ensemble is equal to $|D|^{-1}$. Substituting the probability of degree vector selection into Equation (6.3), the entropy of the distribution of degree vectors is given by

$$H = -\sum_1^{|D|} |D|^{-1} \log |D|^{-1} = -\log \frac{1}{|D|}. \tag{6.4}$$

It is clear from Equation (6.4) that the entropy increases as the cardinality of $D$ increases. Recall from Equation (3.5) that the cardinality of $D$ is equal to the binomial coefficient $\binom{m}{zn}$, which reaches a maximum when $zn = m/2$. Therefore, the entropy of the uniform random selection of degree vectors from the network ensemble is maximized when $zn = m/2$. A maximum entropy network ensemble affords each ND a 50% probability of RB allocation in each frame, resulting in an equal probability that NDs receive equal or disproportionate RB allocations over time. Thus, the probability that individual frames in the simulation return a number of links that is farther from the analytically predicted value increases as the network ensemble approaches maximum entropy. This is reflected in the mAD and MAD statistics as $m$ and $n$ are held constant for increasing $z$.

## 6.3   Time to Initial Unidirectional Connectivity

The probability of initial directed connectivity in each frame given in Equation (5.47) is tested across a range of network sizes and overloading ratios. Similar to the simulation networks for directed connectivity in each frame, the simulation networks for initial directed connectivity are composed of $m \in \{100, 1000, 10000\}$ NDs, $n = 0.1m$ RBs, and overloading ratios $z \in \{2, 3, 4\}$, where each network size is simulated with each overloading ratio. In contrast to the simulation networks for directed connectivity in each frame, network sequences of

$N \in \{100, 60, 45\}$ frames are generated for each set of network parameters through uniform random sampling from the network ensemble where $(z, N) \in \{(2, 100), (3, 60), (4, 45)\}$.[9] The number of directed links from ND $i$ to ND $j$ occurring for the first time in each frame (1 to $N$) are recorded and this process is repeated 1,000 times for each set of network parameters. Experimental PMFs are generated by averaging across all trials for each frame and normalizing by $m(m-1)$ possible links.

The error and distribution similarity results for initial unidirectional connectivity are summarized in Table 6.2. The error statistics behave in a similar manner to the error of directed connectivity in each frame discussed in Section 6.2. The error decreases as the network size increases, but the mAD and MAD increase with $z$ within each fixed network size. However, the JSD decreases as overloading increases within each network size because the probability of initial unidirectional connectivity between each ND pair converges to zero in a smaller number of frames due to the increased availability of transmission resources (i.e., the number of frames over which the PMF is evaluated decreases as $z$ increases, so there are fewer opportunities for dissimilarity).

In terms of the number of incorrectly predicted links, the analytical PMF errs by approximately 1 out of 9,900 links in each frame for the network of 100 NDs, and no error is greater than 12 links. Similarly, for the network of 10,000 NDs, the largest mean error is 930 links, and the maximum error is less than 8,002 out of 99,990,000 potential links in each frame. The error results and distribution similarity analysis lend strong support for the accuracy of the analytical PMF.

Plots comparing the predicted and simulation-generated PMFs across all tested overloading ratios for a 10,000 ND network are depicted in Fig. 6.3–6.5. In each plot, the analytical PMF for the probability of initial directed connectivity is denoted by blue stems, and the simulation PMF is indicated by red "x's.". Support for the accuracy of the analytical results is visually apparent from the close alignment of the PMFs in Figures 6.3–6.5.

---

[9]The number of frames decreases as overloading increases because fewer frames are required for the distribution tail to converge to zero.

Table 6.2. Simulation Error and Distribution Similarity for Initial Unidirectional Connectivity. Source: [76]. © 2022 IEEE

| Network Parameters | | Error | | |
|---|---|---|---|---|
| NDs ($m$) | Overloading ($z$) | Mean AD | Max. AD | JSD |
| | 2 | $6.54 \times 10^{-5}$ | $1.03 \times 10^{-3}$ | $2.66 \times 10^{-5}$ |
| 100 | 3 | $7.57 \times 10^{-5}$ | $7.47 \times 10^{-4}$ | $2.48 \times 10^{-5}$ |
| | 4 | $1.29 \times 10^{-4}$ | $1.17 \times 10^{-3}$ | $2.34 \times 10^{-5}$ |
| | 2 | $1.99 \times 10^{-5}$ | $1.62 \times 10^{-4}$ | $2.94 \times 10^{-6}$ |
| 1,000 | 3 | $2.47 \times 10^{-5}$ | $2.01 \times 10^{-4}$ | $2.23 \times 10^{-6}$ |
| | 4 | $2.75 \times 10^{-5}$ | $2.69 \times 10^{-4}$ | $2.16 \times 10^{-6}$ |
| | 2 | $7.51 \times 10^{-6}$ | $7.46 \times 10^{-5}$ | $3.15 \times 10^{-7}$ |
| 10,000 | 3 | $7.75 \times 10^{-6}$ | $8.00 \times 10^{-5}$ | $2.47 \times 10^{-7}$ |
| | 4 | $9.30 \times 10^{-6}$ | $7.73 \times 10^{-5}$ | $1.20 \times 10^{-7}$ |



Figure 6.3. Comparison of analytical and simulation PMFs for initial unidirectional connectivity with 10,000 NDs, 1,000 RBs, and overloading $z = 2$.
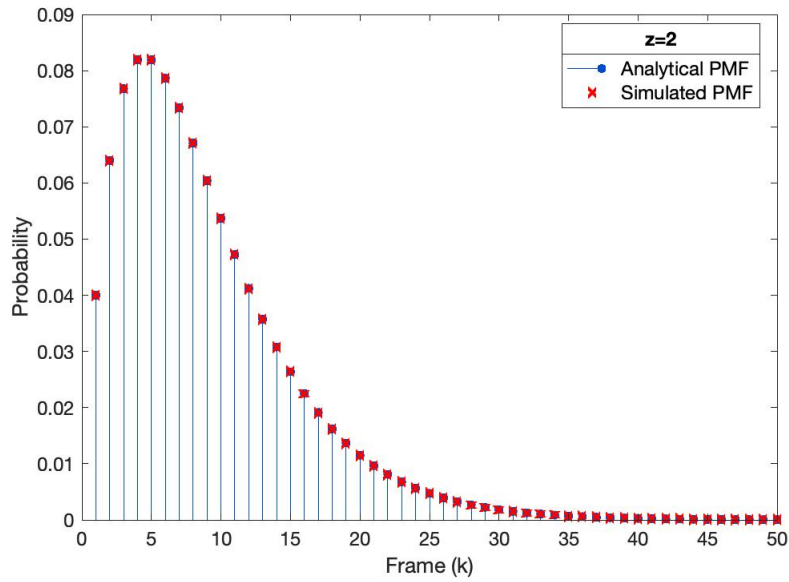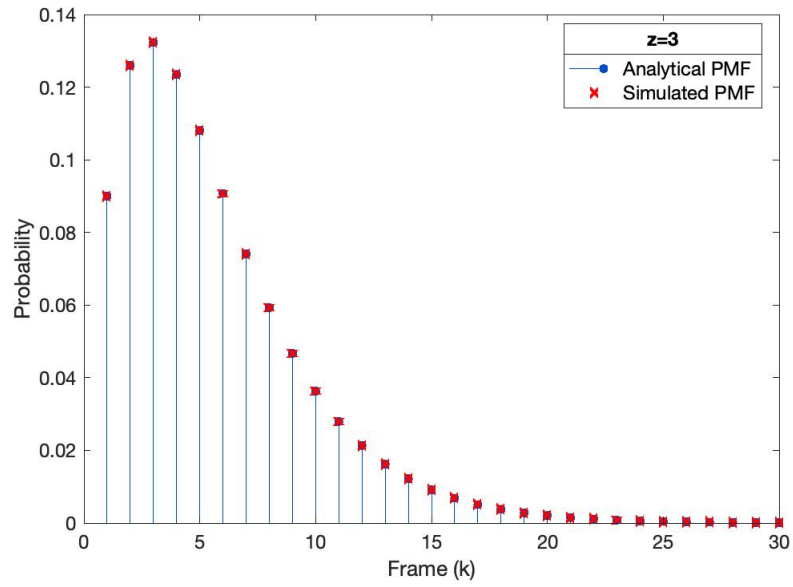
Figure 6.4. Comparison of analytical and simulation PMFs for initial unidirectional connectivity with 10,000 NDs, 1,000 RBs, and overloading $z = 3$.
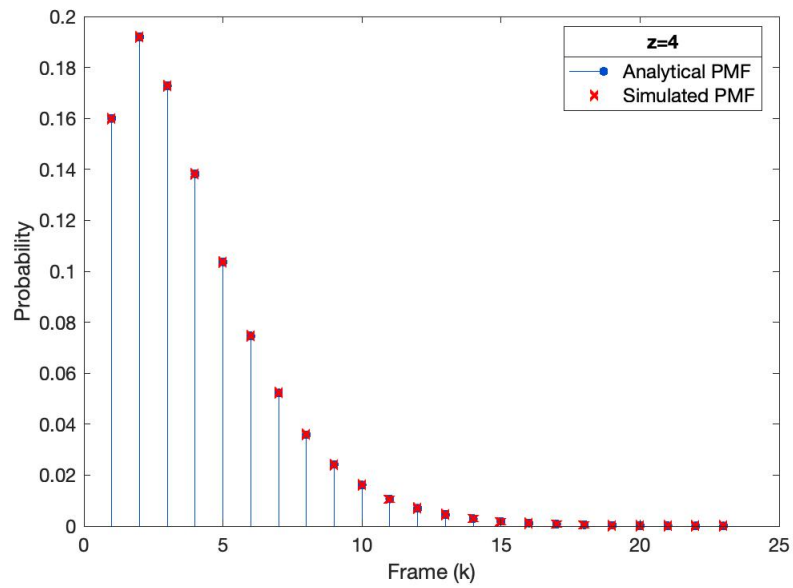


