UNIVERSITY OF CALIFORNIA

Los Angeles

On Information Theoretic and Distortion-based Security

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical and Computer Engineering

by

Gaurav Kumar Agarwal

2019

ABSTRACT OF THE DISSERTATION

On Information Theoretic and Distortion-based Security

by

Gaurav Kumar Agarwal
Doctor of Philosophy in Electrical and Computer Engineering
University of California, Los Angeles, 2019
Professor Christina Panagio Fragouli, Chair

In this thesis, we consider secure communication in the presence of an eavesdropper. With the explosion in the growth of the data produced and communicated, sensitive information such as financial transactions, health records, and control signals for cyber-physical systems, has to be *securely* exchanged. Today, the ever-increasing computational power of adversaries is challenging the state-of-the-art cryptographic encryption mechanisms, as these mechanisms assume adversaries with limited computational power. Thus, with the advent of the quantum computing era, we require new mechanisms to guarantee a secure exchange of information. Moreover, the growing number of small and energy constrained connected devices involved in data exchange calls for lightweight encryption schemes, as low-complexity devices cannot implement complex schemes.

We consider three different scenarios and exploit specific opportunities present in each of these scenarios; we develop lightweight encryption schemes that do not require sharing large keys in advance and are secure against eavesdroppers with unlimited computational capabilities.

The first scenario we consider is *multiple unicast traffic* over wireline networks. In these networks, a single source is connected to m destinations interested in different messages. In designing encryption schemes, we exploit the fact that although the eavesdropper is computationally superpowerful, it might not have capabilities to eavesdrop the entire network. We use the multi-path diversity to securely communicate against the eavesdropper without requiring any pre-shared key.

The second scenario we consider consists of millimeter wave (mmWave) networks. mmWave communication requires deploying networks of relays that communicate through directional beams

to compensate for the high path-loss and the high blockage. Since we need to use beamforming and align beams to activate links, we cannot use all the underlying links of the network simultaneously. However, the degree of freedom in choosing the links to activate can be leveraged for secure communication against an eavesdropper. We show that we can achieve a secure capacity that in some cases, can be very close to the unsecure capacity. Here, capacity refers to the maximum flow of information over the network.

For the third scenario, we consider cyber-physical systems and propose a distortion based security framework where the distortion measures the distance between the eavesdropper's estimates and the ground truth. The primary motivation for this framework is that the messages exchanged in these systems are embedded in a metric space having a notion of distance, and securing raw bits as in traditional encryption schemes might not be necessary. Instead, we show with an example of a linear dynamical system that a carefully designed encryption scheme can significantly distort the eavesdropper's view with just one bit of the pre-shared key.

The dissertation of Gaurav Kumar Agarwal is approved.

Richard D Wesel

Paulo Tabuada

Suhas N. Diggavi

Christina Panagio Fragouli, Committee Chair

University of California, Los Angeles

2019

To my mother

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor Prof. Christina Fragouli. She is the perfect Ph.D. advisor any one can hope for. I feel very fortunate to be working with her. She not only helped me academically - in finding the right problems and giving her valuable insights on each of my results, but she was also just an email away for literally any of my concerns. If I have to name one person whom I could count for any help during my Ph.D. life, she is that one person.

I also benefited from the graduate classes taught by Prof. Suhas Diggavi that formed the working horse of my research. Collaborations with him and Prof. Paulo Tabuada resulting in the work on *distortion based security* formed a significant part of my thesis and exposed me to the exciting field of cyber-physical systems. Weekly meetings with Prof. Tabuada's group, where I used to talk about my research were greatly enriching - I learned something new in every such session and made amazing friends. I would also like to thank Prof. Richard Wesel, Prof. Tabuada, and Prof. Diggavi who kindly agreed to be part of my doctoral committee, and provided valuable feedback during the qualification and the final exam of my Ph.D..

ARNI lab has been like a home for me and I would like to thank Linqi, Yahya, Karmoose, Martina, Ayan, Sundar, and Osama for making it so. Prof. Martina Cardone, who was a post-doc in our lab during my first three years of Ph.D., played a very significant role in my Ph.D. life. I would cherish collaborating with her and the resulting lifelong friendship. Her greatness manifests from the fact that she remains ready for any help even when she is most busy with her academic job. Yahya Ezzlendin started helping me even before I landed in Los Angeles, and no matter how troubled I was, just the mere presence of him, made all problems disappear. I would miss playing munchkins with him. Mohammed Karmoose was another go-to solution for every problem I had. Whenever I was in doubt about something, he would be the first person I would consult. Additionally, I enjoyed collaborating with him on *distortion based security*. I was also glad to work with many amazing undergraduate students who worked in our lab. In particular, I would like to thank Chi-Yo Tsai, William Chen, Nikky Woo, Kathy Daniels, and Tara Sadjadpour with whom I collaborated.

Aayush Jain and Hemant Kumar are like the two brothers of mine here in the US who never let me feel away from home. I shared all my ups and downs with them, and they were always with me to give moral support during the downs. Finally, I would also like to thank my parents for everything they did to raise me. I am also thankful to my brother and my sister, who took care of all the responsibilities back at home in my absence so that I could focus on my Ph.D..

VITA

2010	Summer Undergraduate Researcher, Indian Institute of Technology, Roorkee.
2011	Visiting Student Researcher, Cranfield University, Shrivenhem.
2012	B.Tech. (Electronics and Communication Engineering), Indian Institute of Technology, Roorkee.
2012	Software Engineer, Paypal India Private Limited, Chennai.
2013	Design Engineer, Freescale Semiconductor India Private Limited, Noida.
2015	M.Eng. (Telecommunications), Indian Institute of Science, Bangalore.
2016	Summer Intern, Technicolor Research, Los Altos.
2016–2019	Graduate Student Researcher (GSR), Electrical and Computer Engineering, UCLA.
2016, 2017	Teaching Assistant (TA), Signals and Systems (ECE 102) Electrical and Computer Engineering, UCLA.
2018	Teaching Assistant (TA), Graph Theory (ECE 134) Electrical and Computer Engineering, UCLA.
2019	Teaching Fellow, Statistical Machine Learning (ECE 239AS) Electrical and Computer Engineering, UCLA.

PUBLICATIONS

G. K. Agarwal, M. Cardone, C. Fragouli, "On Secure Network Coding for Multiple Unicast Traffic ," **submitted for** IEEE Trans. Information Theory, 2018.

G. K. Agarwal, M. Karmoose, S. Diggavi, C. Fragouli, P. Tabuada, "Distortion based Light-weight Security for Cyber-Physical Systems," **submitted for** IEEE Trans. Automatic Control, 2018.

G. K. Agarwal, M. Cardone, C. Fragouli, "On Secure Capacity of Multiple Unicast Traffic over Separable Networks ," **submitted for** IEEE International Workshop on Information Theory, 2019.

G. K. Agarwal, M. Karmoose, S. Diggavi, C. Fragouli, P. Tabuada, "Distorting an Adversary's View in Cyber-Physical Systems," in IEEE Conference on Decision and Control, 2018.

G. K. Agarwal, Y. Ezzeldin, M. Cardone, C. Fragouli, "Secure Communication over 1-2-1 Networks," in IEEE International Symposium on Information Theory, 2018.

G. K. Agarwal, M. Cardone, C. Fragouli, "Secure Network Coding for Multiple Unicast: On the Case of Single Source," in International Conference on Information Theoretic Security, 2017.

C.-Y. Tsai, **G. K. Agarwal**, C. Fragouli, S. Diggavi, "A distortion based approach for protecting inferences," in IEEE International Symposium on Information Theory, 2017.

G. K. Agarwal, M. Cardone, C. Fragouli, "On secure network coding for two unicast sessions: studying butterflies," in IEEE Globecom Workshops, 2016.

G. K. Agarwal, M. Cardone, C. Fragouli, "Coding across unicast sessions can increase the secure message capacity," in IEEE International Symposium on Information Theory, 2016.

CHAPTER 1

Introduction

The enormous growth in the data we are communicating over the Internet is such that in the last two years we have exchanged 90% of the data communicated in the entire history [DOM19]. Today, we are exchanging thousands of gigabytes of data per second over communication networks [DOM19]. Moreover, it is not just the communicated data that is growing; the number of connected devices involved in this exchange of information is also rapidly expending [STA19], and we are expected to have around 75 billion connected devices in the next five years.

A large portion of this enormous data exchanged among the billions of small and low-complexity devices is sensitive in nature, such as banking, health, personal, and proprietary information. Therefore, we need to securely exchange this sensitive information against eavesdropping adversaries who have interests in gaining access to this information.

With an increase in the processing power and a move towards an era of quantum computing [Aro19], these adversaries are gaining more and more power. Their empowerment calls for new mechanisms to guarantee the secure exchange of information. The primary reason for this need is that state-of-the-art encryption methods such as RSA public-key cryptosystems rely on the assumption that the adversary has limited computational capabilities. In other words, the encrypted symbols exchanged over the networks in these schemes contain complete information about the messages, but the adversary is assumed to be computationally incapable of extracting the messages from the encrypted text. If the adversary is instead endowed with strong computational power or a quantum computer, however, messages can be decrypted, and RSA cryptosystems are no longer secure. Moreover, due to the low-complexity of the devices such as ones used in the Internet of things (IoT), it is not possible to implement the hash functions used in RSA cryptosystems. Because of these challenges, low-complexity encryption methods that are secure even against

quantum computers, are going to be necessary and are gaining significant research attention.

One such cipher that is secure against quantum computers and that low-complexity devices can implement, known as *one-time-pad*, is illustrated in Fig. 1.1a. This cipher was first introduced in the 19th century and was proved to be secure [Sha49] against an adversary having *infinite* computational power (quantum computers). Fig 1.1a shows a source S exchanging a message W with the destination D using a *key* K. The key K is a random symbol that is agreed upon before the communication, i.e., it is *pre-shared*. However, every time we use such a cipher, we need a new pre-shared key of length equal to the length of the message W. Pre-sharing such large keys and refreshing them becomes impractical with the scale of data and the number of parties involved in today's communication networks.

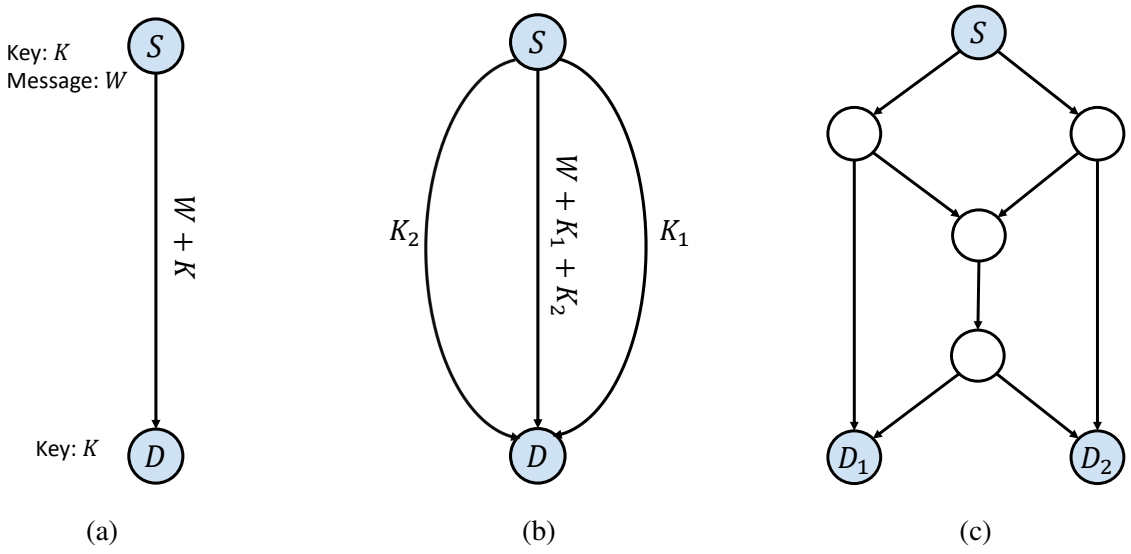


Figure 1.1: (a) One-time-pad. (b) Exploiting limited network presence of the adversary. (c) A network with single source and two destinations.

The primary goal of this thesis is to develop encryption schemes that (1) do not make any assumption on the computational power of the adversary, and (2) are of low-complexity, requiring either no or a small pre-shared key. In this thesis, we focus on three different scenarios, namely (a) multiple unicast traffic over wireline networks, (b) millimeter wave networks, and (c) distortion based security for cyber-physical systems. In each of these scenarios, we exploit specific characteristics of these systems to develop such encryption schemes.

1.1 Multiple unicast traffic over wireline networks

The first scenario we consider is secure communication for *multiple unicast traffic* over a *wireline network model*. The wireline network model abstracts the network connectivity using a graph with edges representing noiseless and non-interfering links. In this scenario, we build on the fact that, although the adversary is assumed to be computationally super-powerful, it may not have the capability to be present everywhere on a network. We exploit this limited presence of the adversary to design secure communication schemes that do not require a pre-shared key.

Cai et al. [CY02] first leveraged the limited network presence of the adversary. We illustrate this in Figure 1.1b with a toy example. In this example, the source S is connected to the destination D with three parallel edges, so at each use of the network, the source S can send three messages to the destination D. However, using two random symbols K_1 and K_2, the source S can instead securely communicate a message W against an adversary eavesdropping *any* two edges of the network. In other words, the destination D can correctly decode the message W based on the information received on three paths whereas the adversary cannot. For the adversary, information eavesdropped on any two edges will be completely useless, i.e., *information theoretically* will have zero *mutual information* [CT06] with the message. This technique can be generalized to any network topology (not just the network with parallel paths in Fig. 1.1b) for traffic consisting of only one source and one destination on the network (also called *unicast* traffic in the literature). Formally, for an arbitrary network with unicast traffic, if the source S can send M messages to the destination D in the absence of any eavesdropper, then, using the scheme of [CY02], the source can instead send $M-k$ messages securely to the destination D against an adversary eavesdropping any k edges. Note that the source and the destination do not know which k edges are eavesdropped.

Traffic consisting of a single source sending **same messages** to a set of destinations on a network is termed *multicast* in the literature (see Fig. 1.1c for an example with both destinations interested in a message W from the source S). The seminal paper by Ahlsewede et. al., [ACL00] showed that for the multicast traffic, if the source can communicate M messages to each destination by exclusively using the entire network resources, then, using coding operations (called *network coding*) at intermediate nodes of the networks, the source can communicate M messages to all the

destinations simultaneously. The result in [CY02] also applies to the multicast traffic, i.e., if the source can communicate M messages to all destinations, it can also simultaneously communicate $M - k$ messages securely to all the destinations against an adversary eavesdropping any k edges.

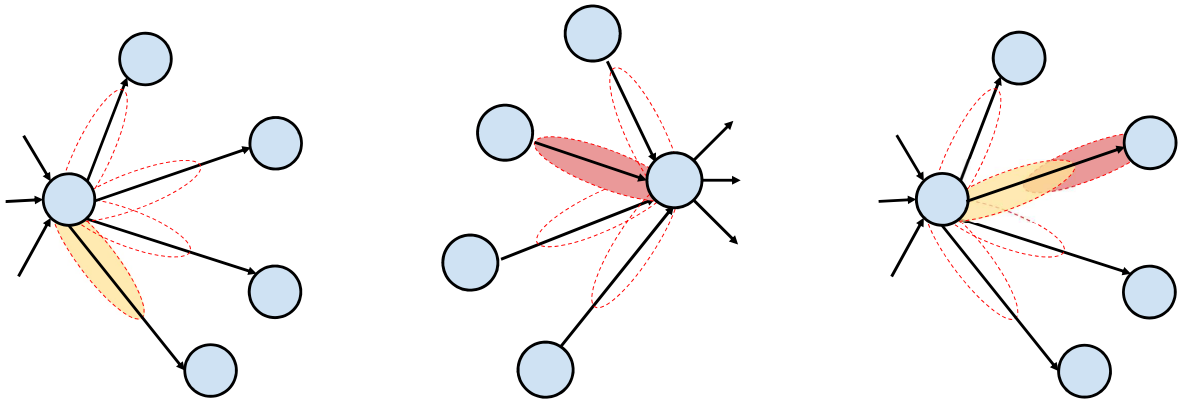
The traffic we consider in this thesis consists of a single source interested in communicating **different messages** to different destinations. In the literature, this traffic is termed as *multiple unicast* (for example, from the source S in Fig. 1.1c, destinations D_1 and D_2 are interested in a message W_1 and a message W_2, respectively). The maximum number of messages the source can communicate to different destinations can be found using the multi-commodity routing algorithm [CLR09, KM03] over the network. We make advances in answering the question of how many messages the source can **securely communicate** to different destinations in the presence of an adversary eavesdropping any k edges of the network. It is worthwhile to note here that a more general traffic consists of multiple source-destination pairs; for each source-destination pair, the destination is interested in a message from the corresponding source. However, characterizing the maximum flow of information over such networks remains an open problem [KTW14] even when there is no eavesdropping adversary. Thus, we restrict ourselves to the case of single source when there is an eavesdropping adversary.

The single source assumption also allows us to exploit the fact that, even if the messages are different for every destination, we can use common random symbols (such as K_1 and K_2 in Fig. 1.1b) for securing the messages against an adversary eavesdropping any k edges. Moreover, we show that in a network, not all the edges are equally suitable for sending common random packets (multicast traffic), and in order to get the optimal performance, (that is to communicate the maximum number of private messages securely), we need to exploit specific sub-networks to send common random packets. We find these sub-networks by first defining the notion of *separable networks* and then using this notion to identify the sub-networks suited for multicasting.

1.2 Millimeter wave networks (1-2-1 networks)

In the second scenario, we consider the problem of secure communication in *millimeter wave networks*. Millimeter wave communication uses a much broader available spectrum (bandwidth)

in the high-frequency region. In particular, it occupies the frequency spectrum from 30 GHz to 300 GHz. The availability of this large spectrum is poised to enable streaming of ultra-high definition videos, and communication among a large number of autonomous vehicle platoons in 5G mobile communication [All15].



(a) Transmitter beamforming in 1-2-1 networks.

(b) Receiver beamforming in 1-2-1 networks.

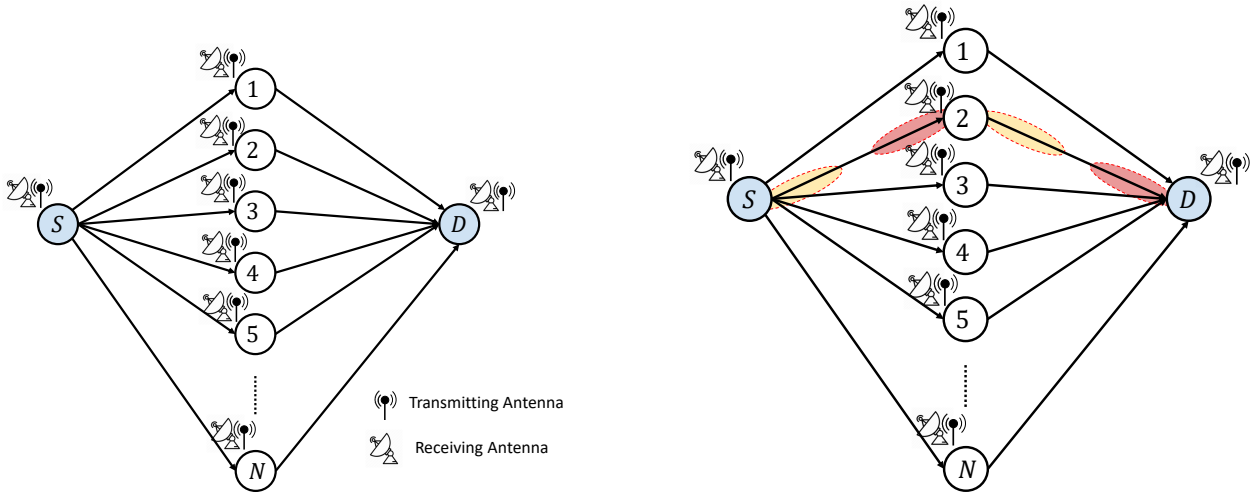
(c) Matching of transmitting and receiving beams.

Figure 1.2: Transmitting and receiving over 1-2-1 networks

However, transmissions at such high frequency suffer from high path-loss and high blockage. Thus, a transmitting node needs to use an array of antennas to beamform in a narrow direction, as shown in Fig. 1.2a. Similarly, a receiving node needs to steer its receiving antenna in a particular direction, as shown in Fig. 1.2b. The 1-2-1 (one-to-one) model [ECF18] abstracts this directivity: to establish a communication link, both the millimeter wave transmitter and receiver employ antenna arrays that electronically steer to direct their beams towards each other, as shown in Fig. 1.2c.

With this constraint of alignment of the beams to establish a communication link, we cannot use all network resources at a time. For example, in Fig. 1.3a, the source S is connected to the destination D via N relays/paths. However, if every node has only one transmitting and one receiving antenna to beamform, out of N paths from the source to the destination, we can only use one path at a time, as shown in Fig. 1.3b.

The freedom to select this path out of N choices provides an opportunity for security against an adversary who eavesdrops a fixed set of edges. For instance, in the example in Fig. 1.3a, the maximum number of messages the source S can send to the destination D is just one, as we can



(a) (b)

Figure 1.3: (a) An example of $1-2-1$ network model. (b) Transmission on one chosen path.

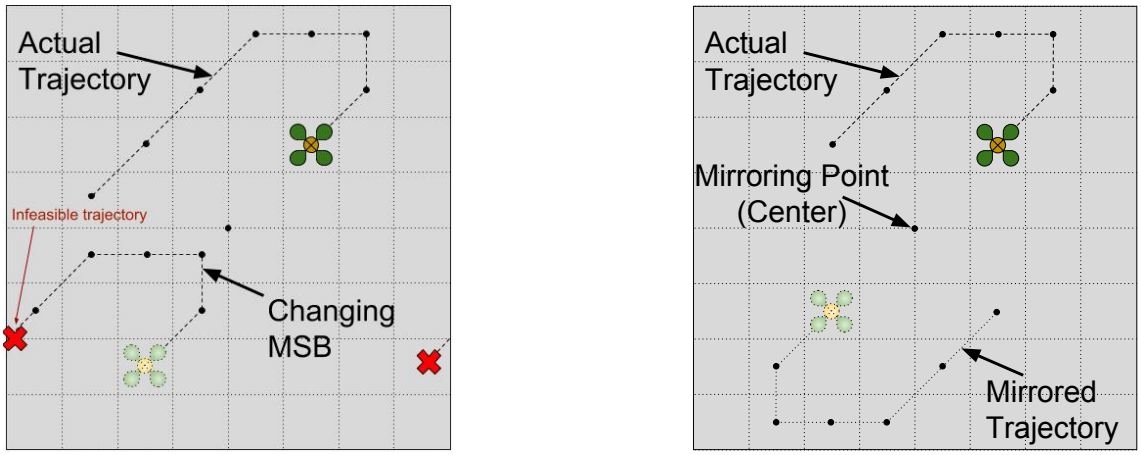
only use one path at a time. However, by employing a time-varying selection for this path, the source can securely transmit at a rate of $(1-\frac{k}{N})$ against an adversary eavesdropping any k edges. In particular, in each use of this network, the source selects a different path and communicates a different symbol. With this, in N uses, the source delivers N different symbols. The adversary eavesdropping on k edges will be able to overhear transmissions of at most k of them. Using a standard coding scheme like Reed-Solomon code, the source can hence securely communicate $N-k$ messages in N network uses. Thus, only a fraction k/N of the packets is lost when securing against the adversary. In contrast, in the wireline network model, where a node can transmit and receive on all outgoing and incoming edges respectively, a rate of k is lost in providing security. In this thesis, we will generalize this for networks having arbitrary topology and characterize the secure capacity for single source and single destination networks by using a time-varying selection of the sub-networks without requiring any pre-shared key.

1.3 Distortion based security for cyber-physical systems

For the third scenario, we consider communication in cyber-physical systems (CPS) where communication, computations, and control are intertwined. The messages exchanged over these net-

works are control signals or the states of the systems. These messages are embedded in a *metric space*, i.e., a space with a notion of distance (for example, the speed of a moving car or the location of a flying drone). We will exploit this fact to design lightweight encryption schemes. In particular, we will design schemes in which, instead of having an adversary who does not learn *anything* about the message, we force the adversary's estimations to be "quite far" from the original message.

For example, consider a drone flying, as depicted in Fig. 1.4a. At every time instance, the drone moves between adjacent squares inside the grid and wants to communicate its location to a control server. An eavesdropper, also interested in the location of the drone, overhears this communication. Since at any time the drone is in one of the 64 positions, it can use an encryption based on the one-time-pad (see Fig. 1.1a) and communicate its position securely by using a 6 bits long secret key per time instance. However, if we only require the adversary's estimate to be sufficiently far from the actual position, we do not require a key of this size. Towards this end, we define the notion of distortion-based security that maximizes the difference between the adversary's estimate and the ground truth (i.e., the actual position).



(a) Protection of the most significant bit.

(b) Mirroring based scheme.

Figure 1.4: Example of drone motion

However, the challenge is in designing encryption schemes that follow the underlying system dynamics. For instance, flipping the most significant bit in the representation of the location, as shown in Fig. 1.4a leads to a trajectory which is inconsistent with the system dynamics. Thus, with this encryption the adversary is able to separate out the fake trajectory form the actual trajectory,

and thus she learns the actual trajectory.

For this particular example, it turns out that if the adversary has to make a random guess that minimizes the distance from the actual location, then the adversary estimates the location of the drone at the origin. Thus, a good encryption scheme strives to keep the adversary's estimate close to the origin even when she has some information. We use the following encryption scheme, where the drone either sends its actual location or a "mirrored" version of it, as shown in Fig. 1.4b. This scheme keeps the adversary's estimate precisely at the origin.

In this thesis, we generalize this mirroring scheme and show that in the distortion-based security framework, a single bit of pre-shared key is sufficient for any linear dynamical system to communicate its state securely (such as the one shown in Fig. 1.4a).

1.4 Main contributions

For the three scenarios considered in this thesis, our major contributions are as follows. The work on distortion based security for cyber-physical systems is joint work with another Ph.D. student Mohammed Karmoose.

Multiple Unicast Traffic over Wireline Networks:

We characterize the secure capacity region for networks with two destinations having arbitrary topology. Using the notion of "separable networks" and their reduction to a two-layer network topology, we characterize the secure capacity region for additional classes of networks having more than two destinations. Finally, we provide a polynomial-time heuristic for securely communicating over networks with arbitrary topology and an arbitrary number of destinations.

Millimeter Wave Networks (1-2-1 Networks):

We consider arbitrary 1-2-1 networks with unit capacity edges and derive lower and upper bounds on the secure capacity. We also characterize the secure capacity for a particular network topology called diamond networks where the edges can have arbitrary edge capacities.

Distortion Based Security for Cyber-Physical Systems:

We identify security measures based on assessing the distance of the adversary's estimates from the ground truth. In particular, we provide both average-case and worst-case performance guarantees. For the average-case distortion, we develop a scheme which uses exactly one bit of the pre-shared key and can provide maximum possible distortion (equivalent to the eavesdropper with no observations) in some cases. For the worst-case distortion, we design a scheme that uses 3 bits of the pre-shared key per dimension and prove that it achieves the maximum possible distortion when the inputs to the systems are independent of the previous states.

1.5 Organization

This thesis is organized as follows. The standard notation used throughout the thesis is given in Section 1.6. Section 1.7 gives an overview of the related literature in information theoretic secrecy over networks, and security over cyber-physical systems. Chapter 2, Chapter 3 and Chapter 4 analyze the three secure communication scenarios of interest, i.e., multiple unicast traffic, millimeter wave networks and distortion based security, respectively. Finally, future directions and open questions are discussed in Chapter 5.

1.6 Notation

Throughout this thesis, we adopt the following notation convention. Calligraphic letters indicate sets; $\emptyset$ is the empty set and $|\mathcal{A}|$ is the cardinality of $\mathcal{A}$; for two sets $\mathcal{A}_1, \mathcal{A}_2$, $\mathcal{A}_1 \subseteq \mathcal{A}_2$ indicates that $\mathcal{A}_1$ is a subset of $\mathcal{A}_2$, $\mathcal{A}_1 \cup \mathcal{A}_2$ indicates the union of $\mathcal{A}_1$ and $\mathcal{A}_2$, $\mathcal{A}_1 \sqcup \mathcal{A}_2$ indicates the disjoint union of $\mathcal{A}_1$ and $\mathcal{A}_2$, $\mathcal{A}_1 \cap \mathcal{A}_2$ is the intersection of $\mathcal{A}_1$ and $\mathcal{A}_2$ and $\mathcal{A}_1 \backslash \mathcal{A}_2$ is the set of elements that belong to $\mathcal{A}_1$ but not to $\mathcal{A}_2$; $[n_1 : n_2]$ is the set of integers from n_1 to $n_2 \geq n_1$; $[n]$ is the set of integers from 1 to $n \geq 1$; $[x]^+ := \max\{0, x\}$ for $x \in \mathbb{R}$; for a vector a, a^T is its transpose vector; $\dim(A)$ is the dimension of the subspace A; $0_{i\times j}$ is the all-zero matrix of dimension $i \times j$; I_j is the identity matrix of dimension j; for a matrix A of dimension $m \times n$, A^T is the transpose of A, A^r is the r-th power of A, $rk(A)$ is the rank of A, and $A|_{\mathcal{S}}$ denotes the submatrix of A of dimension

$|\mathcal{S}| \times n$ where only the rows indexed by the set $\mathcal{S} \subseteq [m]$ are retained. X and X_a denote column vectors, and $X_a^b = [X_a' X_{a+1}' \cdots X_b']'$ for $b \geq a$ and $a, b \in \mathbb{Z}$; $f_X(x)$ denotes the probability density function of a random vector X; for any random vector Y, we denote the mean and covariance matrices of Y by μ_Y and R_Y respectively, (for example, the mean and the covariance matrix of X_a^b are denoted by $\mu_{X_a^b}$ and $R_{X_a^b}$ respectively).

In several parts of this thesis, we represent a network with a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of directed edges. The edges represent orthogonal communication links, which are interference-free. If an edge $e \in \mathcal{E}$ connects a node i to a node j, we refer to node i as the tail and to node j as the head of e, i.e., tail$(e) = i$ and head$(e) = j$. For each node $v \in \mathcal{V}$, we define $\mathcal{I}(v)$ as the set of all incoming edges of node v and $\mathcal{O}(v)$ as the set of all outgoing edges of node v.

1.7 Related work

We first give a brief overview of the information theoretic secure communication for a point to point channel model. Following this, we mention related work in each of the three scenarios considered in this thesis.

Information theoretic security, pioneered by Shannon [Sha49], aims at ensuring a reliable and secure communication among trusted parties inside a network such that a passive external eavesdropper does not learn anything about the content of the information exchanged. For point-to-point channels, information theoretic security can be achieved provided that the communicating trusted parties have a pre-shared key of entropy at least equal to the length of the message [Sha49]. Wyner [Wyn75] showed that, if the adversary's channel is a degraded version of the channel to the legitimate destination, then an information theoretic secure communication can be guaranteed even without the pre-shared keys. Moreover, if public feedback is available, Maurer [Mau93] showed that secure communication can be ensured over erasure networks even when the adversary has a channel of better quality than the legitimate receiver.

Multiple unicast traffic over wireline networks: In [CY02], Cai et al. characterized the information theoretic secure capacity of a noiseless network with unit capacity edges and with

multicast traffic. This seminal works, which was followed by several others [FMS04, ES07], considers a case where a source wishes to multicast the same information to several destinations in the presence of a passive external adversary eavesdropping any k edges of her choice. In [CHK13a], Cui et al. studied networks with non-uniform edge capacities when the adversary is allowed to eavesdrop only some specific subsets of edges. Over the past few years, other notions of information theoretic security have been analyzed, such as the case of weak information theoretic security [BN05, SK09, WYG10]. Moreover, several different scenarios have been studied, that include: (i) the case of an active adversary, who can indeed maliciously corrupt the communication rather than just passively eavesdropping it [JLK07, HLK04, KTT09]; (ii) erasure networks where a public feedback is available [PCF15, ACF16, CPD14, PCF14]; (iii) wireless networks [MSC08, DCN09].

Millimeter wave networks (1-2-1 networks): For our work on millimeter wave networks, we leverage directivity and multipath for security. The fact that directivity can help with security has been observed in the context of MIMO beamforming [MS11, SCA16]; in these works, the main observation is that, by creating a narrow beam, we can limit the locations where the adversary can collect useful information - or at least, significantly weaken her channel, so as to utilize wiretap coding. Exploiting multipath for security over lossless networks with unit capacity links has notably been used in secure network coding [CY02]. The results, however, are only for "traditional networks", where a node can communicate to other nodes using all the edges it is connected with. This is significantly different for millimeter wave networks where a node can only transmit to one among its neighbors at each point in time.

Distortion based security for cyber-physical systems: The study of distortion based security, where the goal is to maximize the distortion of an eavesdropper's estimate on a message, was started by Yamamoto [Yam88]. Schieler and Cuff [SC14] later showed that, in the limit of an infinite block length (n) code, only $\log(n)$ bits of the pre-shared key are needed to achieve the maximum possible distortion. However, Schieler and Cuff also showed that such secrecy is rather fragile: a causal disclosure of even a single message symbol can compromise the secrecy of the entire block. This issue raises because the coding scheme involves infinite block length. We here design schemes with block length equal to one, which obviates the need to wait and accumulate data at the sensor. It also removes the fragility of the distortion based measure as now we do

not need to code a sequence of symbols jointly, and can rather code each symbol independently. Unlike information theoretic security, distortion security of alphabet $\mathcal{A}$, does not imply distortion security of alphabet $\mathcal{B}$ which is a one-to-one mapping of $\mathcal{A}$. Wiese et al. studied a different notion of secure estimation in [WJO16] where they considered zero-error secret capacity. Secure communication in control systems is studied in [TGP17, TGP16, TSS17, MMS13, CDH16]. Securing the system state from an adversary was explored in [TGP17, TGP16], where an asymptotic steady-state analysis was investigated. Information-theoretic security was explored in [TSS17], where the mutual information was used as a privacy measure. Security of the terminal state is considered in [MMS13] where an adversary makes partial noisy measurements of the state trajectory. Differential privacy for control systems was explored in [CDH16], which uses standard statistical indistinguishability which is equally applicable to categorical (non-metric space) data. In our work, we use the estimation error of the adversary in order to quantify privacy, utilizing the fact that CPS data lies in a metric space having a notion of distance, as argued earlier.

CHAPTER 2

Multiple Unicast Traffic over Wireline Networks

Even for different messages requested by various destinations, the source can use the ***same*** *random symbols for all the destinations to securely communicate against an eavesdropper. Moreover, to get the optimal performance, these random symbols have to be multicasted through a suitable selection of sub-network that is identified using the notion of* ***separable networks****.*

2.1 Summary

This chapter investigates the problem of secure communication in a wireline noiseless network model where a source wishes to communicate to a number of destinations in the presence of a passive external adversary. Different from the multicast scenario, where all destinations are interested in receiving the same message, in this setting different destinations are interested in different messages. The main focus of this chapter is on characterizing the secure capacity region.

Towards this end, an outer bound on the secure capacity region is derived, and secure transmission schemes are designed and analyzed in terms of achieved rate performance. It is first shown that, for the case of two destinations, the designed scheme matches the outer bound, hence characterizing the secure capacity region.

To study networks consisting of more than two destinations, a particular class referred to as *two-layer* networks is considered, where the source communicates with the destinations by hopping information through one layer of relays. It is shown that the designed scheme achieves the capacity for any two-layer network for which any of the following three conditions is satisfied: (i) the number of destinations is three, (ii) the number of edges eavesdropped by the adversary is one, (iii) the min-cut capacities satisfy a specific constraint.

We show that the class of two-layer networks is sufficient to model a more general class called *separable* networks. The main feature of separable networks is that they can be partitioned into edge-disjoint networks that satisfy specific min-cut properties. In particular, we prove that the secure capacity region of any separable network can be characterized from the secure capacity region of the corresponding two-layer network.

Finally, for an arbitrary network topology, a polynomial-time two-phase scheme is designed and its performance is compared with the outer bound.

Organization: Section 2.2 formally defines the setup of the multiple unicast wireline noiseless network with a single source and arbitrary number of destinations, and formulates the problem. Section 2.3 derives an outer bound on the secure capacity region. Section 2.4 provides a capacity-achieving secure transmission scheme for networks with two destinations and arbitrary topology. Section 2.5 designs a secure transmission scheme for networks with a two-layer topology and arbitrary number of destinations. Section 2.5 also derives some secure capacity results and shows connections between two-layer networks and separable networks. Section 2.6 provides a two-phase achievable scheme for networks with arbitrary number of destinations and arbitrary topology.

2.2 Setup and problem formulation

The networks considered here have a source node S and m destination nodes $D_i, i \in [m]$. The source node does not have any incoming edges, i.e., $\mathcal{I}(S) = \emptyset$, and each destination node does not have any outgoing edges, i.e., $\mathcal{O}(D_i) = \emptyset, \forall i \in [m]$. Source S has a message W_i for destination $D_i, i \in [m]$. These m messages are assumed to be independent. Thus, the network consists of multiple unicast traffic, where m unicast sessions take place simultaneously and share the network resources. A passive eavesdropper/adversary Eve is also present in the network and can eavesdrop any k edges of her choice. Note that this assumption implies that Eve has limited network presence; this is equivalent to a scenario where there are several *non-collaborating* adversaries, each observing a different subset of k edges. We also highlight that Eve is an external eavesdropper, i.e., she is not one of the destinations.

The symbol transmitted over n channel uses on edge $e \in \mathcal{E}$ is denoted as X_e^n. In addition, for $\mathcal{E}_t \subseteq \mathcal{E}$ we define $X_{\mathcal{E}_t}^n = \{X_e^n : e \in \mathcal{E}_t\}$. We assume that the source node S has infinite sources of randomness Θ, while the other nodes in the network do not have any randomness.

Over this network, we are interested in finding all possible feasible m-tuples $(R_1, R_2, \ldots, R_m)$ such that each destination $D_i, i \in [m]$, reliably decodes the message W_i (with zero error) and Eve receives no information about the content of the messages. In particular, we are interested in ensuring perfect information theoretic secure communication, and hence we aim at characterizing the secure capacity region, which is next formally defined.