Figure 6.5. Comparison of analytical and simulation PMFs for initial unidirectional connectivity with 10,000 NDs, 1,000 RBs, and overloading $z = 4$.

## 6.4　Time Window between Unidirectional Connectivity

The probability of the number of frames between directed connectivity from ND $i$ to ND $j$ (i.e., time window or inter-event time) specified in Equation (5.50) is tested across a range of network sequence lengths and overloading ratios. The simulation networks are composed of $m = 1000$ NDs, $n = 0.1m$ RBs, overloading ratios $z \in \{2, 3, 4\}$, and network sequence lengths $N \in \{1000, \ 5000, \ 10000, \ 20000\}$ frames, where each network sequence length is simulated with each overloading ratio. The network size is held constant at 1,000 NDs to produce a meaningful sample, but ensure a tractable computation time.

The number of frames between each occurrence of directed connectivity from ND $i$ to ND $j$ are recorded, and a PMF is generated by normalizing the total count of each window size by the total number of inter-event times. Each of the network sequences is only run once due to the large sample size produced in one simulation run. The smallest number of inter-event data points produced is $185, 209, 137$ for $(z, N) = (2, 1000)$, which is more than sufficient to generate a probability distribution. The number of inter-event data points only increases as the network sequence lengthens, and the overloading ratio increases, resulting in the simulation with $(z, N) = (4, 20000)$ producing $7,913,212,170$ inter-event data points.

The error for the time window between directed connectivity events is considered through the mAD, root mean squared error (RMSE), and JSD. Unlike the PMFs describing the probability of initial connectivity in each frame, the maximum absolute deviation is not considered because the inter-event times are taken over the entire network sequence rather than the possible number of links each in frame. Thus, the maximum absolute deviation has less meaning since it does not convey the largest link prediction error of any frame when applied to the distribution of inter-event times. The mean absolute deviation is still considered as a complement to the RMSE to measure overall distribution error.

The error and distribution similarity results for the time window between directed connectivity events are summarized in Table 6.3. These results show decreasing error and increasing distribution similarity as the network sequence length grows large, which provides compelling support for the asymptotic predictions in Equation (5.50). These results are similar to the previous error results in Sections 6.2 and 6.3 where overall error decreases as the sample size (network sequence length in this case) grows large, but the error statistics increase with $z$ for fixed network length. This results from the network ensemble entropy increasing with

$z$. Recall from Chapter 5.4 that the inter-event times are directly related to the probability of connectivity in each frame. Thus, as the frame connectivity error increases with the rising network ensemble entropy, so will the inter-event time error since it depends on the frame connectivity events. Conversely, the JSD decreases with increasing $z$ because the sample size of inter-event times increases. The JSD measures overall distribution similarity, so it is not as susceptible to a skewed result as the mAD and RMSE.

Table 6.3. Simulation Error and Distribution Similarity for the Time Window between Unidirectional Connectivity.

| Network Parameters | | Error | | |
|---|---|---|---|---|
| Frames ($k$) | Overloading ($z$) | Mean AD | RMSE | JSD |
| | 2 | $6.81 \times 10^{-4}$ | $1.89 \times 10^{-3}$ | $9.33 \times 10^{-4}$ |
| 1,000 | 3 | $7.92 \times 10^{-4}$ | $2.37 \times 10^{-3}$ | $6.45 \times 10^{-4}$ |
| | 4 | $1.07 \times 10^{-3}$ | $2.97 \times 10^{-3}$ | $4.98 \times 10^{-4}$ |
| | 2 | $2.67 \times 10^{-4}$ | $8.04 \times 10^{-4}$ | $2.19 \times 10^{-4}$ |
| 5,000 | 3 | $3.79 \times 10^{-4}$ | $1.15 \times 10^{-3}$ | $1.47 \times 10^{-4}$ |
| | 4 | $4.99 \times 10^{-4}$ | $1.41 \times 10^{-3}$ | $1.11 \times 10^{-4}$ |
| | 2 | $1.92 \times 10^{-4}$ | $5.76 \times 10^{-4}$ | $1.11 \times 10^{-4}$ |
| 10,000 | 3 | $2.73 \times 10^{-4}$ | $8.09 \times 10^{-4}$ | $7.48 \times 10^{-5}$ |
| | 4 | $3.01 \times 10^{-4}$ | $9.22 \times 10^{-4}$ | $5.52 \times 10^{-5}$ |
| | 2 | $1.35 \times 10^{-4}$ | $4.17 \times 10^{-4}$ | $5.51 \times 10^{-5}$ |
| 20,000 | 3 | $1.64 \times 10^{-4}$ | $5.11 \times 10^{-4}$ | $3.86 \times 10^{-5}$ |
| | 4 | $2.27 \times 10^{-4}$ | $6.59 \times 10^{-4}$ | $2.84 \times 10^{-5}$ |

Plots of the geometric distribution for network sequences of length $N = 1,000$ frames and $N = 20,000$ frames are shown in Figures 6.6 and 6.7, respectively. Notice the closer alignment of the simulation and analytical PMFs when the network sequence is 20,000 frames (Figure 6.7) versus 1,000 frames (Figure 6.6). This clearly demonstrates the convergence to the predicted geometric distribution as the network sequence length goes to infinity.

Figure 6.6. Probability of the number of frames ($k$) between directed connectivity from ND $i$ to ND $j$ for a network with 1,000 NDs, 100 RBs, overloading ratio $z = 2$, and network sequence length $N = 1,000$ frames.
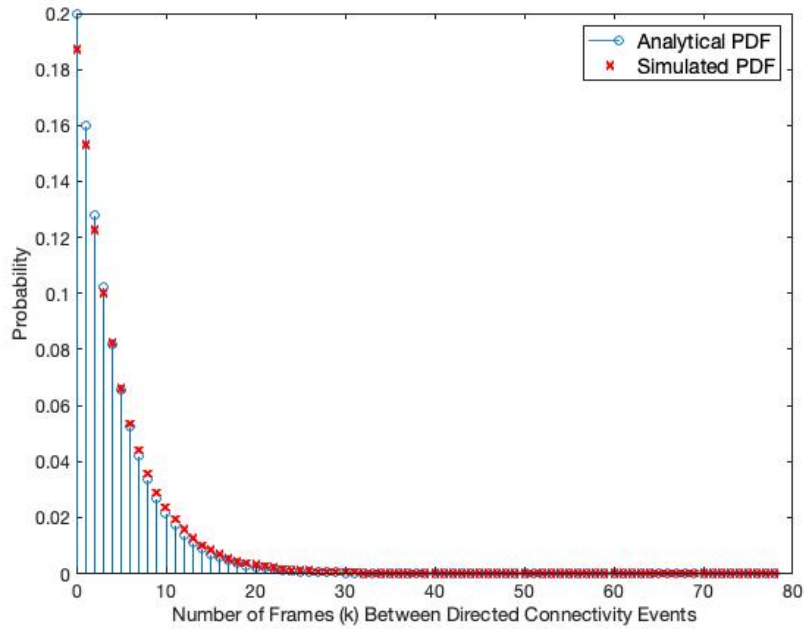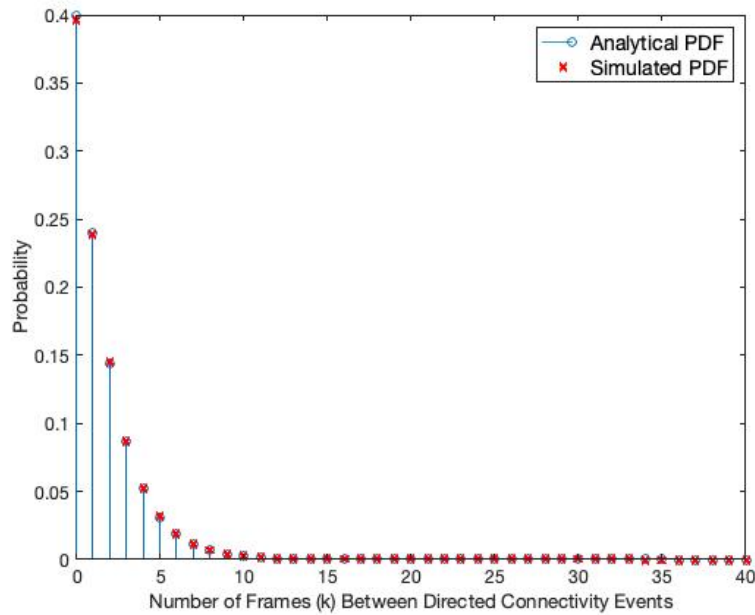


Figure 6.7. Probability of the number of frames ($k$) between directed connectivity from ND $i$ to ND $j$ for a network with 1,000 NDs, 100 RBs, overloading ratio $z = 4$, and network sequence length $N = 20,000$ frames.

## 6.5 Time to Initial Bidirectional Connectivity

The probability of initial bidirectional connectivity in each frame given in Equation (5.61) is tested across the same range of network sizes and overloading ratios as initial unidirectional connectivity in Section 6.3. The simulation networks for initial bidirectional connectivity are composed of $m \in \{100, 1000, 10000\}$ NDs, $n = 0.1m$ RBs, and overloading ratios $z \in \{2, 3, 4\}$, where each network size is simulated with each overloading ratio. Network sequences of $N \in \{100, 60, 45\}$ frames are generated for each set of network parameters through uniform random sampling from the network ensemble where $(z, N) \in \{(2, 100), (3, 60), (4, 45)\}$, and the number of bidirectional links from ND $i$ to ND $j$ occurring for the first time in each frame (1 to $N$) are recorded. This process is repeated 1,000 times for each set of network parameters, and experimental PMFs are generated by averaging across all trials for each frame and normalizing by $m(m - 1)$ possible links.

The error and distribution similarity results for initial bidirectional connectivity are summarized in Table 6.4. The error trends behave in a similar manner to those in the case of initial unidirectional connectivity in Section 6.3. The error statistics decrease as the network size increases, but increase for a fixed network size and increasing overloading. However, the error is larger than the unidirectional case since the PMF incorporates the variability of both the $i$-$j$ and $j$-$i$ links, rather than just the $i$-$j$ links (as in the unidirectional case). Also similar to initial unidirectional connectivity, the JSD decreases as overloading increases within each network size because the probability of initial bidirectional connectivity between each ND pair converges to zero in a smaller number of frames.

Overall, the simulation error and distribution similarity results provide firm support for the accuracy of analytical PMF. Again, in practical terms, the analytical PMF errs by 1 to 3 links out of 9,900 in each frame for the network of 100 NDs, and no error is greater than 24 links. Likewise, the largest mean error for the 10,000 ND network is 1,521 links, and the maximum error is 14,820 out of 99,990,000 potential links in each frame.

A comparison of the analytical and simulated PMFs for all overloading ratios in a 100 ND network is depicted in Figures 6.8-6.10. This increased error, with respect to the case of initial unidirectional connectivity, is visually apparent in the frames where the analytical and simulation results are misaligned, such as frame $k = 8$ in Figure 6.8.

Table 6.4. Simulation Error and Distribution Similarity for Initial Bidirectional Connectivity.

| Network Parameters | | Error | | |
|---|---|---|---|---|
| NDs ($m$) | Overloading ($z$) | Mean AD | Max. AD | JSD |
| | 2 | $1.09 \times 10^{-4}$ | $1.96 \times 10^{-3}$ | $5.74 \times 10^{-5}$ |
| 100 | 3 | $1.54 \times 10^{-4}$ | $9.98 \times 10^{-4}$ | $5.35 \times 10^{-5}$ |
| | 4 | $2.70 \times 10^{-4}$ | $2.38 \times 10^{-3}$ | $5.32 \times 10^{-5}$ |
| | 2 | $3.12 \times 10^{-5}$ | $3.23 \times 10^{-4}$ | $5.90 \times 10^{-6}$ |
| 1,000 | 3 | $3.57 \times 10^{-5}$ | $3.77 \times 10^{-4}$ | $4.72 \times 10^{-6}$ |
| | 4 | $5.83 \times 10^{-5}$ | $4.01 \times 10^{-4}$ | $4.69 \times 10^{-6}$ |
| | 2 | $1.19 \times 10^{-5}$ | $1.31 \times 10^{-4}$ | $6.10 \times 10^{-7}$ |
| 10,000 | 3 | $1.33 \times 10^{-5}$ | $1.16 \times 10^{-4}$ | $4.81 \times 10^{-7}$ |
| | 4 | $1.52 \times 10^{-5}$ | $1.48 \times 10^{-4}$ | $2.47 \times 10^{-7}$ |