Definition 1 (Secure Capacity Region)**.** *A rate m-tuple $(R_1, R_2, \ldots, R_m)$ is said to be securely achievable if there exist a block length n with $R_i = \frac{H(W_i)}{n}$, $\forall i \in [m]$ and a set of encoding functions f_e, $\forall e \in \mathcal{E}$, over a sufficiently large finite field $\mathbb{F}_q$ with*

$$X_e^n = \begin{cases} f_e\left(W_{[m]}, \theta\right) & \text{if tail}(e) = S, \\ f_e\left(\{X_\ell^n : \ell \in \mathcal{I}(tail(e))\}\right) & \text{otherwise}, \end{cases}$$

such that each destination D_i can reliably decode the message W_i i.e.,

$$H\left(W_i | \{X_e^n : e \in \mathcal{I}(D_i)\}\right) = 0, \ \forall i \in [m].$$

Moreover, we also require perfect secrecy, i.e.,

$$I\left(W_{[m]}; X_{\mathcal{E}_\mathcal{Z}}^n\right) = 0, \ \forall\, \mathcal{E}_\mathcal{Z} \subseteq \mathcal{E} \text{ such that } |\mathcal{E}_\mathcal{Z}| \leq k.$$

*The **secure capacity region** is the closure of all such feasible rate m-tuples.*

Definition 2 (Min-cut)**.** *A **cut** is an edge set $\mathcal{E}_\mathcal{A} \subseteq \mathcal{E}$, which separates the source S from a set of destinations $D_\mathcal{A} := \{D_i, \ i \in \mathcal{A}\}$. In a network with unit capacity edges, the minimum cut or **min-cut** is a cut that has the minimum number of edges. Throughout the paper, we denote by $M_\mathcal{A}$ the capacity of the min-cut between the source S and the set of destinations $D_\mathcal{A} := \{D_i, \ i \in \mathcal{A}\}, \mathcal{A} \subseteq [m]$, and we refer to $M_\mathcal{A}$ as the min-cut capacity.*

In Definition 1, we require perfect secrecy, i.e., no matter which (at most) k edges Eve eavesdrops, she does not learn anything about the content of the messages. In particular, throughout the paper, we will use the following condition on perfect secrecy proved in [CY07, Lemma 3.1].

Lemma 1. *Let W be the message vector that has to be transmitted, and K be a vector of uniform i.i.d. symbols independent of W. Then, the vector X representing the symbols transmitted over the edges of the network can be represented in matrix form as*

$$X = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} W \\ K \end{bmatrix},$$

where A and B are the encoding matrices. This transmission scheme is perfectly secure if and only if

$$rk\left(\begin{bmatrix} A & B \end{bmatrix}\Big|_{\mathcal{Z}} \right) = rk\left(B|_{\mathcal{Z}} \right),\ \forall |\mathcal{Z}| \leq k.$$

2.3 Outer bound

In this section, we derive an outer bound on the secure capacity region of a multiple unicast wireline noiseless network with a single source and m destinations. In particular, as stated in Theorem 2, this region depends on the min-cut capacities between the source and different subsets of destinations, and on the number of edges that the adversary eavesdrops. The next theorem provides the outer bound region.

Theorem 2. *An outer bound on the secure capacity region for the multiple unicast traffic over networks with a single source and m destinations is given by*

$$R_{\mathcal{A}} \leq [M_{\mathcal{A}} - k]^+,\ \ \forall \mathcal{A} \subseteq [m]\ \ , \tag{2.1}$$

where $R_{\mathcal{A}} := \sum_{i \in \mathcal{A}} R_i$, and where $M_{\mathcal{A}}$ is defined in Definition 2.

Proof. Let $\mathcal{E}_{\mathcal{A}}$ be a min-cut between the source S and $D_{\mathcal{A}}$ and $\mathcal{E}_{\mathcal{Z}} \subseteq \mathcal{E}_{\mathcal{A}}$ be the set of k edges eavesdropped by Eve, and define $\mathcal{I}(D_{\mathcal{A}}) := \bigcup_{i \in \mathcal{A}} \mathcal{I}(D_i)$. If $|\mathcal{E}_{\mathcal{A}}| < k$, let $\mathcal{E}_{\mathcal{Z}} = \mathcal{E}_{\mathcal{A}}$. We have

$$\begin{aligned} nR_{\mathcal{A}} = H(W_{\mathcal{A}}) &\overset{\text{(a)}}{=} H(W_{\mathcal{A}}) - H(W_{\mathcal{A}}|X^n_{\mathcal{I}(D_{\mathcal{A}})}) \\ &\overset{\text{(b)}}{\leq} H(W_{\mathcal{A}}) - H(W_{\mathcal{A}}|X^n_{\mathcal{E}_{\mathcal{A}}}) \\ &\overset{\text{(c)}}{=} I(W_{\mathcal{A}}; X^n_{\mathcal{E}_{\mathcal{Z}}}, X^n_{\mathcal{E}_{\mathcal{A}} \setminus \mathcal{E}_{\mathcal{Z}}}) \end{aligned}$$

$$
\begin{aligned}
&= I(W_{\mathcal{A}}; X^n_{\mathcal{E}_{\mathcal{Z}}}) + I(W_{\mathcal{A}}; X^n_{\mathcal{E}_{\mathcal{A}}\setminus\mathcal{E}_{\mathcal{Z}}} | X^n_{\mathcal{E}_{\mathcal{Z}}}) \\
&\overset{\text{(d)}}{=} I(W_{\mathcal{A}}; X^n_{\mathcal{E}_{\mathcal{A}}\setminus\mathcal{E}_{\mathcal{Z}}} | X^n_{\mathcal{E}_{\mathcal{Z}}}) \\
&\overset{\text{(e)}}{\leq} H(X^n_{\mathcal{E}_{\mathcal{A}}\setminus\mathcal{E}_{\mathcal{Z}}}) \\
&\overset{\text{(f)}}{\leq} n[M_{\mathcal{A}} - k]^+ \ ,
\end{aligned}
$$

where $W_{\mathcal{A}} = \{W_i, i \in \mathcal{A}\}$ and: (i) the equality in (a) follows because of the decodability constraint (see Definition 1); (ii) the inequality in (b) follows because of the 'conditioning reduces the entropy' principle and since $X^n_{\mathcal{I}(D_{\mathcal{A}})}$ is a deterministic function of $X^n_{\mathcal{E}_{\mathcal{A}}}$; (iii) the equality in (c) follows from the definition of mutual information and since $\mathcal{E}_{\mathcal{A}} = \mathcal{E}_{\mathcal{Z}} \cup \mathcal{E}_{\mathcal{A}\setminus\mathcal{Z}}$; (iv) the equality in (d) follows because of the perfect secrecy requirement (see Definition 1); (v) the inequality in (e) follows since the entropy of a discrete random variable is a non-negative quantity and because of the 'conditioning reduces the entropy' principle; (vi) finally, the inequality in (f) follows since each link is of unit capacity and since $|\mathcal{E}_{\mathcal{A}} \setminus \mathcal{E}_{\mathcal{Z}}| = [M_{\mathcal{A}} - k]^+$. By dividing both sides of the above inequality by n we obtain that $R_{\mathcal{A}}$ in (2.1) is an outer bound on the secure capacity region of the multiple unicast traffic over networks with single source and m destinations. This concludes the proof of Theorem 2. □

2.4 Capacity achieving scheme for networks with two destinations

In this section, we prove that the outer bound in Theorem 2 is tight for the case of $m = 2$ destinations and arbitrary k. Towards this end, we design a secure transmission scheme whose achievable rate region matches the outer bound in Theorem 2. Our scheme follows the works of [CY02] and [Sha49], where the source shares k keys (i.e., uniformly at random generated packets) with each destination, as well as information packets encoded with the k keys. As a result, by observing any k edges, the eavesdropper cannot extract any information about the messages. The main novel observation in our scheme is that, although the source transmits a private message to each receiver, we do not need to necessarily use a private key to encrypt each private message, but instead we can re-use the same key for multiple destinations. Thus, in some cases, we need to multicast keys to the destinations, although we never need to multicast encoded messages. Moreover, this scheme

has the special property that we can isolate the key and encrypted message transmissions: we use some part of the network to convey (potentially multicast) the keys, and the remaining part to communicate the encrypted messages (i.e., the messages encoded with the keys). Our main result is stated in the following theorem.

Theorem 3. *The outer bound in* (2.1) *is tight for the case* $m = 2$*, i.e., the secure capacity region of the multiple unicast traffic over networks with single source and* $m = 2$ *destinations is*

$$R_1 \leq [M_{\{1\}} - k]^+ \ , \tag{2.2a}$$

$$R_2 \leq [M_{\{2\}} - k]^+ \ , \tag{2.2b}$$

$$R_1 + R_2 \leq [M_{\{1,2\}} - k]^+ \ . \tag{2.2c}$$

Proof. Clearly, from the result in Theorem 2, the rate region in (2.2) is an outer bound on the secure capacity region. Hence, we now need to prove that the rate region in (2.2) is also achievable. Towards this end, we start by providing the following definition of *separable* graphs, which we will leverage in the design of our scheme.

Definition 3 (Separable Graph)**.** *A graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *with a single source and* m *destinations is said to be* ***separable*** *if it can be partitioned into* $2^m - 1$ *edge disjoint graphs (graphs with empty edge sets are also allowed). These graphs are denoted as* $\mathcal{G}'_{\mathcal{J}} = (\mathcal{V}, \mathcal{E}'_{\mathcal{J}}), \mathcal{J} \subseteq [m], \mathcal{J} \neq \emptyset$ *and are such that* $\mathcal{E}'_{\mathcal{J}} \subseteq \mathcal{E}$ *and* $\mathcal{E}'_{\mathcal{J}} \cap \mathcal{E}'_{\mathcal{L}} = \emptyset, \ \forall \mathcal{J} \neq \mathcal{L} \subseteq [m]$*. Moreover, their min-cut capacities satisfy the following condition*

$$M_{\mathcal{A}} = \sum_{\substack{\mathcal{J} \subseteq [m] \\ \mathcal{J} \cap \mathcal{A} \neq \emptyset}} M'_{\mathcal{J}}, \ \forall \mathcal{A} \subseteq [m], \tag{2.3}$$

where, for the graph $\mathcal{G}$*,* $M_{\mathcal{A}}$ *is defined in Definition 2, and the graph* $\mathcal{G}'_{\mathcal{J}}$ *has the following min-cut capacities: (i)* $M'_{\mathcal{J}}$ *from the source* S *to any non-empty subset of destinations in* $\mathcal{J}$*, and (ii) zero from the source* S *to the set of destinations* $\{D_i : i \in [m] \setminus \mathcal{J}\}$*.*

To better understand the above definition, consider a graph $\mathcal{G}$ with $m = 2$ destinations. Then, the graph $\mathcal{G}$ is separable if it can be partitioned into 3 edge disjoint graphs such that:

- $\mathcal{G}'_{\{1\}}$ has the following min-cut capacities: $M'_{\{1\}}$ from S to D_1 and zero from S to D_2,

- $\mathcal{G}'_{\{2\}}$ has the following min-cut capacities: zero from S to D_1 and $M'_{\{2\}}$ from S to D_2,
- $\mathcal{G}'_{\{1,2\}}$ has the following min-cut capacities: $M'_{\{1,2\}}$ from S to D_1, $M'_{\{1,2\}}$ from S to D_2 and $M'_{\{1,2\}}$ from S to $\{D_1, D_2\}$,

where the quantities $M'_{\{1\}}$, $M'_{\{2\}}$ and $M'_{\{1,2\}}$ can be computed using the following set of equations:

$$M_{\{1\}} = M'_{\{1\}} + M'_{\{1,2\}}, \tag{2.4a}$$

$$M_{\{2\}} = M'_{\{2\}} + M'_{\{1,2\}}, \tag{2.4b}$$

$$M_{\{1,2\}} = M'_{\{1\}} + M'_{\{2\}} + M'_{\{1,2\}}. \tag{2.4c}$$

An example of separable graph for $m = 2$ and its partition into 3 edge disjoint graphs is shown in Fig. 2.1. We now state the following lemma, which is a consequence of [RW09, Theorem 1] and which we will use to prove the achievability of the rate region in (2.2).

Lemma 4. *[RW09, Theorem 1]: Any graph with a single source and $m = 2$ destinations is separable.*

By leveraging the result in Lemma 4, we are now ready to prove Theorem 3. In particular, we consider two cases depending on the value of k (i.e., the number of edges that the eavesdropper eavesdrops). Without loss of generality, we assume that $k < \min_{i\in[2]} M_{\{i\}}$, as otherwise secure communication to the set of destinations $\{D_i : k \geq M_{\{i\}}, i \in [2]\}$ is not possible at any positive rate, and hence we can just remove this set of destinations from the network. To secure our messages from the adversary, we use uniform random packets generated at the source, which we refer to as *keys*. We will transmit these keys as well as the messages encoded with these keys over the network. The security of our schemes relies on two aspects: (i) a message encoded with a uniform random key is independent of the message and is distributed uniformly, and (ii) the amount of keys that we use is such that the eavesdropper cannot collect a sufficient number of keys and encoded messages to be able to extract any information on the messages.

1. **Case 1:** $k \geq M'_{\{1,2\}}$. In this case, by substituting the quantities in (2.4) into (2.2), we obtain that the constraint in (2.2c) is redundant. Thus, we will now prove that the rate pair

$(R_1, R_2) = (M_{\{1\}} - k, M_{\{2\}} - k)$ is securely achievable, which along with the time-sharing argument proves the achievability of the entire rate region in (2.2).

We denote with $K_1, K_2, \ldots, K_k$ the k key packets and with $W_i^{(1)}, W_i^{(2)}, \ldots, W_i^{(R_i)}$ (with $i \in [2]$) the R_i message packets for D_i. With this, our scheme is as follows:

- We multicast the key packets $K_i, \forall i \in [M'_{\{1,2\}}]$, to both D_1 and D_2 using $\mathcal{G}'_{\{1,2\}}$, which has edges denoted by $\mathcal{E}'_{\{1,2\}}$. This is possible since $\mathcal{G}'_{\{1,2\}}$ has a min-cut capacity $M'_{\{1,2\}}$ to both D_1 and D_2 (see Definition 3).
- We unicast the key packets $K_\ell, \forall \ell \in [M'_{\{1,2\}} + 1 : k]$, to $D_i, \forall i \in [2]$, using $k - M'_{\{1,2\}}$ paths out of the $M'_{\{i\}}$ disjoint paths in $\mathcal{G}'_{\{i\}}$. We denote by $\hat{\mathcal{E}}_{\{i\}}$ the set that contains all the first edges of these paths. Clearly, $|\hat{\mathcal{E}}_{\{i\}}| = k - M'_{\{1,2\}}, \forall i \in [2]$. Notice that $\hat{\mathcal{E}}_{\{i\}} \subseteq \mathcal{E}'_{\{i\}}, \forall i \in [2]$ (see Definition 3).
- We send the $R_i, \forall i \in [2]$, encrypted message packets (i.e., encoded with the keys) of D_i on the remaining $M'_{\{i\}} - k + M'_{\{1,2\}}$ disjoint paths in $\mathcal{G}'_{\{i\}}$. We denote by $\bar{\mathcal{E}}_{\{i\}}$ the set that contains all the first edges of these paths in $\mathcal{G}'_{\{i\}}$. Clearly, $|\bar{\mathcal{E}}_{\{i\}}| = R_i, \forall i \in [2]$, $\bar{\mathcal{E}}_{\{i\}} \subseteq \mathcal{E}'_{\{i\}}$ and $\bar{\mathcal{E}}_{\{i\}} \cap \hat{\mathcal{E}}_{\{i\}} = \emptyset$ (see Definition 3).

This scheme achieves $R_i = M'_{\{i\}} - k + M'_{\{1,2\}} = M_{\{i\}} - k, \forall i \in [1:2]$, where the second equality follows by using the definitions in (2.4). Now we prove that this scheme is also secure. We start by noticing that, thanks to Definition 3, the edge sets $\mathcal{E}'_{\{1,2\}}$, $\hat{\mathcal{E}}_{\{i\}}$ and $\bar{\mathcal{E}}_{\{i\}}$, with $i \in [2]$, are disjoint. We write these transmissions in a matrix form (with G and U being the encoding matrices of size $\ell \times k$ and $(R_1 + R_2) \times k$, respectively) and we obtain

$$\begin{bmatrix} X_{\mathcal{E}'_{\{1,2\}}} \\ X_{\hat{\mathcal{E}}_{\{1\}}} \\ X_{\hat{\mathcal{E}}_{\{2\}}} \end{bmatrix} = \underbrace{\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1k} \\ g_{21} & g_{22} & \cdots & g_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ g_{\ell 1} & g_{\ell 2} & \cdots & g_{\ell k} \end{bmatrix}}_{G} \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_k \end{bmatrix}, \ \ell = |\mathcal{E}'_{\{1,2\}}| + 2\left(k - M'_{\{1,2\}}\right) \ ,$$

$$\begin{bmatrix} X_{\bar{\mathcal{E}}_{\{1\}}} \\ X_{\bar{\mathcal{E}}_{\{2\}}} \end{bmatrix} = \underbrace{\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1k} \\ u_{21} & u_{22} & \dots & u_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{r1} & u_{r2} & \dots & u_{rk} \end{bmatrix}}_{U} \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_k \end{bmatrix} + \begin{bmatrix} W_1^{(1)} \\ \vdots \\ W_1^{(R_1)} \\ W_2^{(1)} \\ \vdots \\ W_2^{(R_2)} \end{bmatrix}, \ r = R_1 + R_2 \ .$$

We here highlight that on the remaining edges $\mathcal{E}\backslash\{\mathcal{E}'_{\{1,2\}} \cup \hat{\mathcal{E}}_{\{1\}} \cup \bar{\mathcal{E}}_{\{1\}} \cup \hat{\mathcal{E}}_{\{2\}} \cup \bar{\mathcal{E}}_{\{2\}}\}$ of the network, we either do not transmit any symbol or simply route the symbols from $X_{\bar{\mathcal{E}}_{\{1\}}}$, $X_{\bar{\mathcal{E}}_{\{2\}}}$, $X_{\hat{\mathcal{E}}_{\{1\}}}$, and $X_{\hat{\mathcal{E}}_{\{2\}}}$ (corresponding to the symbols transmitted on disjoint paths). Thus, without loss of generality, we can assume that Eve eavesdrops at most k edges from $\{\mathcal{E}'_{\{1,2\}} \cup \hat{\mathcal{E}}_{\{1\}} \cup \bar{\mathcal{E}}_{\{1\}} \cup \hat{\mathcal{E}}_{\{2\}} \cup \bar{\mathcal{E}}_{\{2\}}\}$. In what follows, we let: (i) X denote the vector of the symbols transmitted over these edges, (ii) K be the vector of the k random key packets, and (iii) W be the vector of the message packets for both destinations. With this, X can be represented in a matrix form as

$$X = \begin{bmatrix} 0_{\ell \times (R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix} \begin{bmatrix} W \\ K \end{bmatrix}. \tag{2.5}$$

We highlight that the first $|\mathcal{E}'_{\{1,2\}}|$ rows of G (i.e., those that correspond to multicasting the keys) are determined by the network coding scheme for multicasting [ACL00]. As such, they can be constructed in $\mathcal{O}\left(|\mathcal{E}|^3\right)$ by using the multicasting scheme of [JSC05], which requires a finite field of size $m = 2$. Thus, the security follows if we can show that for any choice of G, there exists a U such that (2.5) satisfies the condition in Lemma 1. This is proved in Appendix A.1 where we show that, over a sufficiently large finite field, a random choice of U in (2.5) satisfies the condition in Lemma 1 with high probability. Thus, the rate pair $(R_1, R_2) = (M_{\{1\}} - k, M_{\{2\}} - k)$ is securely achievable.