Figure 6.8. Comparison of analytical and simulation PMFs for initial bidirectional connectivity with 100 NDs, 10 RBs, and overloading ratio $z = 2$.
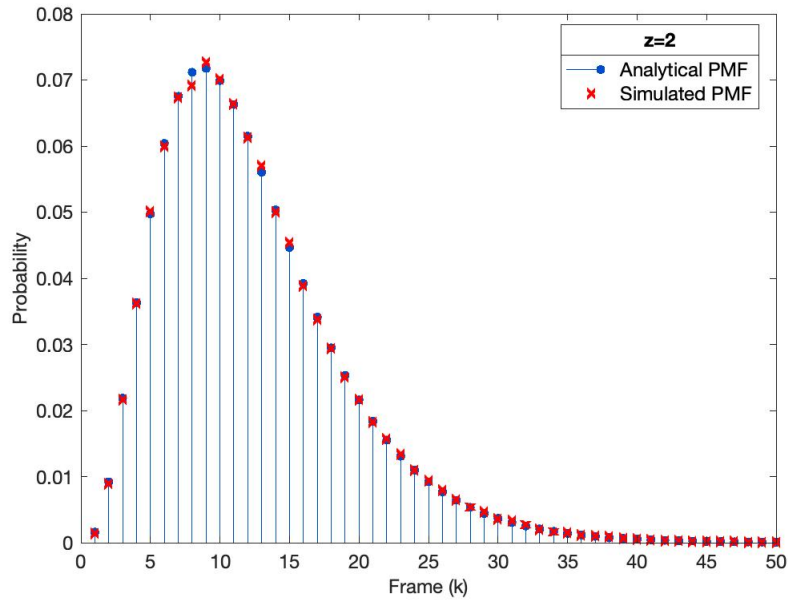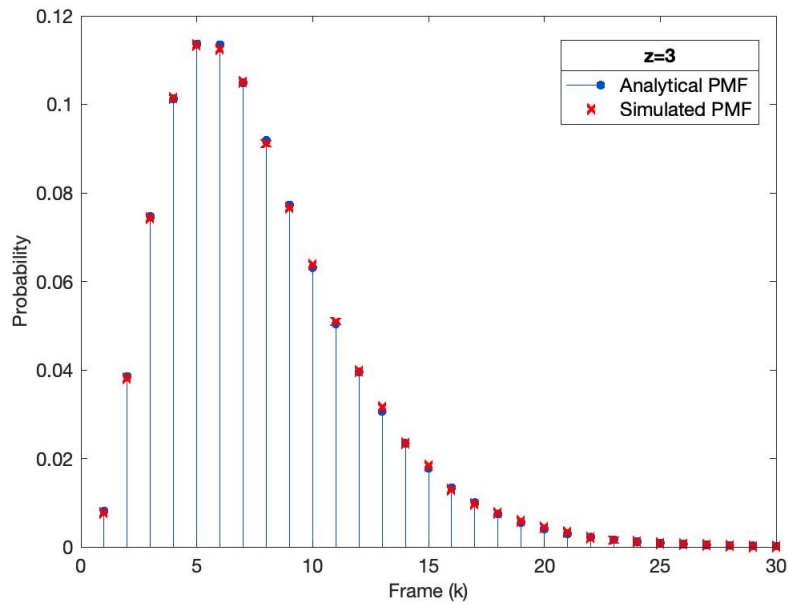
Figure 6.9. Comparison of analytical and simulation PMFs for initial bidirectional connectivity with 100 NDs, 10 RBs, and overloading ratio $z = 3$.
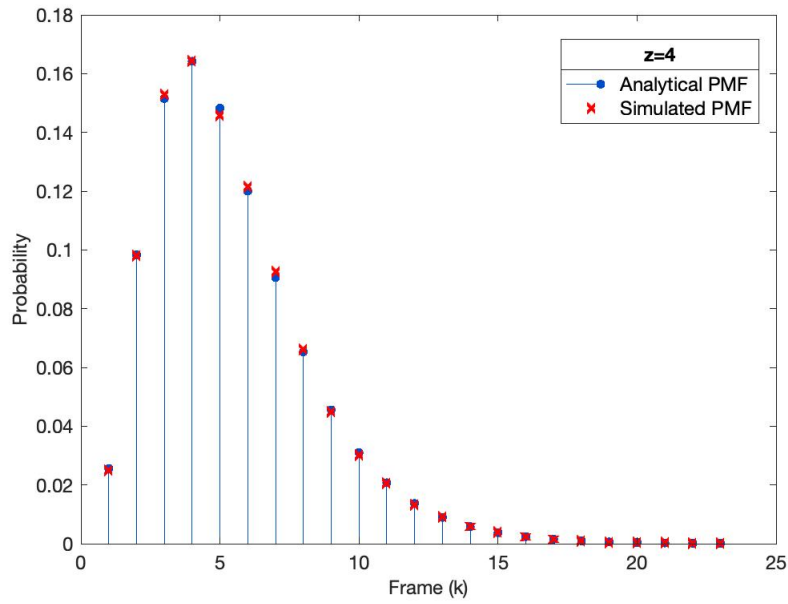


Figure 6.10. Comparison of analytical and simulation PMFs for initial bidirectional connectivity with 100 NDs, 10 RBs, and overloading ratio $z = 4$.

## 6.6 Min. Time to Complete Bidirectional Connectivity

The probability for the minimum time to complete bidirectional connectivity (i.e., the minimum time to achieve a complete affine graph) is tested for each case with a defined expression for $\mathbb{P}\{F_{min}\}$ discussed in Chapter 5.6. NOMA network parameters ($m$, $n$, and $z$) that satisfy each case conditions are selected to generate the corresponding temporal network ensemble. Next, a network sequence of length $N$ is generated through random uniform sampling from the ensemble where $N$ is set equal to $F_{min} \in \{2, 3, 4, 5\}$ based on the corresponding case. The resulting affine graph from each network sequence is calculated using the mathematical framework of Chapter 5.1, 5.5, and Equation (5.62).

A total of $1 \times 10^8$ network sequences are generated for each case and the number of times a complete affine graph occurs is recorded. The number of complete affine graph occurrences is normalized by the number of trials, and the experimental results are compared to the analytical results using the absolute percent error as $|E - P|/E$, where $E$ is the experimental result, and $P$ is the analytical result. The cases, network parameters, and error results are shown in Table 6.5.

All errors except one are less than 1%, which lends considerable support for the accuracy of the analytical results. The largest error is 5.33% and corresponds to the lower bound of Case 2, in which exactly half of the NDs can receive a RB allocation in each frame. This is a maximum entropy ensemble with $\mathbb{P}\{F_{min} = 3\} = 1.57 \times 10^{-5}$ (i.e., a sample of $1 \times 10^8$ network sequences is only expected to produce 1,575 successes), so a larger error is expected. A separate simulation of $1 \times 10^9$ network sequences is generated for this case, with a resulting absolute percent error of $1.33 \times 10^{-3}$. This 40-fold error reduction for a 10-fold sample size increase also supports the accuracy of the analytical results.

Table 6.5. Simulation Error Results for Minimum Time to Complete Bidirectional Connectivity.

| Case | $N$ | NDs ($m$) | RBs ($n$) | $z$ | Absolute Percent Error |
|------|-----|-----------|-----------|-----|------------------------|
| 1 | 2 | 9 | 2 | 4 | $4.90 \times 10^{-4}$ |
| 2 | 3 | 10 | 2 | 4 | $1.50 \times 10^{-4}$ |
| 2 | 3 | 10 | 2 | 2.5 | $5.33 \times 10^{-2}$ |
| 3.1 | 5 | 6 | 2 | 1 | $3.15 \times 10^{-3}$ |
| 3.2 | 4 | 9 | 2 | 2 | $1.12 \times 10^{-3}$ |

## 6.7  Summary

This Chapter discussed the simulations conducted to test the analytical derivations from Chapter 5, and the corresponding results. The simulation environment, parameters, and selected error measures were presented. Error analysis was offered for each simulation, and an entropy-based explanation for the difference in PMF error results was proposed. Additionally, a characterization of the error in terms of link presence was provided for all probabilistic connectivity results. The simulation results strongly support the internal validity of all theoretical developments, and provide a sound basis for continued research.

THIS PAGE INTENTIONALLY LEFT BLANK

# CHAPTER 7:
## Conclusions

This research effort set out to understand the effect of variable overloading on the robustness of NOMA wireless networks. Like so many other research endeavors, the direction of this research was guided by new, and sometimes unexpected, questions that arose along the way. However, all the variations remained true to the original focus on the relationship between overloading and robustness. The following sections review the contributions of this research, limitations, and questions that remain for future work.

## 7.1    Contributions

The research journey began by considering NOMA wireless networks through the theoretical perspective of temporal network theory, and grappling with the representation of a NOMA wireless network as a mixed dependency-connectivity graph. After distilling the mixed dependency-connectivity graph down to degree vectors that preserve connectivity and overloading information, the temporal network ensemble conceptualization of the NOMA wireless network was mathematically formalized.

Equipped with the temporal network ensemble, the research moved forward to recast the network evolution as a Bernoulli random process. In this stochastic context, the implications for the probability of RB allocation and the probability of temporal component membership were considered. The resulting stochastic temporal component framework was the first result of the research that began to get at the core of the research question, providing a measure of network robustness as function of overloading. This development led to a subsequent investigation into the minimum number of frames required for all NDs to join the temporal strongly connected component (i.e., form a complete affine graph), and the likelihood of this event. The knowledge gained from each of these efforts were combined to develop a conference paper accepted to the $55^{th}$ Hawaii International Conference on System Sciences [17]. However, the development of the temporal component framework led to new questions about the impact of queued transmissions on network connectivity, and the number of frames required to reach different states of connectivity between NDs.

These new questions shifted the research focus from temporal components to the probability of temporal connectivity between NDs. First, the probability of connectivity between any two distinct NDs in any frame was studied. This line of inquiry was not long underway before it became evident that queue dynamics were of central importance to the evolution of probabilistic temporal connectivity in each frame. Upon recognizing the queue lent itself to modeling as a random walk, the well-established analytical methods of Markov processes became available to employ. From this point, the research journey quickly navigated from exploring the probability of unidirectional connectivity in each frame, to the probability of initial unidirectional connectivity in each frame, to the probability of initial bidirectional connectivity in each frame, and finally to the time window between unidirectional connectivity events. The results achieved from each of these forays into the realm of unanswered questions provided new insight into the robustness of NOMA networks to variable overloading, and the first two of these efforts became the basis for another conference paper accepted to the 2022 IEEE International Conference on Communications [65].

As is said in the Marine Corps, "no plan survives first contact with the enemy," and this research was no different. The original research plan did not place such a strong emphasis on probabilistic connectivity between distinct ND pairs, but the questions that emerged during the research process led down that path. Ultimately, we submit that progress towards answering the research question was made, and new contributions to knowledge were achieved.

This research developed a novel NOMA temporal graph model that abstracts the physical NOMA implementation and is generalizable to any power-domain or code-domain physical approach. The model directly corresponds to a snapshot representation of a temporal network ensemble, thereby providing a temporal connectivity analysis tool for NOMA network designers. The stochastic temporal component framework developed in Chapter 4, the probabilities and PMFs derived in Chapter 5.2–5.5, and the expressions for complete bidirectional connectivity found in 5.6 all provide generalized measures of NOMA network robustness as a function of variable overloading. These results contribute to wireless network theory, temporal network theory, and provide a starting point for more idiographic analysis in the design of specific physical NOMA wireless network implementations.

## 7.2 Limitations

As discussed throughout the presentation of our work, there are a number of limitations to the results of this research.

First, though intentional, the highly generalizable nature of the NOMA graph model (discussed in Chapter 3.1.2) sacrifices fidelity when applied, without adaptation, to any specific TDD NOMA implementation. Analysis of the specific NOMA network under consideration must be completed to adapt the frame and subframe constructs for more accurate results.

Next, the model is limited in scope by the assumptions outlined in Chapter 3.2.1. Specifically, the model only considers a single BS, uniform overloading across all RBs, and assumes that all NDs have a requirement to transmit in each frame. The model does not consider multi-BS scenarios, device-to-device scenarios, non-uniform overloading among RBs, or ND traffic models for intermittent transmission.

Additionally, as discussed in Chapter 4.1.2, the random sampling process from the network ensemble results in a random RB allocation model, which is different from the prevalent channel-based dynamic scheduling of RBs. While we present the maximum fairness argument for mMTC and URLLC use-cases, this model still limits the applicability of the results, and does not consider cases in which mMTC and URLLC may overlap with eMBB use-cases that require channel-based scheduling.

Finally, the simulation results presented in Chapter 6 are graph-based simulations, and not based on models of wireless networks that simulate the RF environment and network protocol stack. Thus, the simulations provide support for the accuracy of the connectivity behavior of the evolving network graph (i.e., internal validity), but do not assess the relationship between the proposed model and practical network implementations (i.e., external validity).

Each of the research limitations discussed in Section 7.2 represent opportunities to improve and extend the results in future work, as discussed in Section 7.3.

## 7.3  Future Work

Throughout this research effort, questions were encountered for which analytical results have not yet been achieved. The following sections briefly review some ideas of how this research may be continued and extended to address the limitations discussed in Section 7.2.

### 7.3.1  PMF for Complete Bidirectional Connectivity

After delving into the question of how quickly bidirectional connectivity can be achieved between all NDs in a NOMA wireless network, it became apparent that the probability of this event is vanishingly small for all but Case 1 as the network size increases (notice the small number of NDs used for simulation in Table 6.5). Thus, a potentially more useful result to pursue is the expression for the PMF describing the frame, $k$, in which complete bidirectional connectivity is achieved. Some preliminary simulations results suggest that the generalized extreme value distribution or lognormal distribution may reasonably approximate the actual distribution. However, the derivation of an expression for the actual distribution in terms of the NOMA network parameters remains an open question.

### 7.3.2  Probability of Bidirectional Connectivity in Each Frame

Though the probability of *initial* bidirectional connectivity in each frame was derived, the probability of bidirectional connectivity in each frame was not. Similar to the analytical methods employed in Chapter 5, the general approach to this question considers all possible ways in which ND $i$ and ND $j$ can be bidirectionally connected in frame $T$. This can occur in one of three ways:

1. ND $i$ and ND $j$ both establish a directed link to the other in frame $T$
2. ND $i$ has accumulated more directed links to ND $j$ by frame $T$ than ND $j$ has accumulated to ND $j$, and ND $j$ establishes directed connectivity to ND $i$ in frame $T$
3. The inverse of #2 above.

Mathematically, these cases can be stated as

$$\mathbb{P}\left\{A_{ij}^T = 1\right\} = \mathbb{P}\left\{F_{ij}^T = 1,\ F_{ji}^T = 1\right\}\ \cup\ 2\left(\mathbb{P}\left\{\sum_{t=1}^T F_{ij}^t \geq \sum_{t=1}^T F_{ji}^t,\ F_{ji}^T = 1\right\}\right).$$

The probability that $F_{ij}^t = 1$ is already given in equation (5.29), but the probability that $\sum_{t=1}^T F_{ij}^t \geq \sum_{t=1}^T F_{ji}^t$ proved more difficult. This sub-problem is approached by considering the probability that $\sum_{t=1}^T F_{ij}^t = k$ (where $k$ is any positive integer). This appears like a sum of Bernoulli trials (for which the binomial distribution can be used to find the probability) except that the probability that $F_{ij}^t = 1$ changes in each frame. The problem of dynamic probabilities in each trial was approached using the method in [104], but the analytical and simulation results are not yet well-aligned. Thus, the probability of bidirectional connectivity in each frame is ripe for future consideration.

### 7.3.3  Probability of Time Window Between Bidirectional Connectivity

If the result for the probability of the time window between unidirectional connectivity events serves as a guide, finding the time window between bidirectional connectivity events hinges on answering the question posed in Section 7.3.2. The solution to the $\mathbb{P}\left\{A_{ij}^T = 1\right\}$ will likely shed significant light on the question of inter-event time. Thus, both could be considered together in future work.

### 7.3.4  New Models of Overloading, RB Allocation and ND Transmission

This research considered uniform overloading, a random RB allocation model, and a network in which all NDs had a requirement to transmit in each frame. Changes to any of these assumptions provides a wide area of new research.

Future work could reduce the abstraction from the physical NOMA implementation by considering RB allocation that is driven by a spatial distribution of NDs, and the corresponding physical effects on overloading variability (e.g., achievable overloading that varies across the available RF spectrum based on instantaneous channel conditions).

Additionally, a new parameter indicating the time-varying transmission requirements of the NDs could be introduced. This parameter could be driven by additional simulations of user traffic, or be defined according to a probability distribution.

### 7.3.5  High Fidelity Simulation and Physical Experimentation

As discussed in Chapter 3.1.2, this research intentionally developed the ideas with an eye toward generalized applicability. However, this comes at the cost of fidelity when applied

to any specific NOMA implementation. Future work could test the analytical results by associating the general TDD frame structure with a specific wireless network standard (e.g., 3GPP NR), and conducting high-fidelity simulations that account for the RB allocation process defined in that standard.

Similarly, as physical NOMA experimentation capabilities improve [43], the theoretical results in this work could be tested against experimental results from niche or standards-based NOMA implementations. Both high-fidelity simulation and physical experimentation will improve the research by revealing flaws in the model that can be rectified, or providing external validity.