2. **Case 2:** $k < M'_{\{1,2\}}$. By substituting the quantities in (2.4), the rate region in (2.2) becomes

$$R_i \leq M_{\{i\}} - k = M'_{\{i\}} + M'_{\{1,2\}} - k, \forall i \in [2] \ , \tag{2.6a}$$

$$R_1 + R_2 \leq M_{\{1,2\}} - k = M'_{\{1\}} + M'_{\{2\}} + M'_{\{1,2\}} - k \ . \tag{2.6b}$$

We now show that we can achieve the following two corner points i.e., the rate pair

$$\begin{aligned}(R_1, R_2) &= \big((1-\alpha)(M_{\{1,2\}} - M_{\{2\}}) + \alpha(M_{\{1\}} - k), \\ &\qquad (1-\alpha)(M_{\{2\}} - k) + \alpha(M_{\{1,2\}} - M_{\{1\}})\big) \\ &\overset{\text{(a)}}{=} (M'_{\{1\}} + \alpha(M'_{\{1,2\}} - k), M'_{\{2\}} + (1-\alpha)(M'_{\{1,2\}} - k)) \ , \end{aligned} \tag{2.7}$$

for $\alpha \in \{0, 1\}$, where the equality in (a) follows by using the definitions in (2.4). This, along with the time-sharing argument, proves the achievability of the entire rate region in (2.6). We recall that we denote with $K_1, K_2, \ldots, K_k$ the k key packets and with $W_i^{(1)}, W_i^{(2)}, \ldots, W_i^{(R_i)}$ (with $i \in [2]$) the R_i message packets for D_i. With this, our scheme is as follows:

- Using the graph $\mathcal{G}'_{\{1,2\}}$ we multicast to both destinations D_1 and D_2: (i) $K_i, \forall i \in [k]$, (ii) $\alpha(M'_{\{1,2\}} - k)$ encrypted message packets (i.e., formed by encoding W_1 and the keys K) for D_1 and (iii) $(1-\alpha)(M'_{\{1,2\}} - k)$ encrypted message packets (i.e., formed by encoding W_2 and the keys K) for D_2. Recall that the edges of the graph $\mathcal{G}'_{\{1,2\}}$ are denoted by $\mathcal{E}'_{\{1,2\}}$ (see Definition 3). Note that, since all these packets are multicast, then D_1 might also receive packets that are for D_2, and vice versa. However, note that, since the eavesdropper is external, i.e., it is not one of the destinations, then this does not violate the security condition, as long as the adversary, who eavesdrops any k edges of her choice, does not learn anything about the content of the messages. We also highlight that the message packets multicast to the two destinations are encoded using the key packets, where the encoding is based on the secure network coding result on multicasting [CY02], which ensures perfect security from an adversary eavesdropping any k edges.
- We send $M'_{\{i\}}$ encrypted message packets of D_i (i.e., encoded by using the k key packets) on the $M'_{\{i\}}$ disjoint paths to D_i in the graph $\mathcal{G}'_{\{i\}}$, and denote by $\hat{\mathcal{E}}_{\{i\}}$ the set that contains all the first edges of these paths for $i \in [2]$.

This scheme achieves the rate pair in (2.7). Now we prove that this scheme is also secure. For ease of representation, in what follows we let $R'_1 = \alpha(M'_{\{1,2\}} - k)$ and $R'_2 = (1-\alpha)(M'_{\{1,2\}} - k)$. We again notice that, thanks to Definition 3, the edge sets $\mathcal{E}'_{\{1,2\}}$, $\hat{\mathcal{E}}_{\{1\}}$ and

$\hat{\mathcal{E}}_{\{2\}}$ are disjoint. We write these transmissions in a matrix form (with G, S and U being the encoding matrices of sizes $\ell \times k$, $\ell \times t$ and $r \times k$ respectively) and we obtain,

$$X_{\mathcal{E}'_{\{1,2\}}} = \underbrace{\begin{bmatrix} g_{11} & g_{12} & \dots & g_{1k} \\ g_{21} & g_{22} & \dots & g_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ g_{\ell 1} & g_{\ell 2} & \dots & g_{\ell k} \end{bmatrix}}_{G} \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_k \end{bmatrix} + \underbrace{\begin{bmatrix} s_{11} & s_{12} & \dots & s_{1t} \\ s_{21} & s_{22} & \dots & s_{1t} \\ \vdots & \vdots & \ddots & \vdots \\ s_{\ell 1} & s_{\ell 2} & \dots & s_{\ell t} \end{bmatrix}}_{S} \begin{bmatrix} W_1^{(1)} \\ \vdots \\ W_1^{(R'_1)} \\ W_2^{(1)} \\ \vdots \\ W_2^{(R'_2)} \end{bmatrix},$$

where $\ell = |\mathcal{E}'_{\{1,2\}}|$ and $t = R'_1 + R'_2$, and

$$\begin{bmatrix} X_{\hat{\mathcal{E}}_{\{1\}}} \\ X_{\hat{\mathcal{E}}_{\{2\}}} \end{bmatrix} = \underbrace{\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1k} \\ u_{21} & u_{22} & \dots & u_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{r1} & u_{r2} & \dots & u_{rk} \end{bmatrix}}_{U} \begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_k \end{bmatrix} + \begin{bmatrix} W_1^{(R'_1+1)} \\ \vdots \\ W_1^{(R_1)} \\ W_2^{(R'_2+1)} \\ \vdots \\ W_2^{(R_2)} \end{bmatrix}, \; r = R_1 + R_2 - (M'_{\{1,2\}} - k) \; .$$

In what follows, we let: (i) X denote the vector of the symbols transmitted over the edges $\mathcal{E}'_{\{1,2\}}$, $\hat{\mathcal{E}}_{\{1\}}$ and $\hat{\mathcal{E}}_{\{2\}}$, (ii) K be the vector of the k random key packets, and (iii)

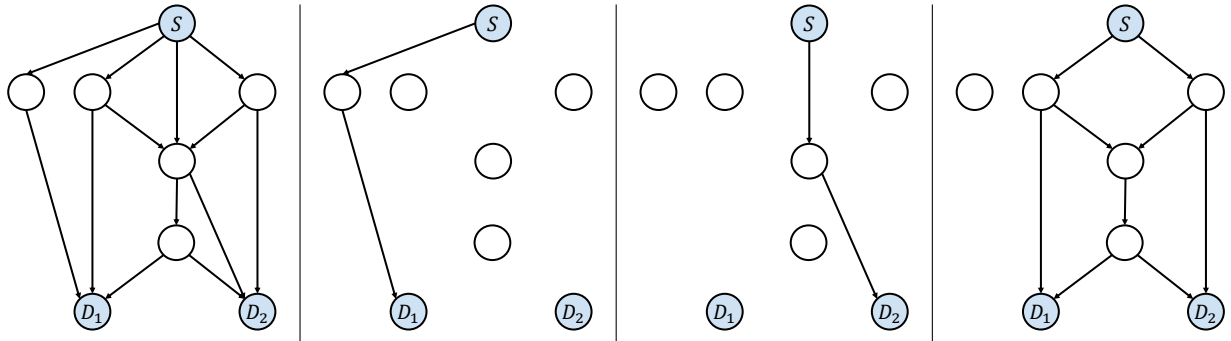
$$W' := \begin{bmatrix} W_1^{(1)} \\ \vdots \\ W_1^{(R'_1)} \\ W_2^{(1)} \\ \vdots \\ W_2^{(R'_2)} \end{bmatrix}, \; W'' := \begin{bmatrix} W_1^{(R'_1+1)} \\ \vdots \\ W_1^{(R_1)} \\ W_2^{(R'_2+1)} \\ \vdots \\ W_2^{(R_2)} \end{bmatrix}.$$

With this, X can be represented in a matrix form as

$$X = \begin{bmatrix} X_{\mathcal{E}'_{\{1,2\}}} \\ X_{\hat{\mathcal{E}}_{\{1\}}} \\ X_{\hat{\mathcal{E}}_{\{2\}}} \end{bmatrix} = \begin{bmatrix} S & 0_{\ell \times r} & G \\ 0_{r \times t} & I_r & U \end{bmatrix} \begin{bmatrix} W' \\ W'' \\ K \end{bmatrix}. \tag{2.8}$$

Similar to Case 1, on the remaining edges $\mathcal{E}\backslash\{\mathcal{E}'_{\{1,2\}} \cup \hat{\mathcal{E}}_{\{1\}} \cup \hat{\mathcal{E}}_{\{2\}}\}$ of the network, we either do not transmit any symbol or simply route the symbols from $\{X_{\hat{\mathcal{E}}_{\{1\}}}, X_{\hat{\mathcal{E}}_{\{2\}}}\}$ (corresponding to the symbols transmitted on disjoint paths). Thus, without loss of generality, we can assume that the eavesdropper eavesdrops at most k edges from $\{\mathcal{E}'_{\{1,2\}} \cup \hat{\mathcal{E}}_{\{1\}} \cup \hat{\mathcal{E}}_{\{2\}}\}$. We highlight that the matrices G and S are determined by the secure network coding scheme for multicasting [CY02]. As such, they can be constructed in $\mathcal{O}\left(k^2|\mathcal{E}|^{k+2}\right)$ by using the scheme of [KOK17], which requires a finite field of size $|\mathcal{E}|^k$. Thus, security follows if we can show that for any choice of S and G satisfying the security condition in Lemma 1, i.e., $rk\left(\left[\begin{array}{cc} S & G \end{array}\right]\Big|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right), \forall |\mathcal{Z}| \leq k$, there exists a choice of U such that the security condition in Lemma 1 is satisfied for (2.8). This is proved in Appendix A.2 where we show that, over a sufficiently large finite field, a random choice of U in (2.8) satisfies the condition in Lemma 1 with high probability.

This concludes the proof of Theorem 3. □



(a) $\mathcal{G}_0$. (b) $\mathcal{G}'_{\{1\}}$. (c) $\mathcal{G}'_{\{2\}}$. (d) $\mathcal{G}'_{\{1,2\}}$.

Figure 2.1: A 2-destination separable network $\mathcal{G}_0$ in (a) and its partition into 3 edge disjoint graphs in (b), (c) and (d). Here, $M'_{\{1\}} = M'_{\{2\}} = 1$, and $M'_{\{1,2\}} = 2$.

Example 1: We here illustrate the above described scheme for the network $\mathcal{G}_0$ in Fig. 2.1(a). We first note that $\mathcal{G}_0$ has min-cut capacities $M_{\{1\}} = M_{\{2\}} = 3$ and $M_{\{1,2\}} = 4$, and it can be partitioned into three edge disjoint graphs $\mathcal{G}'_{\mathcal{J}}, \mathcal{J} \subseteq \{1,2\}, \mathcal{J} \neq \emptyset$ as shown in Figs. 2.1(b)-(d), with min-cut capacities equal to $M'_{\{1\}} = M'_{\{2\}} = 1$ and $M'_{\{1,2\}} = 2$, respectively. We assume that the adversary eavesdrops any $k = 2$ edges of her choice. For this case, the source should be able to securely

communicate at a rate $(R_1, R_2) = (1, 1)$ towards the $m = 2$ destinations. This rate pair can be achieved using two key packets K_1 and K_2 and operations over $\mathbb{F}_4$ as follows:

1. Over the set of edges in $\mathcal{G}'_{\{1\}}$, the source transmits $W_1 + K_1 + 2K_2$; the intermediate node simply routes this transmission to D_1;

2. Over the set of edges in $\mathcal{G}'_{\{2\}}$, the source transmits $W_2 + K_1 + 3K_2$; the intermediate node simply routes this transmission to D_2;

3. Over the set of edges in $\mathcal{G}'_{\{1,2\}}$, the source multicasts K_1 and K_2 to the receivers. It transmits K_1 to one intermediate node and K_2 to the other intermediate node. The intermediate node denoted as i in Fig. 2.1(d) receives K_1 and K_2 and transmits $K_1 + K_2$ on its outgoing edges. Thus D_1 and D_2 receive both K_1 and K_2. It therefore follows that $D_i, i \in [2]$, can successfully recover W_i. ■

We conclude this section with some observations on separable graphs. As highlighted in the proof of Theorem 3, given the separation of a graph into subgraphs, our capacity achieving scheme is polynomial-time. However, identifying the subgraphs with the required min-cut properties is not an easy problem [RW09], and it is not clear if it can be performed in polynomial-time. Moreover, although for the case of $m = 2$ destinations any graph is separable (see [RW09, Theorem 1]), in general the same does not hold for $m \geq 3$, as the following example illustrates.

Example 2: Consider the network in Fig. 2.2, which consists of $m = 3$ destinations and has the following min-cut capacities: $M_{\{1\}} = 1$, $M_{\{2\}} = 1$, $M_{\{3\}} = 1$, $M_{\{1,2\}} = 2$, $M_{\{2,3\}} = 2$, $M_{\{1,3\}} = 2$ and $M_{\{1,2,3\}} = 2$. With this, we can find $M'_{\mathcal{J}},\ \mathcal{J} \subseteq [3]$, by solving (2.3). In particular, we obtain: $M'_{\{1\}} = M'_{\{2\}} = M'_{\{3\}} = 0$, $M'_{\{1,2\}} = M'_{\{2,3\}} = M'_{\{1,3\}} = 1$ and $M'_{\{1,2,3\}} = -1$. Since a graph can not have a negative min-cut capacity, we readily conclude that a separation of the form defined in Definition 3 is not possible. ■

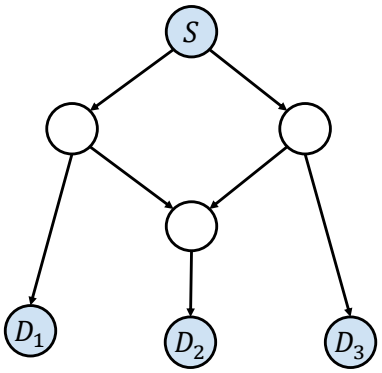


Figure 2.2: Example of a non-separable graph.

2.5 Secure scheme for two-layer networks

In Section 2.4, we characterized the secure capacity region of networks with $m = 2$ destinations, by leveraging the separability property. In this section, we discuss separable networks with arbitrary number of destinations and characterize the capacity region of networks with: (i) $m = 3$ destinations, where the adversary eavesdrops any arbitrary k edges of her choice, (ii) networks with arbitrary number m of destinations, where the adversary eavesdrops any $k = 1$ edge of her choice, and (iii) networks with arbitrary values of k and m for which the min-cut capacities satisfy certain properties. Towards this end, we will first consider a special class of separable networks, namely networks having a two-layer topology, and design a secure scheme for this class of networks. We will then show that, in order to characterize the secure capacity region of any separable network, it is sufficient to study two-layer networks. In particular, we will prove that any separable network can be modeled as a two-layer network with the same min-cut capacities, and that a secure scheme for a two-layer network can be transformed into a secure scheme for its corresponding separable network. We now proceed by formally defining the two-layer network topology.

Definition 4. *A two-layer network consists of one source S that wishes to communicate with m destinations, by hopping information through one layer of t relays. As such, a two-layer network is parameterized by: (i) the integer t, which denotes the number of relays in the first layer; (ii) the integer m, which indicates the number of destinations in the second layer; (iii) m sets $\mathcal{M}_i$, $i \in [m]$, such that $\mathcal{M}_i \subseteq [t]$, where $\mathcal{M}_i$ contains the indexes of the relays connected to destination D_i.*

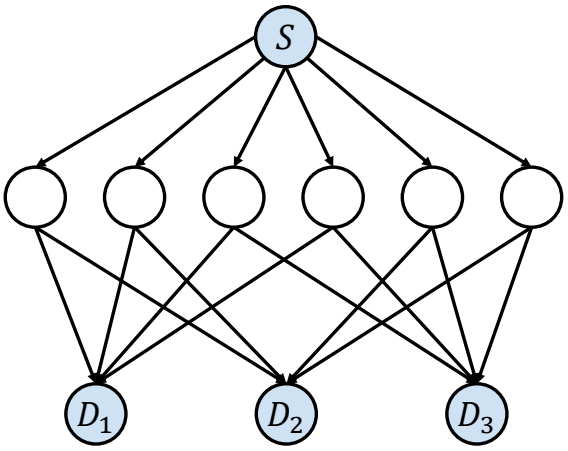


Figure 2.3: Two-layer network example which illustrates that using different parts of the network to transmit the keys and the encrypted messages is not optimal. In this network $\mathcal{M}_1 = \{1, 2, 3, 4\}$, $\mathcal{M}_2 = \{1, 2, 5, 6\}$ and $\mathcal{M}_3 = \{3, 4, 5, 6\}$.

An example of a two-layer network is shown in Fig. 2.3, for which $t = 6$, $m = 3$, $\mathcal{M}_1 = \{1, 2, 3, 4\}$, $\mathcal{M}_2 = \{1, 2, 5, 6\}$ and $\mathcal{M}_3 = \{3, 4, 5, 6\}$.

Before delving into the study of such two-layer networks, recall that the capacity-achieving scheme for $m = 2$ destinations described in Section 2.4 uses some parts of the network to convey (potentially multicasting) the keys and the remaining part to communicate the encrypted messages. Therefore, we now ask the following question: can we extend this idea to get a capacity-achieving scheme for separable networks with arbitrary number of destinations? In other words, can we spatially isolate the key from the message transmission? The next example shows that this is not possible through an example.

Example 3: Consider the two-layer network shown in Fig. 2.3, which consists of $m = 3$ destinations, and where the adversary can eavesdrop any $k = 3$ edges of her choice. For this network we have the following min-cut capacities: $M_{\{1\}} = M_{\{2\}} = M_{\{3\}} = 4$, $M_{\{1,2\}} = M_{\{1,3\}} = M_{\{2,3\}} = M_{\{1,2,3\}} = 6$. We would like to show that the triple $(R_1, R_2, R_3) = (1, 1, 1)$ – obtained from the outer bound in Theorem 2 – can not be achieved when the key packets and the encrypted messages are transmitted over different parts of the network. It is not difficult to see that, out of the 6 outgoing edges from the source, multicasting 3 keys[1] requires a number of edges strictly greater than

[1]Note that 3 keys are required since the adversary eavesdrops $k = 3$ edges of her choice.

4. Thus, we would be left with strictly less than 2 edges, which are not sufficient to transmit 3 message packets, i.e., one for each destination. It therefore follows that, with this strategy, the rate triple $(R_1, R_2, R_3) = (1, 1, 1)$ can not be securely achieved.

However, let the source transmits the following symbols on its outgoing edges

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 4 \\ 1 & 0 & 0 & 1 & 3 & 2 \\ 4 & 6 & 4 & 1 & 4 & 2 \\ 2 & 4 & 2 & 1 & 5 & 4 \end{bmatrix}}_{B} \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ K_1 \\ K_2 \\ K_3 \end{bmatrix}, \tag{2.9}$$

where $B \in \mathbb{F}_7^{6\times 6}$ is the encoding matrix. If the intermediate nodes simply route the received symbols, then we can achieve the rate tuple $(1, 1, 1)$. This is because, the encoding matrix B can be written as

$$B = \begin{bmatrix} B' & B'' \end{bmatrix},$$

where B' contains the first three columns of B in (2.9), and B'' contains the last three columns of B in (2.9). Thus, it follows that

$$rk\left(\begin{bmatrix} B' & B'' \end{bmatrix}\Big|_{\mathcal{Z}}\right) = rk\left(B''\Big|_{\mathcal{Z}}\right), \ \forall |\mathcal{Z}| \leq 3,$$

which, from Lemma 1, implies that the encoding in (2.9) is secure.

Moreover, each destination can decode its respective message as follows:

- Destination 1: $W_1 = 6X_1 + 3X_2 + 4X_3 + X_4$,
- Destination 2: $W_2 = 6X_1 + 4X_2 + 3X_5 + X_6$,
- Destination 3: $W_3 = 5X_3 + 6X_4 + X_5 + 2X_6$.

Thus, the rate triple $(R_1, R_2, R_3) = (1, 1, 1)$ can be securely achieved. This example shows that using different parts of the network to transmit the keys and the encrypted messages, in general is not optimal. This is partially due to the fact that destinations do not need to decode each key individually, as long as they can successfully recover their message. ■

2.5.1 Secure transmissions scheme

For two-layer networks, we have $M_{\mathcal{A}} = |\cup_{i\in\mathcal{A}}\mathcal{M}_i|$. For notational convenience, we let $M_{\cap\{i,j\}} = |\mathcal{M}_i \cap \mathcal{M}_j|$ and $M_{\cap\{i,\mathcal{A}\}} = |\mathcal{M}_i \cap (\cup_{j\in\mathcal{A}}\mathcal{M}_j)\,|$. Moreover, we also assume that $M_{\{i\}} \geq k, \forall i \in [m]$ (otherwise secure communication is not possible) with $M_{\emptyset} := k$ for consistency.