### 7.3.6   Design and Creativity

The ideas for future research discussed in Section 7.3 only represent a small sample of an immense space. Whether one claims the title of engineer or scientist, I reassert that both are fundamentally engaged in the creative act of design. As Herbert Simon profoundly observed, the bridge between participants in disparate intellectual pursuits may be the recognition of the "common creative [design] activity in which they are both engaged" [5, p. 137]. In this light, the possibilities for future research within this space, and the potential for interdisciplinary collaboration, are only limited by the creativity of the researcher.

## 7.4   Final Thoughts

Nonorthogonal multiple access is a dynamic area of research that is developing to support future iterations of 5G and the next generation of 6G wireless networks. I hope the contributions of this work prove useful to other wireless network researchers, and support the integration of emerging wireless network capabilities into Marine Corps tactical networks. I would be especially pleased to learn that the theory developed in this research finds use in ways and disciplines that I never considered.

# APPENDIX: Marine Corps Relevance

The Marine Corps doctrine for command and control asserts that "no single activity in war is more important than command and control," and defines "the basic elements of our command and control system" as "people, information, and the command and control support structure" [105, pp. 35, 47]. Fundamentally, this research contributes to the understanding of *information* movement through an exploration of the underlying *command and control support structure*. The results find specific applicability in the context of Force Design 2030 and the emerging fifth generation (5G) of mobile wireless networks.

The future battlefield environment envisioned in Force Design 2030 is characterized by the proliferation of unmanned, autonomous, and intelligent systems across all warfighting functions and echelons of command [106]. Just as the explosion of intelligent devices designed to automate various aspects of civilian life has given rise to the commercial Internet of Things (IoT) [107], the complex ecosystem of sensors and actuators that will accomplish tasks ranging from intelligence collection to kinetic operations represents the evolution towards the Internet of Battlefield Things (IoBT) [108], [109].

## A.1   Operational Problem

Fifth generation (5G) mobile wireless networks provide the capability necessary to support the commercial IoT through a heavy reliance on beamforming technology that requires multiple input multiple output (MIMO) antenna arrays with large numbers of elements (i.e., massive MIMO (mMIMO)). However, this approach is ill-suited for the size, weight, and power (SWaP) constraints of future tactical wireless networks that may be partially or entirely composed of mobile access points (MAPs) or mobile base stations (MBSs), such as low-altitude unmanned aerial vehicles (UAVs) (see Chapter 1.3 for a technical example of mMIMO array requirements). Providing the capability of 5G wireless networks to support the IoBT requires the consideration of a different approach.

## A.2   Technical Approach

Nonorthogonal multiple access (NOMA) wireless networks provide an alternative path to

achieve 5G capabilities that do not impose the SWaP constraints of mMIMO. In contrast to previous generations of multiple access technology, NOMA allows multiple network users to simultaneously use the same transmission resources (e.g., electromagnetic spectrum (EMS), time slots, spatial beams) to send or receive information. As discussed in Chapter 1.3, this technique is called *overloading*, and it allows NOMA wireless networks to increase the total information capacity of the network, increase the number of devices able to access the network, use the EMS more efficiently, and reduce the time required to access the network.

## A.2.1 Example Scenario

Consider a future stand-in force (SIF), as defined in the Marine Corps Concept for Stand-in Forces, operating in Southeast Asia that is tasked with maintaining a large network of intelligence, surveillance, reconnaissance and targeting (ISR-T) sensors in order to "gain and maintain custody of potential targets," and contribute to the joint force "kill web" [110, p. 5]. Accomplishing this mission requires processing of sensor data and shared real-time situational awareness (SA). However, in accordance with [110], the SIF must also maintain a high level of mobility, minimize sustainability requirements, and maintain a low signature. As a result, the SIF foregoes centralized computing resources (i.e., servers) in favor of distributed computing among their tactical mobile devices.

After deciding on their sensing and computing architecture, the SIF focuses their planning on the communications network that must support it. The data-rate requirements for the large sensor network are generally low, but some produce high-resolution imagery and video. Additionally, the selection of a distributed computing model introduces additional communications overhead as computing tasks are segmented among the tactical mobile devices (e.g., performing federated learning algorithms, such as image recognition, on inputs from the sensor network where each tactical mobile device only stores a small portion of the image database [31]). Finally, integrating into the joint force kill-web imposes strict latency requirements to support fire control systems. Recognizing their need for a communications network that supports a large sensor network, and high-throughput/low latency for real-time SA/distributed computing, the SIF employs a 5G tactical communications network supported by MBSs mounted on small UAVs.

In contrast to other non-fixed infrastructure mobile wireless networks (e.g., a cell-on-

116

wheels [111]), the small UAV network reduces lift requirements, autonomously adapts MBS positions to ensure coverage as the SIF maneuvers, and minimizes persistent emissions from the same location that support adversary targeting. However, the small UAV size cannot accommodate the typical 5G mMIMO antenna arrays at the lower frequency ranges that ensure communications in non-line-of-sight situations. Instead, the SIF employs NOMA technology and a smaller MIMO array to handle the device density of the sensor network, and throughput/latency requirements of the distributed computing applications.

## A.3   Research Relevance

The overloading ratio is a critical network parameter that must be considered before deploying a network using NOMA technology. The amount of overloading achieved impacts the throughput, latency, and connection density of the network. The throughput and latency are important for the SIF distributed computing applications that must support real-time SA and post-processing of sensor data. The connection density is important for the large sensor network the SIF must maintain. Correspondingly, the impact of changes in the overloading ratio must be understood to design a network that effectively supports all requirements.

The overloading ratio may change as the SIF executes their scheme of maneuver and the MBSs change locations to maintain network coverage. The EMS conditions at each location are different (e.g., different levels of interference, signal absorption by natural or artificial terrain), and the spatial distribution of the connected devices is different. Both of these factors affect the achievable overloading ratio. Rather than considering methods to maintain a specific overloading ratio, this research assumes it will change over the course of a mission, and investigates the resulting impact to time-varying network connectivity. This research provides a generalized model and analytical tools (in terms of the overloading ratio and other network parameters) to help network designers answer questions such as:

1. How likely are two devices to communicate with each other at a specific time during the evolution of the network?
2. How long does it take for device 1 to receive a message from device 2 for the first time? What is the probability that it takes that long?
3. After device 1 receives a message from device 2, how much time passes before device 1 receives another message from device 2? What is the probability that the

time window is that size?

4. How long does it take for two devices to both receive a message from the other for the first time (i.e. two-way hand shake)? What is the probability that it takes that long?

5. How long will it take for each device in the network to communicate with all others? What is the probability that it takes that long?

The answers to questions like these assist network designers in determining whether or not the network is likely to support all requirements, and determine how much risk they are willing to (or must) accept given a set of network resource constraints.

# Supplemental

MATLAB® Code and Simulation Data.

This supplemental material includes all MATLAB® scripts employed to run the NOMA network model simulations described in Chapter 6, and the resulting data (saved as MAT-LAB® workspaces) on which the error and distribution similarity analysis was performed.

The supplemental material is available through the Dudley Knox Library at the Naval Postgraduate School.

Dudley Knox Library
Naval Postgraduate School
411 Dyer Rd. Bldg. 339
Monterey, CA 93943

Phone: (831) 656-2947
Email: circdesk@nps.edu

THIS PAGE INTENTIONALLY LEFT BLANK

# List of References

[1] GSA, "LTE & 5G subscribers - update," Global mobile Suppliers Association (GSA), Farnham, GU9 1LU, UK, 2022 [Online]. Available: https://gsacom.com/paper/subscription-update-january-2022/

[2] United States Census Bureau, "U.S and world population clock," Jan. 14, 2022 [Online]. Available: https://www.census.gov/popclock/

[3] R. L. Ackoff, "From data to wisdom," *Journal of Applied Systems Analysis*, vol. 16, no. 1, pp. 3–9, 1989.

[4] H. Borko, "Information science: what is it?" *American Documentation*, vol. 19, no. 1, pp. 3–5, 1968.

[5] H. A. H. A. Simon, *The Sciences of the Artificial*, 3rd ed. Cambridge, MA, USA: MIT Press, 1996.

[6] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. Burlington, MA, USA: Academic Press, 2011.