We here propose a polynomial-time (see Lemma 6) secure transmission scheme for two-layer networks. In Section 2.5.2, we will then derive its achieved rate region. The source S encodes the message packets with k random packets and transmits these packets on its outgoing edges to the t relays. We can write the received symbols at the t relays as

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_t \end{bmatrix} = \left[\begin{array}{c|c} & \\ M & V \\ & \end{array}\right] \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_m \\ K \end{bmatrix}, \tag{2.10}$$

where: (i) $W_i, i \in [m]$ is a column vector of R_i message packets for destination D_i, (ii) K is a column vector which contains the k random packets, (iii) M is an encoding matrix of dimension $t \times (\sum_{i=1}^{m} R_i)$ (as we will show below such a matrix can always be constructed so that all the destinations correctly decode their intended message), and (iv) V is a Vandermonde matrix of dimension $t \times k$. The matrix V is chosen for *s*ecurity purposes, i.e., any set of k rows of V are linearly independent and hence Lemma 1 ensures that, no matter which k rows (i.e., edges) Eve eavesdrops, she will learn nothing about the messages $W_{[m]}$.

Remark 1. *The only property of V that we require in our scheme is the Maximum Distance Separable (MDS) property (i.e., any k rows of V are linearly independent). This implies that, even if we select a random matrix $\tilde{V}$ instead of V, with high probability (close to 1 for large field size) we will have a secure scheme for the two-layer network. This also implies that a finite field of size $\mathcal{O}(|\mathcal{E}|)$ can deterministically provide such a matrix $\tilde{V}$.*

Each relay $i \in [t]$ will then forward the received symbol X_i in (2.10) to the destinations to which it is connected. As such, each destination will observe a subset of symbols from $\{X_1, X_2, \ldots, X_t\}$

(depending on which of the t relays it is connected to). Finally, destination $D_i, i \in [m]$ selects a decoding vector and performs the inner product with $[X_1, X_2, \ldots, X_t]$. In particular, this decoding vector is chosen such that it has two characteristics: (1) it is in the left null space of V, i.e., in the right null space of V^T; this ensures that each destination is able to cancel out the random packets (encoded with the message packets); (2) it has zeros in the positions corresponding to the relays it is not connected to; this ensures that each destination uses only the symbols that it actually observes. In other words, all the decoding vectors that D_i can choose belong to the right null space N_i of the matrix V_i defined

$$V_i = \begin{bmatrix} V^T \\ C_i \end{bmatrix}, \tag{2.11}$$

where C_i is a matrix of dimension $\bar{t} \times t$, with $\bar{t}$ being the number of relays to which D_i is not connected to. In particular, each row of C_i has all zeros except a one in the position corresponding to a relay to which D_i is not connected to.

For instance, for the network in Fig. 2.3, we have

$$C_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; C_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \; C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We let T be a matrix of dimension $(\sum_{i=1}^{m} R_i) \times t$ that, for each destination $D_i, i \in [m]$, contains R_i decoding vectors that belong to the right null space of the matrix V_i in (2.11), denoted as N_i. Mathematically, we have

$$T = \begin{bmatrix} - - - - d_1^{(1)} - - - - \\ - - - - d_2^{(1)} - - - - \\ \vdots \\ - - - - d_{R_1}^{(1)} - - - - \\ - - - - d_1^{(2)} - - - - \\ \vdots \\ - - - - d_{R_m}^{(m)} - - - - \end{bmatrix}, \tag{2.12}$$

where $d_j^{(i)}$ denotes the j-th decoding vector (of length t) selected from the null space N_i, with $i \in [m], j \in [R_i]$. Note that, if for all $i \in [m]$, we can select R_i decoding vectors from N_i such

that all the $d_j^{(i)}$ in (2.12) are linearly independent (i.e., such that T has a full row rank), then it is possible to construct the matrix M in (2.10) such that

$$TM = I_{\left(\sum_{i=1}^{m} R_i\right)}, \tag{2.13}$$

which ensures that all the destinations are able to correctly decode their intended message as

$$\begin{bmatrix} \hat{W}_1 \\ \vdots \\ \hat{W}_m \end{bmatrix} = T \begin{bmatrix} X_1 \\ \vdots \\ X_t \end{bmatrix} \overset{(2.10)}{=} \begin{bmatrix} TM & TV \end{bmatrix} \begin{bmatrix} W_1 \\ \vdots \\ W_m \\ K \end{bmatrix}$$
$$= TM \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix} + TVK = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix}.$$

In Appendix A.3, we propose an iterative algorithm (of polynomial-complexity as formally proved in Lemma 6) to select $R_i, i \in [m]$ decoding vectors from N_i such that T in (2.12) has indeed a full row rank. The performance of the proposed algorithm is provided in the following lemma, which is also proved in Appendix A.3.

Lemma 5. *For any given permutation $\pi = \{\pi(1), \ldots, \pi(m)\}$ of $[m]$, it is possible to select*

$$R_{\pi(i)} = \textit{dim}\left(\sum_{j=1}^{i} N_{\pi(j)}\right) - \textit{dim}\left(\sum_{j=1}^{i-1} N_{\pi(j)}\right),\ i \in [m], \tag{2.14}$$

vectors from $N_{\pi(i)}$ so that all the $\sum_{i=1}^{m} R_i$ selected vectors are linearly independent.

Remark 2. *Note that, since there are $m!$ possible permutations of $[m]$, then Lemma 5 offers $m!$ possible choices for selecting $R_i, i \in [m]$ vectors from N_i so that all the $\sum_{i=1}^{m} R_i$ selected vectors are linearly independent. We prove in Lemma 7 that these choices form the corner points of the secure rate region achieved by our scheme.*

Remark 3. *The result in Lemma 5 implies that rate m-tuple $(R_1, R_2, \ldots, R_m)$, with $R_i, i \in [m]$ being defined in* (2.14)*, can be securely achieved by our proposed scheme.*

The following lemma analyzes the complexity of designing our proposed secure scheme.

Lemma 6. *The complexity of designing the secure transmission scheme in* (2.10) *equals* $\mathcal{O}(m|\mathcal{E}|^4)$. *Moreover, a field size of dimension* $q \geq |\mathcal{E}|$ *is sufficient.*

Proof. To achieve any rate m-tuple $(R_1, R_2, \ldots, R_m)$ using our scheme, we need to find a basis of null spaces $N_i,\ \forall i \in [m]$ and then use the iterative algorithm proposed in Appendix A.3 to form the decoding matrix T in (2.12). A basis of the null space N_i can be found using the Gaussian elimination algorithm, which has a complexity of $\mathcal{O}(|\mathcal{E}|^3)$ [AHM07]. The iterative algorithm in Appendix A.3 for selecting decoding vectors in these null spaces requires discarding dependent vectors, which has a complexity of $\mathcal{O}(m|\mathcal{E}||\mathcal{E}|^3)$. This follows since: (i) there are at most $m|\mathcal{E}|$ vectors in the basis of these null spaces, and (ii) to check if each vector is dependent on the previously selected vectors, we require $\mathcal{O}(|\mathcal{E}|^3)$ computations using the Gaussian elimination algorithm. Finally, given the decoding matrix T, we require the computation of the encoding matrix M which, as highlighted in (2.13), is the right inverse of T. Thus, computing M requires $\mathcal{O}(|\mathcal{E}|^3)$ operations by again using the Gaussian elimination algorithm. It therefore follows that the overall complexity of our secure transmission scheme is $\mathcal{O}(m|\mathcal{E}|^4)$.

As discussed in Remark 1, to ensure security we are using only the MDS property of the Vandermonde matrix V in (2.10). The size of this matrix is $t \times k$, and $t \leq |\mathcal{E}|$. Thus, a field size of dimension $|\mathcal{E}|$ is sufficient. This concludes the proof of Lemma 6. □

In the next section, we will leverage the result in Lemma 5 and Remark 2 to derive the secure rate region achieved by our proposed scheme.

2.5.2 Achieved secure rate region

In this section, we derive the rate region achieved by the secure scheme described in Section 2.5.1. In particular, we have the following lemma, whose proof is in Appendix A.4.

Lemma 7. *The secure rate region achieved by the proposed scheme is given by*

$$0 \leq \sum_{i \in \mathcal{A}} R_i \leq \textit{dim}\left(\sum_{i \in \mathcal{A}} N_i\right),\ \forall \mathcal{A} \subseteq [m], \tag{2.15}$$

where N_i *is the right null space of the matrix* V_i *in* (2.11).

In the remainder of this section, we prove that the secure rate region in (2.15) is indeed the secure *c*apacity region when: (i) the adversary eavesdrops any $k = 1$ edge of her choice (and arbitrary m); (ii) there are $m = 3$ destinations (and arbitrary k); (iii) k and m are arbitrary, but the network has some special structure in terms of minimum cut.

2.5.3 Secure capacity for $k = 1$, m arbitrary

In this section, we consider the case where Eve eavesdrops any $k = 1$ edge of her choice, and characterize the secure capacity region. In particular, we prove the following theorem.

Theorem 8. *For the two-layer network when Eve eavesdrops any $k = 1$ edge of her choice, the secure capacity region is*

$$\sum_{i\in\mathcal{A}} R_i \leq M_{\mathcal{A}} - C_{\mathcal{A}},\ \forall \mathcal{A} \subseteq [m], \tag{2.16}$$

with $C_{\mathcal{A}}$ being the number of connected components in an undirected graph where: (i) there are $|\mathcal{A}|$ nodes, i.e., one for each $i \in \mathcal{A}$; (ii) an edge between node i and node j, $\{i, j\} \in \mathcal{A}$, $i \neq j$, exists if $\mathcal{M}_i \cap \mathcal{M}_j \neq \emptyset$.

2.5.3.1 Outer bound

We show that the outer bound in Theorem 2 can be equivalently written as in (2.16). Let $\mathcal{V}_i$, $i \in [C_{\mathcal{A}}]$, represent the set of nodes in the i-th component of the graph constructed as explained in Theorem 8. Then, clearly $\mathcal{A} = \bigsqcup_{i=1}^{C_{\mathcal{A}}} \mathcal{V}_i$ and we can write

$$\begin{aligned}
\sum_{i\in\mathcal{A}} R_i &= \sum_{i\in\mathcal{V}_1} R_i + \sum_{i\in\mathcal{V}_2} R_i + \ldots + \sum_{i\in\mathcal{V}_{C_{\mathcal{A}}}} R_i \\
&\overset{\text{(a)}}{\leq} (M_{\mathcal{V}_1} - k) + (M_{\mathcal{V}_2} - k) + \ldots + \left(M_{\mathcal{V}_{C_{\mathcal{A}}}} - k\right) \\
&\overset{\text{(b)}}{=} M_{\mathcal{V}_1\cup\mathcal{V}_2\cup\ldots\cup\mathcal{V}_{C_{\mathcal{A}}}} - kC_{\mathcal{A}} \\
&\overset{\text{(c)}}{=} M_{\mathcal{A}} - kC_{\mathcal{A}} \\
&\overset{k=1}{=} M_{\mathcal{A}} - C_{\mathcal{A}},
\end{aligned}$$

where: (i) the inequality in (a) follows by applying (2.1) for each set $\mathcal{V}_i, i \in [C_{\mathcal{A}}]$, (ii) the equality in (b) follows since, by construction, $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$ for all $i \in \mathcal{V}_x$ and $j \in \mathcal{V}_y$ with $x \neq y$, and

(iii) the equality in (c) follows since $\mathcal{A} = \bigsqcup_{i=1}^{C_{\mathcal{A}}} \mathcal{V}_i$. Thus, (2.1) implies (2.16). Moreover, since $C_{\mathcal{A}} \geq 1$, (2.16) implies (2.1). This shows that the rate region in Theorem 8 is an outer bound on the secure capacity region when $k = 1$.

We now consider an example of a two-layer network and show how the upper bound derived above applies to it.

Example 4: Let $\mathcal{A} = \{2, 3, 4\}$, and assume that $\mathcal{M}_1 = \{1, 2\}$, $\mathcal{M}_2 = \{3, 4\}$, $\mathcal{M}_3 = \{4, 5, 6\}$ and $\mathcal{M}_4 = \{7, 8\}$. Then, we construct an undirected graph such that: (i) it has 3 nodes since $|\mathcal{A}| = 3$ and (ii) it has an edge between node 2 and node 3 since $\mathcal{M}_2 \cap \mathcal{M}_3 = \{4\} \neq \emptyset$. It therefore follows that this graph has $C_{\mathcal{A}} = 2$ components. In particular, we have

$$\sum_{i \in \mathcal{A}} R_i = \sum_{i \in \mathcal{V}_1} R_i + \sum_{i \in \mathcal{V}_2} R_i \leq M_{\{2,3,4\}} - 2k \overset{k=1}{=} 4, \tag{2.17}$$

where $\mathcal{V}_1 = \{2, 3\}$ and $\mathcal{V}_2 = \{4\}$. ■

2.5.3.2 Achievable rate region

We here show that the rate region in Theorem 8 is achieved by the scheme described in Section 2.5.1. In particular, we show that

$$\dim\left(\sum_{i \in \mathcal{A}} N_i\right) \geq M_{\mathcal{A}} - C_{\mathcal{A}}, \ \forall \mathcal{A} \subseteq [m], \tag{2.18}$$

where recall that $\dim\left(\sum_{i \in \mathcal{A}} N_i\right)$ is the secure rate performance of our proposed scheme in Section 2.5.1 (see Lemma 7). The condition in (2.18) can be equivalently written as $\forall \mathcal{A} \subseteq [m]$,

$$\begin{aligned} M_{\mathcal{A}} - C_{\mathcal{A}} \leq \dim\left(\textstyle\sum_{i \in \mathcal{A}} N_i\right) &\overset{\text{(a)}}{=} \dim\left(\left(\cap_{i \in \mathcal{A}} V_i\right)^{\perp}\right) \\ &= t - \dim\left(\cap_{i \in \mathcal{A}} V_i\right), \end{aligned}$$

where the equality in (a) follows by using the property of the dual space and rank nullity theorem, and $V_i, i \in \mathcal{A}$ is defined in (2.11). In other words, we next show that

$$\forall \mathcal{A} \subseteq [m], \ \dim\left(\cap_{i \in \mathcal{A}} V_i\right) \leq t - M_{\mathcal{A}} + C_{\mathcal{A}}. \tag{2.19}$$

Towards this end, we would like to count the number of linearly independent vectors $x \in \mathbb{F}_q^t$ that belong to $(\cap_{i \in \mathcal{A}} V_i)$.

We note that, by our construction: (i) V^T consists of one row (since $k = 1$) of t ones, and (ii) C_i has zeros in the positions indexed by $\mathcal{M}_i$. Hence, if a vector belongs to V_i, then all its components indexed by $\mathcal{M}_i$ have to be the same, i.e., either they are all zeros, or they are all equal to a multiple of one. Thus, we have q choices to fill such positions indexed by $\mathcal{M}_i$.

Now, consider V_j with $j \in \mathcal{A}$ and $j \neq i$. By using the same logic as above, if a vector belongs to V_j, then all its components indexed by $\mathcal{M}_j$ have to be the same and we have q choices to fill these. We now need to count the number of such choices that are consistent with the choices made to fill the positions indexed by $\mathcal{M}_i$. Towards this end, we consider two cases:

1. **C**ase 1: $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$. In this case, there is no overlap in the elements indexed by $\mathcal{M}_i$ and $\mathcal{M}_j$ and hence we can select all the available q choices for the positions indexed by $\mathcal{M}_j$;

2. **C**ase 2: $\mathcal{M}_i \cap \mathcal{M}_j \neq \emptyset$. In this case, there is some overlap in the elements indexed by $\mathcal{M}_i$ and $\mathcal{M}_j$. Thus, since we have already fixed the elements indexed by $\mathcal{M}_i$, we do not have any choice for the elements indexed by $\mathcal{M}_j$ (since all the elements have to be the same).

By iterating the same reasoning as above for all $i \in \mathcal{A}$, we conclude that we can fill all the positions indexed by $\cup_{i\in\mathcal{A}}\mathcal{M}_i$ of a vector $x \in \mathbb{F}_q^t$ and make sure that $x \in (\cap_{i\in\mathcal{A}}V_i)$ in $q^{C_\mathcal{A}}$ ways. This is because, there are $C_\mathcal{A}$ connected components, and for each of these components we have only q choices to fill the corresponding positions in the vector x (i.e., the positions that correspond to the relays to which at least one of the destinations inside that component is connected). Once we fix any position inside a component, in fact all the other positions inside that component have to be the same, and thus we have no more freedom in choosing the other positions. Moreover, the remaining $t - M_\mathcal{A}$ positions of x can be filled with any value in $\mathbb{F}_q$ and for this we have $q^{t-M_\mathcal{A}}$ possible choices. Therefore, the number of vectors $x \in \mathbb{F}_q^t$ that belong to $(\cap_{i\in\mathcal{A}}V_i)$ is at most $q^{C_\mathcal{A}+t-M_\mathcal{A}}$, which implies

$$\forall \mathcal{A} \subseteq [m],\ \dim\left(\cap_{i\in\mathcal{A}}V_i\right) \leq t - M_\mathcal{A} + C_\mathcal{A}.$$

This proves that the secure scheme in Section 2.5.1 achieves the rate region in Theorem 8. We now illustrate our method of identifying vectors that belong to $\cap_{i\in\mathcal{A}}V_i$ through an example.

Example 5: Let $t = 8$, $m = 4$, $\mathcal{M}_1 = \{1, 2\}$, $\mathcal{M}_2 = \{3, 4\}$, $\mathcal{M}_3 = \{4, 5, 6\}$ and $\mathcal{M}_4 = \{7, 8\}$. Let $\mathcal{A} = \{2, 3, 4\}$. With this, we can construct $V_i, i \in [4]$, as described in (2.11), where V^T consists of one row of 8 ones. We now want to count the number of vectors $x \in \mathbb{F}_q^8$ such that $x \in V_2 \cap V_3 \cap V_4$. We use the following iterative procedure:

1. For x to belong to V_2 its elements in the 3rd and 4th positions have to be the same since $\mathcal{M}_2 = \{3, 4\}$. Thus, we have q choices to fill the 3rd and 4th position.

2. For x to belong to V_3, its elements in the 4th, 5th and 6th positions have to be the same since $\mathcal{M}_3 = \{4, 5, 6\}$. However, the element in the 4th position has already been fixed in selecting vectors that belong to V_2. Thus, there is no further choice in filling the 5th and 6th positions.

3. For x to belong to V_4, its elements in the 7th and 8th positions have to be the same since $\mathcal{M}_4 = \{7, 8\}$. Since in the previous two steps, we have not filled yet the elements in these positions, then we have q possible ways to fill the elements in the 7th and 8th positions.

4. Moreover, we can fill the elements in the 1st and 2nd positions of x in q^2 possible ways.

With the above procedure we get that $\dim\left(\cap_{i\in\{2,3,4\}} V_i\right) = 4$, which is equal to the upper bound that we computed in (2.17) for the same example. ■

2.5.4 Secure capacity for $m = 3$, k arbitrary

In this section, we consider the case $m = 3$, and prove the following theorem.

Theorem 9. *For a two-layer network with $m = 3$ destinations, the secure capacity region is given by*

$$\sum_{i\in\mathcal{A}} R_i \leq M_{\mathcal{A}} - k, \ \forall \mathcal{A} \subseteq [m]. \tag{2.20}$$

Clearly, from our result in Theorem 2, the rate region in (2.20) is an outer bound on the secure capacity region and can be equivalently written as

$$\sum_{i\in\mathcal{A}} R_i \leq \min_{\mathcal{P} \,:\, \bigsqcup_{\mathcal{Q}\in\mathcal{P}} \mathcal{Q} = \mathcal{A}} \left\{ \sum_{\mathcal{Q}\in\mathcal{P}} M_{\mathcal{Q}} - |\mathcal{P}|k \right\}, \ \forall \mathcal{A} \subseteq [m],$$

where $\mathcal{P}$ is a disjoint partition of $\mathcal{A}$. We will now show that for every $\mathcal{A} \subseteq [m]$,

$$\dim\left(\sum_{i\in\mathcal{A}} N_i\right) \geq \min_{\mathcal{P}\,:\,\bigsqcup\limits_{\mathcal{Q}\in\mathcal{P}} \mathcal{Q} = \mathcal{A}} \left\{\sum_{\mathcal{Q}\in\mathcal{P}} M_{\mathcal{Q}} - |\mathcal{P}|k\right\}. \tag{2.21}$$

We prove (2.21) by considering three different cases.

Case 1: $|\mathcal{A}| = 1$**, i.e.,** $\mathcal{A} = \{i\}$. For this case, V_i in (2.11) has $k + t - M_{\{i\}}$ rows. In particular, all these rows are linearly independent since: (i) the rows of V^T are linearly independent as V is a Vandermonde matrix, (ii) C_i is full row rank by construction, and (iii) any linear combination of the rows of V^T will have a weight of at least $t - k + 1$ (from the Vandermonde property), whereas any linear combination of the rows of C_i will have a weight of at most $t - M_{\{i\}} \leq t - k$. It therefore follows that, $\forall i \in [3]$, we have that

$$\dim(N_i) = t - \dim(V_i) = t - (k + t - M_{\{i\}}) = M_{\{i\}} - k,$$

where the first equality follows by using the rank-nullity theorem. Thus, (2.21) is satisfied.

Case 2: $|\mathcal{A}| = 2$**, i.e.,** $\mathcal{A} = \{i, j\}$. For this case, $\forall (i, j) \in [3]^2, i \neq j$, we have that

$$\begin{aligned}\dim(N_i + N_j) &= \dim(N_i) + \dim(N_j) - \dim(N_i \cap N_j) \\ &= M_{\{i\}} + M_{\{j\}} - 2k - \dim(N_i \cap N_j),\end{aligned} \tag{2.22}$$

where the second equality follows by using $\dim(N_i)$ derived in Case 1. Thus, we need to compute $\dim(N_i \cap N_j)$. Note that, by definition, $N_i \cap N_j$ is the right null space of

$$V_{ij}^{\star} = \begin{bmatrix} V_i \\ V_j \end{bmatrix} \overset{(2.11)}{=} \begin{bmatrix} V^T \\ C_i \\ C_j \end{bmatrix} = \begin{bmatrix} V^T \\ C_{ij} \end{bmatrix},$$

where the last equality follows by removing one copy of the common rows in C_i and C_j, i.e., C_{ij} is a matrix of dimension $(t - M_{\cap\{i,j\}}) \times t$, with all unique rows. Using a similar argument as in Case 1 (i.e., any vector in the span of V^T has a minimum weight of $t - k + 1$ and any linear combination of the rows of C_{ij} will have a weight of at most $t - M_{\cap\{i,j\}}$), the number of linearly independent rows of $V_{ij}^{\star}$ is $\min\{t, t - M_{\cap\{i,j\}} + k\}$. Thus,

$$\dim(N_i \cap N_j) = t - \min\{t, t - M_{\cap\{i,j\}} + k\}$$

$$= \max\{0, M_{\cap\{i,j\}} - k\} = [M_{\cap\{i,j\}} - k]^+,$$

where the first equality follows from the rank-nullity theorem. We can now write $\dim(N_i + N_j)$ from (2.22) as

$$\dim(N_i + N_j) = \min\left\{M_{\{i\}} + M_{\{i\}} - 2k, M_{\{i,j\}} - k\right\}.$$

Thus, the condition in (2.21) is satisfied.

Case 3: $\mathcal{A} = \{1, 2, 3\}$. For this case, we will compute $\dim(N_1 + N_2 + N_3)$ as

$$\dim(N_1 + N_2 + N_3) = t - \dim(V_1 \cap V_2 \cap V_3). \tag{2.23}$$

Towards this end, we would like to compute the number of linearly independent vectors $x \in \mathbb{F}_q^t$ that belong to $V_1 \cap V_2 \cap V_3$. We start by noting that, similar to the case $k = 1$, we have $t - M_{\{1,2,3\}}$ degrees of freedom to fill the positions of x corresponding to $[t] \setminus \cup_{i\in[3]}\mathcal{M}_i$. We now select a permutation (i, j, ℓ) of $(1, 2, 3)$. In order for x to belong to V_i, the positions of x corresponding to $\mathcal{M}_i$ can be filled with k degrees of freedom. This is because: (i) C_i in (2.11) has zeros in the positions specified by $\mathcal{M}_i$, and (ii) V^T has k rows. Then, to fill the positions of x specified by $\mathcal{M}_j$ so that $x \in V_j$, we have at most $[k - M_{\cap\{i,j\}}]^+$ degrees of freedom. This is because the positions of x corresponding to $\mathcal{M}_i \cap \mathcal{M}_j$ have already been fixed (when filling the positions of x specified by $\mathcal{M}_i$). Finally, to fill the positions of x corresponding to $\mathcal{M}_\ell$ such that $x \in V_\ell$, we have at most $[k - M_{\cap\{\ell,\{i,j\}\}}]^+$ degrees of freedom. This is because, the positions of x corresponding to $\mathcal{M}_\ell \cap (\mathcal{M}_i \cup \mathcal{M}_j)$ are already fixed. Thus, we obtain

$$\begin{aligned}\dim(V_1 \cap V_2 \cap V_3) \leq k &+ [k - M_{\cap\{i,j\}}]^+ \\ &+ [k - M_{\cap\{\ell,\{i,j\}\}}]^+ + t - M_{\{1,2,3\}}.\end{aligned}$$

In Appendix A.5, we further tighten the bound for the quantity $\dim(V_1 \cap V_2 \cap V_3)$ and show that, when substituted in (2.23), it satisfies the condition in (2.21). This proves that the scheme described in Section 2.5.1 securely achieves the rate region in Theorem 9.

2.5.5 Secure capacity for arbitrary values of k and m

We here provide sufficient conditions for which the secure scheme in Section 2.5.1 is capacity achieving for arbitrary values of k and m. In particular, we have the following lemma.

Lemma 10. *The scheme in Section 2.5.1 achieves the secure capacity region of a two-layer network with arbitrary values of k and m whenever $\mathcal{M}_{\cap\{i,j\}} \geq k$ for all $(i,j) \in [m]^2, i \neq j$.*

Proof. We can compute $\dim(\cap_{i\in\mathcal{A}} V_i)$ as follows

$$\begin{aligned}\dim(\cap_{i\in\mathcal{A}} V_i) &\overset{\text{(a)}}{\leq} k + [k - \mathcal{M}_{\cap\{i_1,i_2\}}]^+ \\ &\quad + [k - \mathcal{M}_{\cap\{i_3,\{i_1,i_2\}\}}]^+ + \dots \\ &\quad + [k - \mathcal{M}_{\cap\{i_{|\mathcal{A}|},\{i_1,i_2,\dots,i_{|\mathcal{A}|-1}\}\}}]^+ + t - M_{\mathcal{A}} \\ &\overset{\text{(b)}}{=} k + t - M_{\mathcal{A}},\end{aligned}$$

where $(i_1, i_2, \dots, i_{|\mathcal{A}|})$ represents a permutation of the elements of $\mathcal{A}$ and: (i) the inequality in (a) follows by extending to arbitrary values of m the iterative algorithm proposed for Case 3 in Section 2.5.4 to fill the vector x so that $x \in \cap_{i\in\mathcal{A}} V_i$ and (ii) the equality in (b) follows since

$$\mathcal{M}_{\cap\{i_j,\{i_1,i_2,\dots,i_{j-1}\}\}} \geq \mathcal{M}_{\cap\{i_j,i_{j-1}\}} \geq k.$$

By using the property of dual spaces and the rank-nullity theorem, we obtain $\dim(\sum_{i\in\mathcal{A}} N_i) \geq M_{\mathcal{A}} - k$, which satisfies the condition in (2.21) $\forall \mathcal{A} \subseteq [m]$. This concludes the proof of Lemma 10. □

Example 6: An example of a two-layer network that satisfies the condition in Lemma 10 is characterized by the following parameters (see Definition 4): $t = 10, m = 4, k = 6$,

$$\begin{aligned}\mathcal{M}_1 &= \{1,2,3,4,5,6,7,8\} \\ \mathcal{M}_2 &= \{3,4,5,6,7,8,9,10\} \\ \mathcal{M}_3 &= \{1,2,5,6,7,8,9,10\} \\ \mathcal{M}_4 &= \{1,2,3,4,7,8,9,10\}.\end{aligned}$$

■

The results presented in this section provide the secure capacity region characterization for networks with: (i) arbitrary value m of destinations, and $k = 1$ edge eavesdropped by the adversary; (ii) arbitrary value k of edges eavesdropped and $m = 3$ destinations; (iii) arbitrary values for k and m under certain conditions on the min-cut capacities (see Lemma 10).

For arbitrary values of m and k for which the condition in Lemma 10 is not satisfied, we performed numerical evaluations by randomly constructing two-layer networks and, for all the cases we tried, we could not find any network for which the scheme is not optimal. In particular, in our simulations, we considered up to $m = 8$ destinations and, for different choices of t and k, we connected each destination to a randomly chosen set of relays. We constructed 100 such network instances, and verified that the rate region achieved by our designed scheme given in Lemma 7 equals the outer bound in (2.1). This suggests that our designed scheme could indeed be optimal for arbitrary values of m and k, and we conjecture this result to hold.

Conjecture 1. *Consider a two-layer network with m destinations, where an adversary eavesdrops any k edges of her choice. The secure capacity region is given by*

$$\sum_{i\in\mathcal{A}} R_i \leq |\cup_{i\in\mathcal{A}}\mathcal{M}_i| - k,\ \forall \mathcal{A} \subseteq [m],$$

where $\mathcal{M}_i \subseteq [t],\ i \in [m]$ denotes the destination connection sets.

2.5.6 Secure capacity scheme for arbitrary separable networks

In this section, we will first show that for any separable network, a corresponding two-layer network can be created such that both networks have the same min-cut capacities $M_{\mathcal{A}}$ for all $\mathcal{A} \subseteq [m]$. We will then show that a secure scheme designed for a two-layer network can be converted to a secure scheme for the corresponding separable network.

By Definition 3, a separable network $\mathcal{G}$ with m destinations, can be separated into $2^m - 1$ networks $\mathcal{G}'_{\mathcal{J}},\ \mathcal{J} \subseteq [m], \mathcal{J} \neq \emptyset$ where $\mathcal{G}'_{\mathcal{J}}$ has min-cut capacity $M'_{\mathcal{J}}$ to every subset of destinations in $\mathcal{J}$. To construct the corresponding two-layer network, we use the following iterative procedure:

1. We place the source node S in layer 0 of our network, and the m destination nodes $D_i, i \in [m]$, in layer 2 of our network;

2. For each $\mathcal{J} \subseteq [m]$, we add $M'_{\mathcal{J}}$ relays in layer 1 of our network;

3. For each $\mathcal{J} \subseteq [m]$, we connect: (i) the source in layer 0 with all the added $M'_{\mathcal{J}}$ relays, and (ii) all the added $M'_{\mathcal{J}}$ relays with the destinations $D_i, i \in \mathcal{J}$ in layer 2.

By following the above procedure, it is not difficult to verify that, for each $\mathcal{A} \subseteq [m]$, the min-cut capacity in the constructed two-layer network is $M_{\mathcal{A}}$ as given in (2.3). As such, the new constructed two-layer network has the same min-cut capacity $M_{\mathcal{A}}$ of the corresponding separable network. In what follows, we refer to the original separable network as *parent* separable network, and to the corresponding two-layer network as *child* two-layer network.

We now show that a secure scheme designed for the child two-layer network can be leveraged to build a secure scheme for the corresponding parent separable network. Towards this end, we assume that we have a secure scheme for the child two-layer network, i.e., as described in (2.10) in Section 2.5.1, we have

$$X = \begin{bmatrix} M & V \end{bmatrix} \begin{bmatrix} W \\ K \end{bmatrix}.$$

Recall that, as highlighted in Remark 1, even if we select a random matrix $\tilde{V}$ instead of the Vandermonde matrix V, with a high probability (close to 1 for large field size) we will have a secure scheme for the child two-layer network.

To transform the above secure scheme into a secure scheme for the parent separable network, we proceed as follows. On every graph $\mathcal{G}'_{\mathcal{J}}$ in the parent separable network, we transmit (multicast) the symbols that were transmitted in the child two-layer network from the source node S in layer 0 to the set of $M'_{\mathcal{J}}$ relays in layer 1 that were added when constructing the child two-layer network for $\mathcal{G}'_{\mathcal{J}}$. Note that this multicast towards all destinations $D_i, i \in \mathcal{J}$, is possible since $\mathcal{G}'_J$ has min-cut capacity $M'_{\mathcal{J}}$. With such a strategy, at the end of the transmissions every destination in the parent separable graph still receives the same set of packets as it would have received in the child two-layer network. Thus, all the destinations can still decode their respective messages. We now prove that this scheme is also secure. Let Y be the collection of the symbols transmitted (multicast) on the parent separable network, as described above. Since multicasting involves network coding, we

have

$$Y = \left[\begin{array}{c} G \end{array} \right] X, \tag{2.24}$$

where G is an encoding matrix of dimension $|\mathcal{E}| \times M_{[m]}$, which can be constructed in $\mathcal{O}\left(m|\mathcal{E}|^3\right)$ by using the multicasting scheme of [JSC05], which requires a finite field of dimension m. Thus,

$$\begin{aligned} Y &= \left[\begin{array}{c} G \end{array} \right] \left[\begin{array}{cc} M & \tilde{V} \end{array} \right] \left[\begin{array}{c} W \\ K \end{array} \right] \\ &= \left[\begin{array}{cc} GM & G\tilde{V} \end{array} \right] \left[\begin{array}{c} W \\ K \end{array} \right]. \end{aligned}$$

From the security condition in Lemma 1, it follows that the scheme above is secure if we can show that for any choice of G, there exists a $\tilde{V}$ such that $\tilde{V}$ is an MDS matrix (i.e., any k rows of $\tilde{V}$ are linearly independent) and

$$rk\left(\left[\begin{array}{cc} GM & G\tilde{V} \end{array} \right]\Big|_{\mathcal{Z}} \right) = rk\left(\left[\begin{array}{c} G\tilde{V} \end{array} \right]\Big|_{\mathcal{Z}} \right), \ \forall |\mathcal{Z}| \leq k. \tag{2.25}$$

This is shown in Appendix A.6, where we prove that over a sufficiently large finite field, with high probability a random choice of $\tilde{V}$ is an MDS matrix and satisfies the condition in (2.25).

2.6 Two-Phase scheme for networks with arbitrary topologies and arbitrary number of destinations

We now propose the design of a secure transmission scheme for networks with arbitrary topologies and arbitrary number of destinations. This scheme consists of two phases, namely the key generation phase (in which secret keys are generated between the source and the m destinations) and the message sending phase (in which the message packets are first encoded using the secret keys and then transmitted to the m destinations). In particular, this scheme is inspired by the work in [CPF11], where it was shown that for multicast and single unicast connections, such a two-phase scheme that separates over time the transmissions of keys and messages indeed achieves the secure capacity. However, it turns out that this is no longer the case for multiple unicast sessions, as we discuss in detail in the following.

The achievable rate region of this two-phase scheme is presented in Theorem 11.

Theorem 11. *Let $(\hat{R}_1, \hat{R}_2, \ldots, \hat{R}_m)$ be an achievable rate m-tuple in the absence of the eavesdropper. Then, the rate m-tuple $(R_1, R_2, \ldots, R_m)$ with*

$$R_i = \hat{R}_i \left(1 - \frac{k}{M}\right), \forall i \in [m] \ , \tag{2.26}$$

where M is the minimum min-cut capacity between the source and any destination, is securely achievable in the presence of an adversary who eavesdrops any k edges of her choice.

Proof. Let $M_{\{i\}}$ be the min-cut capacity between the source and the destination D_i with $i \in [m]$. We define M as the minimum among all these individual min-cut capacities, i.e., $M = \min_{i\in[m]} M_{\{i\}}$. Let $(\hat{R}_1, \hat{R}_2, \ldots, \hat{R}_m) \in \mathbb{R}^m$ be the rate m-tuple achieved in the absence of the eavesdropper. We start by approximating this rate m-tuple with rational numbers; notice that this is always possible since the set of rationals $\mathbb{Q}$ is dense in $\mathbb{R}$. Since this rational rate m-tuple might involve fractional flows on the edges, we replace each edge with T parallel edges. This number T is chosen such that: (i) we achieve the rate m-tuple $(T\hat{R}_1, T\hat{R}_2, \ldots, T\hat{R}_m)$ over T network uses of the original network, and (ii) every edge carries an integer flow. We denote this new network as $\mathcal{G}_T$. In what follows, we describe our coding scheme and show that

$$(R_1, R_2, \ldots, R_m) = \left(1 - \frac{k}{M}\right)(\hat{R}_1, \hat{R}_2, \ldots, \hat{R}_m) \tag{2.27}$$

is securely achievable. In particular, our scheme consists of two phases, and in each phase we use the network $\mathcal{G}_T$. We also highlight that we allow the adversary to eavesdrop any Tk edges of $\mathcal{G}_T$. In other words, at the i-th network use of the original network, with $i \in [T]$, we allow the adversary to eavesdrop a set of k edges that might be different from those eavesdropped in the previous network uses. We next describe the two phases of our scheme.

- *K*ey generation. This first phase – in which secure keys are generated between the source and the destinations – consists of k subphases. In each subphase, the source multicasts $T(M-k)$ random packets securely to all destinations. This is possible thanks to the secure network coding result of [CY02], since the minimum min-cut capacity of $\mathcal{G}_T$ is TM and Eve has access to Tk edges. Thus, at the end of this phase, a total of $Tk(M-k)$ secure keys are generated and exchanged between the source and the m destinations.

- *M*essage sending. This phase consists of $M-k$ subphases. We choose Tk packets out of the $Tk(M-k)$ securely shared (in the key generation phase) random packets. For each choice of Tk packets, we convert the unsecure scheme achieving $(T\hat{R}_1, T\hat{R}_2, \ldots, T\hat{R}_m)$ to a secure scheme achieving the same rate m-tuple. Towards this end, we expand the Tk shared packets into $\sum_{j=1}^{m} T\hat{R}_j$ packets using an MDS code matrix. With this, we have the same number of random packets as the message packets. We then add the message packets with the random packets and transmit them as it was done in the corresponding unsecure scheme (i.e., in absence of the eavesdropper).

Proof of security. For each of the $M-k$ subphases of the message sending phase, we denote by W_i the $T\hat{R}_i$ messages for D_i, and define W as

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \cdots \\ W_m \end{bmatrix}$$

Moreover, we let K be the vector containing the Tk securely shared random packets. With this, for each of the $M-k$ subphases of the message sending phase, we can write the transmissions over the network $\mathcal{G}_T$ as

$$X = \begin{bmatrix} M_{us} & V \end{bmatrix} \begin{bmatrix} W \\ K \end{bmatrix}, \tag{2.28}$$

where M_{us} is the encoding matrix used in absence of the eavesdropper and V is the Vandermonde matrix of size $\sum_{i=1}^{m} T\hat{R}_i \times Tk$. Because of the property of Vandermonde matrices, (2.28) satisfies the security condition in Lemma 1, and hence the scheme above is secure.