[7] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. Cambridge, MA, USA: Academic Press, 2018.

[8] J. Boccuzzi, "Introduction to cellular mobile communications," in *Multiple Access Techniques for 5G Wireless Networks and Beyond*, M. Vaezi, Z. Ding, and H. V. Poor, Eds. Berlin, Germany: Springer, 2019, pp. 3–37.

[9] Pew Research Center, "Mobile fact sheet," Apr. 7, 2021 [Online]. Available: https://www.pewresearch.org/internet/fact-sheet/mobile/

[10] J. Park, "S. Korea first to roll out 5G services, beating U.S. and China," *Reuters*, Apr. 3, 2019 [Online]. Available: https://www.reuters.com/article/southkorea-5g/s-korea-first-to-roll-out-5g-services-beating-u-s-and-china-idUSL3N21K114

[11] ITU-R, "IMT vision–framework and overall objectives of the future development of IMT for 2020 and beyond," *Recommendation ITU-R M.2083*, September, 2015.

[12] S. Ahmadi, *5G NR: Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standards*. Cambridge, MA, USA: Academic Press, 2019.

[13] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.

[14] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[15] M. Vaezi, Z. Ding, and H. V. Poor, Eds., *Multiple Access Techniques for 5G wireless Networks and Beyond*. Berlin, Germany: Springer, 2019.

[16] Y. Liu, S. Zhang, X. Mu, Z. Ding, R. Schober, N. Al-Dhahir, E. Hossain, and X. Shen, "Evolution of NOMA toward next generation multiple access (NGMA) for 6G," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1037–1071, 2022.

[17] B. Pimentel, A. Bordetsky, and R. Gera, "Robustness in nonorthogonal multiple access 5G networks," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, pp. 7444–7453.

[18] 3GPP, "Release 16: study on non-orthogonal multiple sccess (NOMA) for NR," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.812, Dec 2018, version 16.0.0. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3236

[19] TSG-RAN, "Meeting report for TSG RAN meeting: 82," 3rd Generation Partnership Project (3GPP), Meeting Report (RP) 190643, Dec 2018. Available: https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_82/Report/RP-190643.zip

[20] B. Makki, K. Chitti, A. Behravan, and M.-S. Alouini, "A survey of NOMA: Current status and open research challenges," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, 2020.

[21] Y. Yuan, Z. Yuan, and L. Tian, "5G non-orthogonal multiple access study in 3GPP," *IEEE Communications Magazine*, vol. 58, no. 7, pp. 90–96, 2020.

[22] T. Hou, Y. Liu, Z. Song, X. Sun, and Y. Chen, "Exploiting NOMA for UAV communications in large-scale cellular networks," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 6897–6911, 2019.

[23] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[24] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "UAV-enabled communication using NOMA," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5126–5138, 2019.

[25] Z. Na, Y. Liu, J. Shi, C. Liu, and Z. Gao, "UAV-supported clustered NOMA for 6G-enabled internet of things: Trajectory planning and resource allocation," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 041–15 048, 2020.

[26] 3GPP, "Release 15: Study on New Radio (NR) to support non-terrestrial networks," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.811, Sep 2020, version 15.4.0. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3234

[27] 3GPP, "Release 16: Solutions for NR to support non-terrestrial networks (NTN)," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.821, May 2021, version 16.1.0. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3525

[28] E. Bjornson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 100–108, 2019.

[29] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[30] Y. Sun, Z. Ding, X. Dai, and O. A. Dobre, "On the performance of network NOMA in uplink CoMP systems: A stochastic geometry approach," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5084–5098, 2019.

[31] Y. Shi, K. Yang, T. Jiang, J. Zhang, and K. B. Letaief, "Communication-efficient edge AI: Algorithms and systems," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2167–2191, 2020.

[32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005.

[33] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2294–2323, 2018.

[34] M. Vaezi and H. V. Poor, "NOMA: An information-theoretic perspective," in *Multiple Access Techniques for 5G Wireless Networks and Beyond*, M. Vaezi, Z. Ding, and H. V. Poor, Eds., Berlin, Germany.

[35] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," in *11th International Symposium on Wireless Communications Systems*. IEEE, 2014, pp. 781–785.

[36] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1501–1505, 2014.

[37] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, 2015.

[38] M. Shirvanimoghaddam, M. Condoluci, M. Dohler, and S. J. Johnson, "On the fundamental limits of random non-orthogonal multiple access in cellular massive IoT," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2238–2252, 2017.

[39] M. Vaezi, R. Schober, Z. Ding, and H. V. Poor, "Non-orthogonal multiple access: Common myths and critical questions," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 174–180, 2019.

[40] M. Vaezi, G. A. A. Baduge, Y. Liu, A. Arafa, F. Fang, and Z. Ding, "Interplay between NOMA and other emerging technologies: A survey," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 4, pp. 900–919, 2019.

[41] M. Elbayoumi, M. Kamel, W. Hamouda, and A. Youssef, "NOMA-assisted machine-type communications in UDN: State-of-the-art and challenges," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1276–1304, 2020.

[42] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for IoT: A survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1805–1838, 2020.

[43] Y. Qi, X. Zhang, and M. Vaezi, "Over-the-air implementation of NOMA: New experiments and future directions," *IEEE Access*, vol. 9, pp. 135 828–135 844, 2021.

[44] Y. Chen, A. Bayesteh, Y. Wu, B. Ren, S. Kang, S. Sun, Q. Xiong, C. Qian, B. Yu, Z. Ding *et al.*, "Toward the standardization of non-orthogonal multiple access for next generation wireless networks," *IEEE Communications Magazine*, vol. 56, no. 3, pp. 19–27, 2018.

[45] S. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2016.

[46] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—A novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, 2017.

[47] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, 2013, pp. 332–336.

[48] M. Moltafet, N. Mokari, M. R. Javan, H. Saeedi, and H. Pishro-Nik, "A new multiple access technique for 5g: Power domain sparse code multiple access (PSMA)," *IEEE Access*, vol. 6, pp. 747–759, 2018.

[49] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938–947, 2006.

[50] A. Haghighat, S. N. Nazar, S. Herath, and R. Olesen, "On the performance of IDMA-based non-orthogonal multiple access schemes," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[51] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 364–375, 1996.

[52] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, 2013, pp. 770–774.

[53] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[54] K. Higuchi and A. Benjebbour, "Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access," *IEICE Transactions on Communications*, vol. 98, no. 3, pp. 403–414, 2015.

[55] M. Newman, *Networks*. Oxford, UK: Oxford University Press, 2018.

[56] G. Bianconi, *Multilayer Networks: Structure and Function*. Oxford, UK: Oxford University Press, 2018.

[57] G. Chartrand and P. Zhang, *A First Course in Graph Theory*. Mineola, NY, USA: Dover Publications, Inc., 2013.

[58] V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, and V. Latora, "Components in time-varying graphs," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 023101, pp. 1–11, 2012.

[59] P. Holme, "Modern temporal network theory: A colloquium," *The European Physical Journal B*, vol. 88, no. 9, pp. 1–30, 2015.

[60] N. Masuda and R. Lambiotte, *A Guide to Temporal Networks*, 2nd ed. (Series on Complexity Science). Hackensack, NJ, USA: World Scientific, 2020, vol. 6.

[61] Q. Liang and E. Modiano, "Survivability in time-varying networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 9, pp. 2668–2681, 2016.

[62] G. Khanna, S. Chaturvedi, and S. Soh, "Time varying communication networks: Modelling, reliability evaluation and optimization," in *Advances in Reliability Analysis and Its Applications*, M. Ram and H. Pham, Eds. Cham, Switzerland: Springer, 2020, pp. 1–30.

[63] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, "Communicability across evolving networks," *Physical Review E*, vol. 83, no. 4, p. 046120, 2011.

[64] V. Nicosia, J. Tang, C. Mascolo, M. Musolesi, G. Russo, and V. Latora, "Graph metrics for temporal networks," in *Temporal Networks*, P. Holme and J. Saramäki, Eds. New York, USA: Springer, 2013, pp. 15–40.

[65] B. Pimentel, A. Bordetsky, R. Gera, A. Conti, and M. Z. Win, "Temporal connectivity as a robustness measure in NOMA wireless networks," in *IEEE International Conference on Communications*, 2022, pp. 3911–3917.

[66] Huawei and HiSilicon, "LLS results for uplink multiple access," R1-164037, 3GPP TSG RAN WG1 Meeting #85, Nanjing, China, May 23-27 2016.

[67] Y. Huang, J. Wang, and J. Zhu, "Optimal power allocation for downlink NOMA systems," in *Multiple Access Techniques for 5G Wireless Networks and Beyond*, M. Vaezi, Z. Ding, and H. V. Poor, Eds. Berlin, Germany: Springer, 2019, pp. 195–227.

[68] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.

[69] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205–230, 2009.

[70] H. ElSawy, A. Sultan-Salem, M.-S. Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 167–203, 2016.

[71] G. Chisci, H. ElSawy, A. Conti, M. Alouini, and M. Z. Win, "Uncoordinated massive wireless networks: Spatiotemporal models and multiaccess strategies," *IEEE/ACM Transactions on Networking*, vol. 27, no. 3, pp. 918–931, 2019.

[72] M. Emara, H. ElSawy, and G. Bauch, "A spatiotemporal model for peak AoI in uplink IoT networks: Time versus event-triggered traffic," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6762–6777, 2020.

[73] K. S. Ali, M. Haenggi, H. ElSawy, A. Chaaban, and M. Alouini, "Downlink non-orthogonal multiple access (NOMA) in Poisson networks," *IEEE Transactions on Communications*, vol. 67, no. 2, pp. 1613–1628, 2018.

[74] X. Yue, Z. Qin, Y. Liu, S. Kang, and Y. Chen, "A unified framework for non-orthogonal multiple access," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5346–5359, 2018.

[75] T. Hou, Y. Liu, Z. Song, X. Sun, and Y. Chen, "UAV-to-everything (U2X) networks relying on NOMA: A stochastic geometry model," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7558–7568, 2020.

[76] C. Zhang, Y. Liu, and Z. Ding, "Semi-grant-free NOMA: A stochastic geometry model," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 1197–1213, 2022.

[77] Z. Ding, R. Schober, and H. V. Poor, "A new QoS-guarantee strategy for NOMA assisted semi-grant-free transmission," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7489–7503, 2021.

[78] R. R. Brooks, B. Pillai, S. Racunas, and S. Rai, "Mobile network analysis using probabilistic connectivity matrices," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 4, pp. 694–702, 2007.

[79] S. Dasgupta, G. Mao, and B. Anderson, "A new measure of wireless network connectivity," *IEEE Transactions on Mobile Computing*, vol. 14, no. 9, pp. 1765–1779, 2015.

[80] S. Scellato, I. Leontiadis, C. Mascolo, P. Basu, and M. Zafer, "Evaluating temporal robustness of mobile networks," *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 105–117, 2013.

[81] P. K. Sangdeh, H. Pirayesh, Q. Yan, K. Zeng, W. Lou, and H. Zeng, "A practical downlink NOMA scheme for wireless LANs," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2236–2250, 2020.

[82] R. R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple-access system," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 918–930, 1988.

[83] Y. Zhong, M. Haenggi, T. Q. Quek, and W. Zhang, "On the stability of static Poisson networks under random access," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2985–2998, 2016.

[84] C. Therrien and M. Tummala, *Probability and Random Processes for Electrical and Computer Engineers*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2012.

[85] A. Papoulis and U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. Boston, MA, USA: McGraw-Hill, 2002.

[86] B. Bollobás, "A probabilistic proof of an asymptotic formula for the number of labelled regular graphs," *European Journal of Combinatorics*, vol. 1, no. 4, pp. 311–316, 1980.

[87] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, "Activity driven modeling of time varying networks," *Scientific Reports*, vol. 2, no. 1, pp. 1–7, 2012.

[88] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[89] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature Physics*, vol. 8, no. 1, pp. 40–48, 2012.

[90] M. M. Danziger, I. Bonamassa, S. Boccaletti, and S. Havlin, "Dynamic interdependence and competition in multilayer networks," *Nature Physics*, vol. 15, no. 2, pp. 178–185, 2019.

[91] S. Chen, Y. Gao, X. Liu, J. Gao, and S. Havlin, "Robustness of interdependent networks based on bond percolation," *EPL (Europhysics Letters)*, vol. 130, no. 3, p. 38003, 2020.

[92] P. Erdös and A. Rényi, "On random graphs," *Publ Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[93] G. Perarnau and G. Petridis, "Matchings in random biregular bipartite graphs," *The Electronic Journal of Combinatorics*, vol. 20, no. 1, pp. 1–30, 2013.

[94] B. D. McKay and F. Skerman, "Degree sequences of random digraphs and bipartite graphs," *Journal of Combinatorics*, vol. 7, no. 1, pp. 21–49, 2016.

[95] Y. Yuan, S. Wang, Y. Wu, H. V. Poor, Z. Ding, X. You, and L. Hanzo, "NOMA for next-generation massive IoT: Performance potential and technology directions," *IEEE Communications Magazine*, vol. 59, no. 7, pp. 115–121, 2021.

[96] J. Mun, *Quantitative Research Methods*. Dublin, CA, USA: IIPER Press, 2019.

[97] F. J. Caro-Lopera, V. Leiva, and N. Balakrishnan, "Connection between the Hadamard and matrix products with an application to matrix-variate Birnbaum–Saunders distributions," *Journal of Multivariate Analysis*, vol. 104, no. 1, pp. 126–139, 2012.

[98] S. J. Leon, *Linear Algebra with Applications*, 8th ed. Upper Saddle River, NJ, USA: Pearson/Prentice Hall, 2010.

[99] V. I. Bogachev, *Measure Theory*. Berlin, Heidelberg: Springer-Verlag, 2007.

[100] G. Neglia, G. Calbi, D. Towsley, and G. Vardoyan, "The role of network topology for distributed machine learning," in *IEEE Conference on Computer Communications*, 2019, pp. 2350–2358.

[101] T. Park, W. Saad, and B. Zhou, "Centralized and distributed age of information minimization with nonlinear aging functions in the Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 8437–8455, 2020.

[102] A. Nedić, A. Olshevsky, and M. G. Rabbat, "Network topology and communication-computation tradeoffs in decentralized optimization," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 953–976, 2018.

[103] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991.

[104] M. A. Thomas and A. E. Taub, "Calculating binomial probabilities when the trial probabilities are unequal," *Journal of Statistical Computation and Simulation*, vol. 14, no. 2, pp. 125–131, 1982.

[105] *Command and Control*, MCDP 6, U.S. Marine Corps, Washington, DC, USA, 1996.

[106] *Force Design 2030*, Commandant of the Marine Corps (CMC), Washington, DC, USA, 2020 [Online]. Available: https://www.marines.mil/News/News-Display/Article/2711668/force-design-2030

[107] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, "6G Internet of Things: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 359–383, 2022.

[108] S. Russell and T. Abdelzaher, "The Internet of Battlefield Things: The next generation of command, control, communications and intelligence (C3I) decision-making," in *IEEE Military Communications Conference*, 2018, pp. 737–742.

[109] T. Abdelzaher, N. Ayanian, T. Basar, S. Diggavi, J. Diesner, D. Ganesan, R. Govindan, S. Jha, T. Lepoint, B. Marlin *et al.*, "Toward an Internet of Battlefield Things: A resilience perspective," *Computer*, vol. 51, no. 11, pp. 24–36, 2018.

[110] *A Concept for Stand-in Forces*, Commandant of the Marine Corps (CMC), Washington, DC, USA, 2021 [Online]. Available: https://www.hqmc.marines.mil/LinkClick.aspx?fileticket=IjfhTYwXBJk%3d&tabid=10341&portalid=142&mid=21287

[111] CellSite Solutions, "Cell on wheels," Accessed Feb. 3, 2022 [Online]. Available: https://cellsitesolutions.com/portfolio-view/cows/

# Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California