Analysis of the achieved rate m-tuple. The secure scheme described above requires a total of M subphases, where the first k subphases are from phase 1 (i.e., key generation) and the next $M-k$ subphases are from phase 2 (i.e., message sending). In particular, in the first k subphases, we generate the secure keys and in the remaining $M-k$ subphases, we securely transmit at rates of $(T\hat{R}_1, T\hat{R}_2, \ldots, T\hat{R}_m)$. Thus, the achieved secure message rate $(R_1, R_2, \ldots, R_m)$ is

$$R_j = \frac{M-k}{M}\hat{R}_j = \left(1 - \frac{k}{M}\right)\hat{R}_j, \forall j \in [m] \ . \tag{2.29}$$

This concludes the proof of Theorem 11. □

It is worth noting that the capacity region in absence of the eavesdropper was determined in [KM03, Theorem 9], and is given by

$$\sum_{i\in\mathcal{A}} R_i \leq M_{\mathcal{A}},\ \forall \mathcal{A} \subseteq [m].$$

By leveraging this result and (2.26), we can therefore compute the rate region achieved by our secure two-phase scheme, which is given in the next corollary.

Corollary 12. *The achievable secure rate region of the two-phase scheme is given by*

$$\sum_{i\in\mathcal{A}} R_i \leq M_{\mathcal{A}} - k\left(\frac{M_{\mathcal{A}}}{M}\right),\ \forall \mathcal{A} \subseteq [m],$$

where $M = \min_{i\in[m]} M_{\{i\}}$.

We now comment on the design complexity of this two-phase scheme and provide a trivial upper bound on the field size.

Lemma 13. *The complexity of designing the secure two-phase transmission scheme equals* $\mathcal{O}(m^3|\mathcal{E}|^3)$. *Moreover, a field size of* $\mathcal{O}(m + |\mathcal{E}|)$ *suffices.*

Proof. To design our two-phase scheme for a rate tuple $(R_1, R_2, \ldots, R_m)$, we need to use k sub-phases for multicasting the keys and $M - k$ sub-phases for routing the multi-commodity information flow. Recall that M is the minimum of the min-cut capacities from the source to any destination $D_i,\ i \in [m]$, and as such M can be found in $\mathcal{O}(m|\mathcal{E}|^{2.5})$ [Vai89] by solving m linear programs. The design of the deterministic matrix for multicasting the keys has a time complexity of $\mathcal{O}(|\mathcal{E}|mM(M+m))$ [JSC05], which can be further upper bounded as $\mathcal{O}(|\mathcal{E}|^3 m)$. Finally, the design of the routing for multi-commodity flow can also be performed using a linear program which has a time complexity of $\mathcal{O}((m|\mathcal{E}|)^{2.5})$ [Vai89]. Thus, the overall time complexity is $\mathcal{O}(m^3|\mathcal{E}|^3)$.

The field size required for constructing the deterministic multicast matrix is m [JSC05]. We also require an MDS code for encoding the message packets before routing. This requires a field size of $|\mathcal{E}|$ (corresponding to the Vandermonde or similar MDS matrix). Thus, a trivial bound on

the field size requirement is $\mathcal{O}(m + |\mathcal{E}|)$. However, since our encoding schemes are linear, we believe that vector encoding schemes, such as the subspace coding scheme in [KSK09], could be adapted to this case and leveraged to achieve a small finite field size. This is part of our current investigation. □

We conclude this section, by highlighting two fundamental features of our two-phase scheme:

1. Different from the scheme designed in Section 2.5, which only applies to separable networks, the two-phase scheme applies to networks with *arbitrary* topologies.

2. The two-phase scheme is oblivious to the network structure, and uses all the network resources in both phases. In other words, different from the optimal scheme of Section 2.4 for $m = 2$ destinations, the two-phase scheme does not seek to optimally separate the information and key flows. This causes the scheme to be sub-optimal (see also Corollary 12) as also remarked by the detailed analysis in [ACF17, Section 4.3].

CHAPTER 3

Millimeter Wave Networks (1-2-1 Networks)

High-frequency communication using millimeter waves employs beamforming to compensate for the high path-loss and the high blockage. Although beamforming restricts the communication to be in one direction, the additional degree of freedom in choosing the direction can make the secure capacity arbitrary close to the unsecure capacity. We show this by designing an encryption scheme that uses a time-varying selection of the network and does require any pre-shared key.

3.1 Summary

In this chapter, we consider a 1-2-1 network model and study its secure capacity. The 1-2-1 model is the abstracted noiseless model of millimeter wave networks and has been shown to approximate the Gaussian capacity by Ezzeldin et al. [ECF18]. In this model, any two connected nodes communicate by aligning their transmitting and receiving antenna known as beamforming.

Unlike the wireline network model, a node receiving different symbols on each of its incoming edges and similarly transmitting different symbols on the outgoing links, is not possible in 1-2-1 network model due to the beamforming. Any node in the 1-2-1 network model can receive only on one incoming edge and can also transmit only on one outgoing link. Our main observation is that the choice in picking the direction to communicate helps in securing against an eavesdropper without requiring any pre-shared key.

In this chapter, we characterize the secure capacity of unicast traffic against an eavesdropper, using a time-varying selection of the network. We also derive sharp lower and upper bounds when the source and the destination are assumed to have more than one antenna, and therefore, simultaneously transmitting and receiving on multiple edges. Finally, we characterize the secure capacity

for a particular network topology called *diamond networks* where the edges can have arbitrary edge capacities.

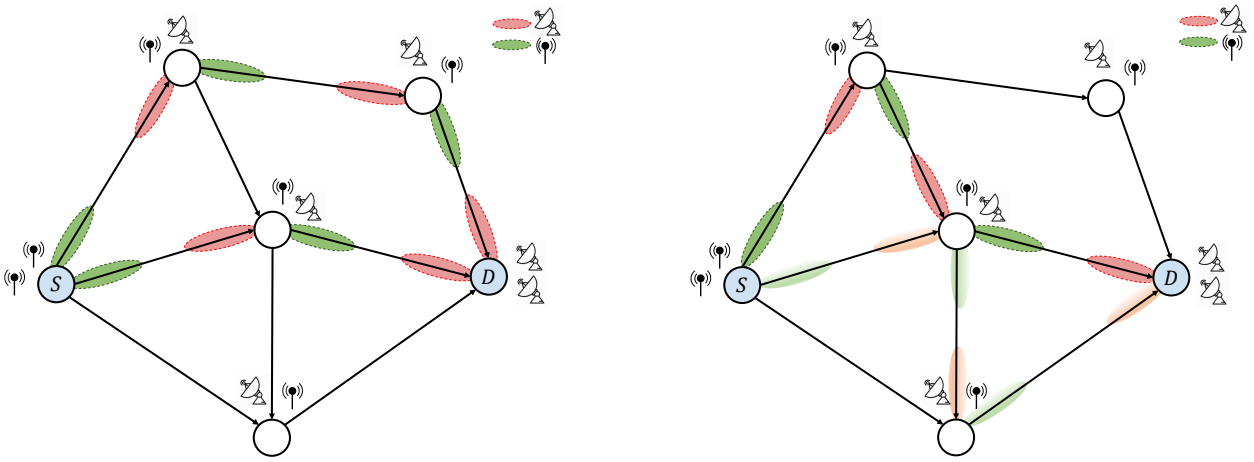
Organization: Section 3.2 formally describes the 1-2-1 network model and characterizes the unsecure capacity for arbitrary network topology with unit edge capacities. Section 3.3 characterizes the secure capacity results for arbitrary networks with unit edge capacities and provides sharp bound when the source and the destination have multiple antennas. Section 3.4 presents the secure capacity result for *diamond networks* with arbitrary edge capacities.

3.2 System model and unsecure capacity

The work in [ECF18] examined the unsecure capacity characterization of the Gaussian mmWave network by modeling it as a *1-2-1 network*. In [ECF18], it was shown that the capacity of a Gaussian 1-2-1 network can be approximately characterized to within a constant gap by a lossless 1-2-1 network where the schedule does not depend on the transmitted messages in the network. In this chapter, we examine security over such 1-2-1 networks, that we describe next.

We assume that each link in the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with edges of fixed finite capacities, can be activated according to the 1-2-1 constraints. That is, at any particular time, an intermediate node can simultaneously receive and transmit but it can at most listen to one node (one incoming edge) and direct its transmission to one node (one outgoing edge) in the network. For the source (respectively, destination) we allow it to transmit (respectively, receive from) m other nodes i.e., on m outgoing edges (respectively, on m incoming edges), simultaneously with no interference. An example with $m = 2$ antennas at the source and the destination is shown in Fig. 3.1. The Fig. 3.1a depicts a valid network use whereas it is not possible to use the network as shown in Fig. 3.1b since the intermediate node has to receive from two nodes and also has to transmit to two nodes at a time.

Adversary model and security: We assume that the source wishes to communicate a message W of entropy rate R securely from a passive external adversary Eve who can wiretap any k edges of her choice. If Eve wiretaps edges in the set $\mathcal{S} \subseteq \mathcal{E},\ |\mathcal{S}| = k$, and the symbols transmitted on these



(a) Valid use. (b) Invalid use.

Figure 3.1: An example of network with 1-2-1 constraints.

edges over n network uses are denoted by $\{T_e^n,\ e \in \mathcal{S}\}$, then we require that:

$$I(W; \{T_e^n,\ e \in S\}) = 0,\ \forall \mathcal{S} \subseteq \mathcal{E},\ |\mathcal{S}| = k. \tag{3.1}$$

We are interested in characterizing the secure message capacity C, that is the maximum rate at which the source can communicate with the destination with zero error under (3.1).

Unsecure capacity: Here, we derive the capacity in the absence of the eavesdropper Eve. 1-2-1 networks with arbitrary edge capacities and $m = 1$, under Gaussian channel models were analyzed in [ECF18], where the main result is that over such networks, one can **approximately** (i.e., up to a gap that only depends on the number of nodes) achieve the capacity by routing information across paths; moreover, out of an exponential number of paths that potentially connect the source to the destination, capacity can be achieved by utilizing at most a linear number (in the number of nodes) of them. In this section, we derive an additional result, namely the exact capacity for any m when all the edges are of unit capacity.

Theorem 14. *For arbitrary 1-2-1 networks with unit capacity edges, the capacity in absence of any eavesdropper is given by,*

$$C_u = \min(m, H_v), \tag{3.2}$$

where H_v is the maximum number of vertex disjoint paths.

Proof. **Achievability:** Let $p_{[H_v]}$ be the H_v vertex disjoint paths. These paths are of fundamental importance under the 1-2-1 constraints. This is because intermediate nodes can transmit and receive from only one node each, and this ensures that multiple vertex disjoint paths (depending on the number of source and destination beams) can be simultaneously operated at each time. We pick $\min(m, H_v)$ such paths and use these for the transmission and thus achieve a rate of $\min(m, H_v)$.
Outer bound: Whenever there are direct edges from the source to the destination, we add a virtual node in between, so that a direct edge turns into a two-hop path. This does not change the transmission rate as if there was a transmission on the direct edge in $\mathcal{G}$, it can also be performed using the added virtual node with no extra resources. Thus, we can assume that there are no direct edges from the source to the destination.

Now, we consider the minimum vertex cut of the network, i.e., the minimum number of vertices (excluding the source and the destination), such that when we remove them there is no path from the source to the destination. This minimum number of vertices is equal to the maximum number of vertex disjoint paths, i.e., H_v. We denote these vertices as $V_1,\ V_2,\ , \ldots,\ V_{H_v}$. Each of these intermediate nodes can transmit only on one of its outgoing edges. We denote the symbols transmitted on the outgoing edges of these nodes over n network uses as $T^n_{V_{[H_v]}}$, where $T^n_{V_i}$ denotes the symbols transmitted by vertex V_i. We represent the symbols received by the destination as T^n_D. By Fano's inequality, we obtain

$$\begin{aligned} nR &\leq H(W) \\ &\overset{(a)}{=} H(W) - H(W|T^n_D) \\ &\overset{(b)}{\leq} H(W) - H(W|T^n_{V_{[H_v]}}) \\ &= I(W; T^n_{V_{[H_v]}}) \\ &\leq H(T^n_{V_{[H_v]}}) \\ &\overset{(c)}{\leq} nH_v, \end{aligned} \tag{3.3}$$

$$\begin{aligned} nR &\leq H(W) - H(W|T^n_D) \\ &= I(W; T^n_D) \leq H(T^n_D) \\ &\overset{(d)}{\leq} Hn, \end{aligned} \tag{3.4}$$

where (a) is due to the reliable decoding constraint; (b) follows from the 'conditioning does not increase the entropy' principle and since $V_{[H_v]}$ is a vertex cut and thus all the information going to the destination passes through these vertices (i.e., T_D^n is a deterministic function of $T_{V_{[H_v]}}^n$); (c) is because there are H_v symbols for every instance and there are n such instances; and (d) holds because the destination can receive only on m incoming edges from m nodes. □

3.3 Arbitrary networks with unit edge capacities

In this section, we prove lower and upper bounds on the secure capacity.

Theorem 15. *Consider an arbitrary 1-2-1 network with unit capacity edges.*

*(a) For $m = 1$: If H_e is the maximum number of **edge disjoint** paths connecting the source to the destination on the underlying graph, then the 1-2-1 secure capacity C can be lower bounded as follows:*

$$C \geq \left(1 - \frac{k}{H_e}\right). \tag{3.5}$$

*(b) For $m > 1$: If H_v is the maximum number of **vertex disjoint** paths connecting the source to the destination on the underlying graph, then the 1-2-1 secure capacity C can be lower bounded as follows:*

$$C \geq \min(m, H_v)\left(1 - \frac{k}{H_v}\right). \tag{3.6}$$

Proof. The main intuition behind the proof is that we can apply the optimal secure communication scheme we would have used on the underlying graph if we did not have the 1-2-1 constraints, and then use this scheme under the 1-2-1 constraints, as described in what follows.

(a) **For** $m = 1$: Let $p_{[H_e]}$ be the edge disjoint paths. We start by generating k uniform random packets and make H_e linear combinations of these using an MDS code matrix of size $k \times H_e$. We denote these packets as $X_{[H_e]}$, and refer to them as "keys". Any k of these combinations are mutually independent. Next, we take $H_e - k$ message packets, and add (i.e., encode) these with

the first $H_e - k$ random packets. In other words, after this coding operation we obtain

$$T_i = \begin{cases} W_i + X_i & \text{if } i \leq H_e - k, \\ X_i & \text{else,} \end{cases}$$

where $W_{[H_e-k]}$ are message packets.

We use the network H_e times, and in each instance we use one of the paths from $p_{[H_e]}$. Thus, we would be able to communicate all encoded symbols in H_e time instances. Moreover, the destination will be able to cancel out the keys and thereby decode $H_e - k$ messages, as there are k symbols $T_{[H_e-k+1:T_{H_e}]}$, which are just independent combinations of the k random packets we started with.

Moreover, in each instance, Eve will receive a symbol if the edges she eavesdrops are part of the path that is used in that particular instance. Since her k edges can at most be part of k paths, Eve will receive at most k symbols, all of which are encoded with independent keys. Thus, the scheme securely transmits $H_e - k$ message packets in H_e uses of the network. Hence, we get a rate $R = \frac{H_e-k}{H_e} = 1 - \frac{k}{H_e}$, which is precisely the one in (3.5). Note that security follows from the security of the underlying scheme, that is a standard scheme for multipath security.

(b) **For $m > 1$:** Let $p_{[H_v]}$ be the vertex disjoint paths. Again, the fact that paths are vertex disjoint is crucial under the 1-2-1 constraints. This is because intermediate nodes can transmit and receive from only one node each, and this ensures that $\min(m, H_v)$ paths can be simultaneously operated at each time (note that having vertex disjoint paths is a sufficient but not a necessary condition).

Let $\hat{m} = \min(m, H_v)$. We start by generating $k\binom{H_v-1}{\hat{m}-1}$ random packets and extend them to $\hat{m}\binom{H_v}{\hat{m}}$ packets using an MDS code matrix. Then, similar to the case $m = 1$, we take the first $\hat{m}\binom{H_v}{\hat{m}} - K\binom{H_v-1}{\hat{m}-1}$ of these random packets and add (i.e., encode) them with the same amount of message packets. More formally, if $\left\{X_i, i \in \left[\hat{m}\binom{H_v}{\hat{m}}\right]\right\}$ are the random packets after the extension using the MDS code matrix, and $\left\{W_i, i \in \left[\hat{m}\binom{H_v}{\hat{m}} - k\binom{H_v-1}{\hat{m}-1}\right]\right\}$ are the message packets, then

$$T_i = \begin{cases} X_i + W_i & \text{if } i \leq \hat{m}\binom{H_v}{\hat{m}} - k\binom{H_v-1}{\hat{m}-1} \\ X_i & \text{else} \end{cases}.$$

We use this network $\binom{H_v}{\hat{m}}$ times, and in each instance we use a different choice of $\hat{m}$ paths to communicate. It is not difficult to see that each of the k edges eavesdropped by the adversary will intersect with $\binom{H_v-1}{\hat{m}-1}$ such network uses. This is because, for a fixed choice of edge, there are

$\binom{H_v-1}{\hat{m}-1}$ network instances where a symbol in carried via this edge. Hence, in total the adversary will receive only $k\binom{H_v-1}{\hat{m}-1}$ symbols, which are encoded with independent keys. The receiver, after the $\binom{H_v}{\hat{m}}$ network uses will be able to cancel out the keys. Thus, we can securely communicate $\hat{m}\binom{H_v}{\hat{m}} - k\binom{H_v-1}{\hat{m}-1}$ over $\binom{H_v}{\hat{m}}$ instances of the network, and achieve a rate R equal to

$$\begin{aligned} R &= \frac{\hat{m}\binom{H_v}{\hat{m}} - k\binom{H_v-1}{\hat{m}-1}}{\binom{H_v}{\hat{m}}} \\ &= \hat{m} - \frac{k\hat{m}}{H_v} \\ &= \min(m, H_v)\left(1 - \frac{k}{H_v}\right), \end{aligned}$$

which is precisely the one in (3.6). This concludes the proof of Theorem 15. □

Theorem 16. *Let H_e be the maximum number of **edge disjoint** paths connecting the source to the destination on the underlying directed graph, then the 1-2-1 secure capacity C can be upper bounded as follows:*

$$C \leq \min(m, H_e)\left(1 - \frac{k}{H_e}\right).$$

Proof. From the min-cut, max-flow theorem there are H_e edges such that, when removed, the source gets disconnected from the destination. Let $e_1,\ e_2,\ \ldots,\ e_{H_e}$ denote these edges. Assume that the network is used n times, and let $T^n_{e_i},\ i \in \{1, 2, \ldots, H_e\}$ be the symbols transmitted on these H_e edges over n uses of the network. By denoting the symbols transmitted by the source on n network instances by T^n_S, then,

$$\begin{aligned} nH &\geq H(T^n_S) \\ &\overset{(a)}{=} H(T^n_S, \{T^n_{e_i},\ i \in [H_e]\}) \\ &\geq H(\{T^n_{e_i},\ i \in [H_e]\}), \end{aligned}$$

where (a) follows because $\{T^n_{e_i},\ i \in [H_e]\}$ is a deterministic function of T^n_S. Moreover, $H(\{T^n_{e_i},\ i \in [H_e]\}) \leq nH_e$. Thus,

$$H(\{T^n_{e_i},\ i \in [H_e]\}) \leq \min(nH_e, nm). \tag{3.7}$$

In the remaining part of the proof, we use the result in the following lemma.

Lemma 17. *$\forall k, \ell \in \mathbb{Z},\ 0 \leq k \leq \ell$, $\exists\, \mathcal{S} \subset [\ell], |\mathcal{S}| = k$, such that*

$$H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\}) \leq \frac{\ell - k}{\ell} H(X_{[\ell]}).$$

Proof. Assume for all choices of $\mathcal{S} \subset [\ell],\ |\mathcal{S}| = k$,

$$H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\}) > \frac{\ell - k}{\ell} H(X_{[\ell]}).$$

Then, $\binom{\ell}{k} H(X_{[\ell]})$

$$\begin{aligned}
&\overset{(a)}{=} \sum_{\substack{\mathcal{S}\subset[\ell]\\|\mathcal{S}|=k}} \left(H(\{X_i,\ i \in \mathcal{S}\}) + H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\})\right)\\
&\overset{(b)}{\geq} \sum_{\substack{\mathcal{S}\subset[\ell]\\|\mathcal{S}|=k}} \left(\left(\sum_{i\in\mathcal{S}} H(X_i|\{X_j,\ j<i\})\right) + H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\})\right)\\
&\overset{(c)}{=} \binom{\ell-1}{k-1}\left(\sum_{i\in[\ell]} H(X_i|\{X_j,\ j<i\})\right) + \sum_{\substack{\mathcal{S}\subset[\ell]\\|\mathcal{S}|=k}} H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\})\\
&\overset{(d)}{=} \binom{\ell-1}{k-1} H(\{X_i,\ i \in [\ell]\}) + \sum_{\substack{\mathcal{S}\subset[\ell]\\|\mathcal{S}|=k}} H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\})\\
&\overset{(e)}{>} \binom{\ell-1}{k-1} H(\{X_i,\ i \in [\ell]\}) + \sum_{\substack{\mathcal{S}\subset[\ell]\\|\mathcal{S}|=k}} \frac{\ell-k}{\ell} H(\{X_i,\ i \in [\ell]\})\\
&= \binom{\ell-1}{k-1} H(\{X_i,\ i \in [\ell]\}) + \binom{\ell}{k}\frac{\ell-k}{\ell} H(\{X_i,\ i \in [\ell]\})\\
&= \binom{\ell}{k}\left(\frac{k}{\ell} H(\{X_i,\ i \in [\ell]\}) + \frac{\ell-k}{\ell} H(\{X_i,\ i \in [\ell]\})\right)\\
&= \binom{\ell}{k} H(\{X_i,\ i \in [\ell]\}),
\end{aligned}$$

and hence we get a contradiction. Here (a) is because there are $\binom{\ell}{k}$ ways of breaking $\{X_i,\ i \in [\ell]\}$ into two sets of size k and $\ell - k$, and then it follows from the chain rule of entropy; (b) follows because for any $\mathcal{S} \subset [\ell]$, we can order $\{X_i,\ i \in \mathcal{S}\}$ according to their index, and then we use the chain rule of entropy followed by the condition reduces entropy principle; (c) follows because for each $i \in [\ell]$, there will be $\binom{\ell-1}{k-1}$ choices of $\mathcal{S}$ where this i will be part of $\mathcal{S}$; (d) follows again from the chain rule of entropy; and (e) follows because of the assumption in the proof that for all choices of $\mathcal{S} \subset [\ell], |\mathcal{S}| = k$, $H(\{X_i,\ i \in \mathcal{S}^c\}|\{X_i,\ i \in \mathcal{S}\}) > \frac{\ell-k}{\ell} H(\{X_i,\ i \in [\ell]\})$. □

For $\ell = H_e$, we assume $\mathcal{S} = [k] \subset [H_e]$ in Lemma 17. Then, starting with Fano's inequality, we have

$$\begin{aligned}
nR &\leq H(W) = H(W) - H(W|\{T_{e_i}^n,\ i \in [H_e]\}) \\
&= I(W; \{T_{e_i}^n,\ i \in [H_e]\}) \\
&= I(W; \{T_{e_i}^n,\ i \in [k]\}) + \\
&\quad I(W; \{T_{e_i}^n,\ i \in [H_e] \setminus [k]\} | \{T_{e_i}^n,\ i \in [k]\}) \\
&\overset{(a)}{=} I(W; \{T_{e_i}^n,\ i \in [H_e] \setminus [k]\} | \{T_{e_i}^n,\ i \in [k]\}) \\
&\leq H(\{T_{e_i}^n,\ i \in [H_e] \setminus [k]\} | \{T_{e_i}^n,\ i \in [k]\}) \\
&\overset{(b)}{\leq} \frac{H_e - k}{H_e} \min(nH_e, nm) \\
&\implies R \leq \min(m, H_e) \left(1 - \frac{k}{H_e}\right),
\end{aligned}$$

where (a) follows since, for security, $I(W; \{T_{e_i}^n,\ i \in [k]\}) = 0$ and (b) is because of Lemma 17 and (3.7). This concludes the proof of Theorem 16. □

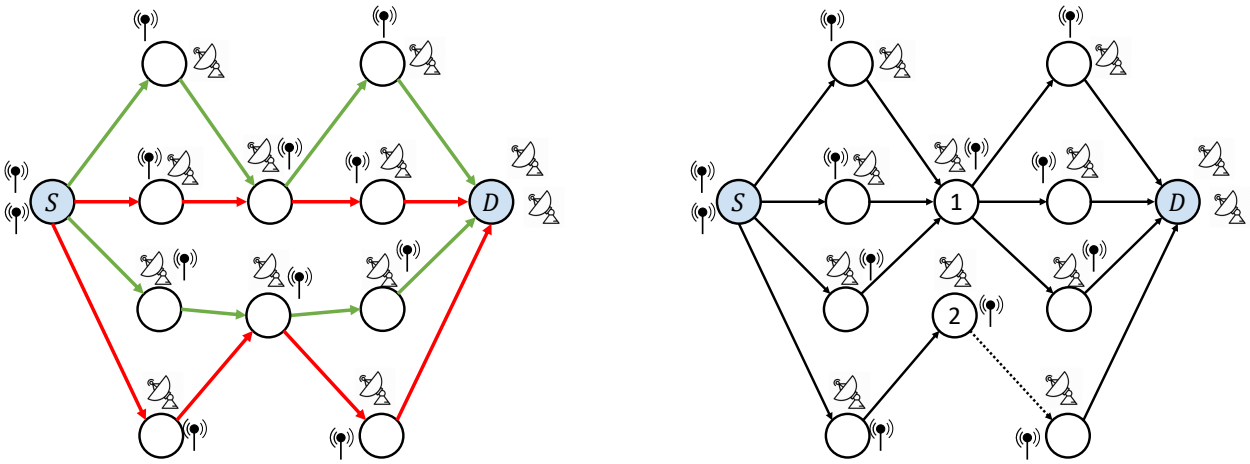
For some special cases, we can exactly characterize the capacity (i.e., the upper and lower bounds previously derived coincide). In particular, these include:

- Networks where the number of edge disjoint paths is equal to the number of vertex disjoint paths. For these networks, the capacity is given by $C = \min(m, H_e)(1 - \frac{k}{H_e})$.
- For networks where the source and the destination have one transmit and one receive beam each, i.e., $m = 1$. For these networks, the capacity is given by $C = 1 - \frac{k}{H_e}$.

We next provide two different network examples where:

- Example 7 where the upper bound is tight,
- Example 8 where the upper bound is not tight, but the lower bound is tight .

Example 7: In Fig. 3.2a, there are four edge disjoint paths from the source to the destination, i.e., $H_e = 4$. Assume that $m = 2$, i.e., both the source and the destination can transmit and receive



(a) (b)

Figure 3.2: Network example with $H_e = 4$ and $m = 2$. (a) The upper bound is tight. (b) The lower bound is tight.

from two nodes and $k = 1$, i.e., Eve wiretaps any one edge of her choice. We refer to the paths in Fig. 3.2a as $p_1,\ p_2,\ p_3$ and p_4, where they are ordered from top to bottom (and also represented with different line patterns). To achieve the outer bound, one can first use p_1 and p_4 and then use p_2 and p_3 to communicate two symbols in each instance of network use. Thus, on two time instances, one can communicate 4 messages (3 securely since $k = 1$). This gives a secure rate of $\frac{3}{2}$, which matches the outer bound.

Example 8: Fig. 3.2b has also $H_e = 4$. However, for $m = 2$ and $k = 1$, it can be shown that the secure capacity is 1, whereas our outer bound in Theorem 16 is $\frac{3}{2}$. In order to achieve a secure rate of one, we can select two node disjoint paths (one through node 1 and the other through node 2) and use them to communicate. We next derive an outer bound for the network in Fig. 3.2b that is tighter than the one in Theorem 16. Assume that, at any time instant t, node 1 transmits symbol $Y_1^{(t)}$ (it can transmit only one symbol even though it has three outgoing edges) and node 2, transmits $Y_2^{(t)}$. Suppose that the network is used n times, then by Fano's inequality,

$$\begin{aligned} nR \leq H(W) &= H(W) - H\left(W|\{Y_i^{(t)},\ i \in [2],\ t \in [n]\}\right) \\ &= I(W; \{Y_i^{(t)},\ i \in [2],\ t \in [n]\}) \\ &= I(W; \{Y_2^{(t)},\ t \in [n]\}) + \end{aligned}$$

$$I(W; \{Y_1^{(t)},\ t \in [n]\} | \{Y_2^{(t)},\ t \in [n]\})$$
$$\overset{(a)}{\leq} n \implies R \leq 1,$$

where (a) is because, if Eve wiretaps the edge outgoing from node 2, then $I(W; \{Y_2^{(t)},\ t \in [n]\}) = 0$ and there are only n symbols in $\{Y_1^{(t)},\ t \in [n]\}$.

3.4 Diamond networks with different path capacities

For the N-relay diamond network (shown in Fig. 3.3) with unit edge capacities ($C_i = 1, \forall i$), the lower and upper bounds in Theorem 15 and Theorem 16 match (since all the N edge disjoint paths are also vertex disjoint, namely $H_e = H_v = N$), and thus the secure capacity equals $C = \min(m, N)(1 - \frac{k}{N})$.

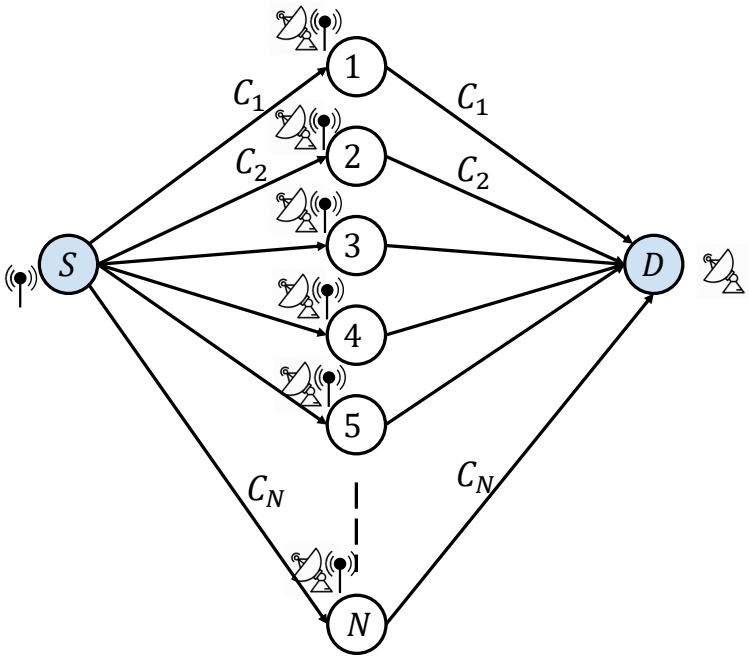


Figure 3.3: Diamond network with different path capacities.

We next consider the case where the edges have non-uniform capacities and $m = 1$. In particular, we assume that for relay $i \in [N]$, both links connecting to the source and the destination have capacity C_i, as depicted in Fig. 3.3. In general, even over traditional networks, the problem of security over unequal capacity edges is everything but easily solvable [CHK13b]. The main reason

is that we need to consider all possible subsets of edges that Eve may wiretap.

Theorem 18. *For the diamond network with $m = 1$ and N relays as shown in Fig. 3.3, the secure capacity equals*

$$C = \max_{\substack{\sum_{i=1}^{N} f_i = 1 \\ f_i \geq 0}} \left[\sum_{i=1}^{N} f_i \, C_i - \max_{\substack{\mathcal{S} \subseteq [N] \\ |\mathcal{S}| = k}} \sum_{i \in \mathcal{S}} f_i \, C_i \right]. \tag{3.8}$$

Proof. **Achievability:** It is clear that we can transmit $\sum_{i=1}^{N} f_i C_i$ symbols from the source to the destination, by using for a fraction f_i of time the path with capacity C_i. Thus, each of the N outgoing edges from the source (and similarly each of the N incoming edges to the destination) will carry $f_1 C_1, \ f_2 C_2, \ \ldots \ f_N C_N$ packets, respectively. The adversary, in the worst case wiretaps k edges, which carry the maximum number of packets. Using a similar encryption scheme as we designed in Section 3.3, ensures a secure rate $\left[\sum_{i=1}^{N} f_i C_i - \max_{\mathcal{S}} \sum_{i \in \mathcal{S}} f_i C_i \right]$, where $\mathcal{S} \subseteq [N], \ |\mathcal{S}| = k$. By optimizing over the f_i's we get that the C in (3.8) is achievable.

Outer bound: Since $m = 1$, at any time instant, the source can transmit on at most one of its N outgoing edges. We let $\{T^t_{e_{i_t}}, \ t \in [n]\}$ be the symbols transmitted over n such instances, where e_{i_t} denotes the edge used in the t-th instance. Some of these symbols will flow through e_1, some through e_2, and similarly some through e_N, where e_i is the edge of capacity C_i outgoing from the source. Let T_{e_i} denote the symbols transmitted on e_i in all such instances. Thus, $\{T^t_{e_{i_t}}, \ t \in [n]\} = \{T_{e_i}, i \in [N]\}$. Let $|T_{e_i}| = n_i, \ i \in [N]$ such that $\sum_i n_i = n$. Because of the edge capacity constraints we have $H(T_{e_i}) \leq n_i C_i, \ \forall i \in [N]$. Now, by Fano's inequality,

$$\begin{aligned}
nR &\leq H(W) = H(W) - H(W|\{T^t_{e_{i_t}}, \ t \in [n]\}) \\
&= I(W; \{T^t_{e_{i_t}}, \ t \in [n]\}) = I(W; \{T_{e_i}, i \in [N]\}) \\
&= I(W; \{T_{e_i}, i \in \mathcal{S}\}) + I(W; \{T_{e_i}, i \notin \mathcal{S}\} | \{T_{e_i}, i \in \mathcal{S}\}) \\
&\overset{(a)}{\leq} \min_{\substack{\mathcal{S} \subseteq [N] \\ |\mathcal{S}| = k}} I(W; \{T_{e_i}, i \notin \mathcal{S}\} | \{T_{e_i}, i \in \mathcal{S}\}) \\
&\leq \min_{\substack{\mathcal{S} \subseteq [N] \\ |\mathcal{S}| = k}} H(\{T_{e_i}, i \notin \mathcal{S}\} | \{T_{e_i}, i \in \mathcal{S}\})
\end{aligned}$$

$$
\begin{aligned}
&\leq \min_{\substack{\mathcal{S}\subseteq[N]\\|\mathcal{S}|=k}} \sum_{i\notin\mathcal{S}} n_i C_i \\
&= \sum_{i\in[N]} n_i C_i - \max_{\substack{\mathcal{S}\subseteq[N]\\|\mathcal{S}|=k}} \sum_{i\in\mathcal{S}} n_i C_i \\
&\implies R \leq \sum_{i\in[N]} f_i C_i - \max_{\substack{\mathcal{S}\subseteq[N]\\|\mathcal{S}|=k}} \sum_{i\in\mathcal{S}} f_i C_i,
\end{aligned}
$$

where (a) follows from the security condition and the choice of $\mathcal{S}$ to have the tightest bound, and $f_i = \frac{n_i}{\sum_{i\in[N]} n_i} \geq 0, \sum_{i\in[N]} f_i = 1$. Optimizing over all such choices of $n_i,\ i \in [N]$, we get that C in (3.8) is an outer bound on the secure capacity. This concludes the proof of Theorem 18. □

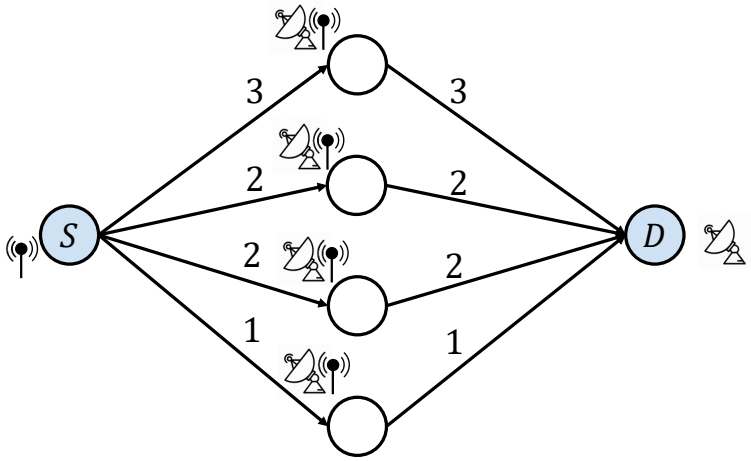


Figure 3.4: Diamond network for Example 9

Example 9: Consider a diamond network with $N = 4$, and $C_1 = 3$, $C_2 = 2$, $C_3 = 2$ and $C_4 = 1$ and assume $k = 1$ as shown in Fig. 3.4. If we were to use each path the same number of times, we would get a secure rate of $\frac{5}{4}$. In contrast, the optimal scheme from Theorem 18 uses the first path twice, the second and third three times each, and does not use the last path, achieving a secure rate of $\frac{3}{2}$. Thus, we see that different from non 1-2-1 networks, here we might need to discard some of the resources in order to get an optimal secure rate.

CHAPTER 4

Distortion Based Security for Cyber-Physical Systems

Traditionally encryption schemes are designed to secure raw bits of the messages. However, when the messages are embedded in a metric space having a notion of distance (for instance, sensor measurements), securing raw bits might not be necessary and is an overkill. In fact, with just a single bit of pre-shared key, a carefully designed encryption scheme can ***distort*** *the eavesdropper's view* ***significantly enough****. In this chapter, we show this with an example of a linear dynamical system that securely communicates its states in the presence of an eavesdropping adversary.*

4.1 Summary

This chapter introduces a new framework called *distortion based security* for the communication in cyber-physical systems (CPS) in the presence of an eavesdropping adversary. We argue that in CPS, the notion of distortion based security is more appropriate and provides theoretical guarantees with a minimal requirement on the pre-shared key. We show that it is possible to confuse the eavesdropper with as little as just one bit of pre-shared key. In particular, we will show that a linear dynamical system can communicate its state to a remote cloud in a manner that prevents an eavesdropper from accurately learning the state.

The linear dynamical system considered in this chapter transmits its state at each time instance, and thus communicates a co-related stream of messages. We identify two notions of distortion at eavesdropper's end and develop encryptions scheme towards maximizing the distortion for each of these two notions. In particular, we define an average-case distortion measure and a worst-case distortion measure. The average-case distortion measure considers an average distance of adversary's estimates of the state to the ground truth. Here, we average over the randomness in

the message transmitted at each time instance and also average across the entire time horizon. On the other hand, the notion of the worst-case distortion is a stronger metric where we consider the closest the adversary comes to the ground truth among all time instances as the measure of security.

The main challenge in designing such encryption schemes is that the encrypted sequences should also follow the same underlying system dynamics in order to not let the eavesdropper separate out the fake transitions from the actual ones. Towards this, our encryption schemes hide the actual state transitions in a set of state transitions, all following the same underlying system dynamics yet having a maximal spatial separation.

For the notion of average-case distortion, our scheme does this by extending the idea of mirroring (illustrated in Chapter 2) to a more general light-weight mappings for dynamical systems in higher dimensional spaces. For the worst-case distortion, the encryption scheme is more involved and achieves a near-perfect distortion with 3 bits of the shared key per dimension (i.e., 9 bits of the pre-shared key for a three-dimensional motion).

Organization: This chapter is organized as follows. Section 4.2 formally defines the problem, and the two notions called the average-case distortion and the worst-case distortion. Section 4.3 and Section 4.5 discusses encryption schemes for these two notions, respectively. Section 4.4 describes transformations which maintain a *symmetry* - this symmetry is used to guarantee the security of encryption schemes with just one bit of the pre-shared key.

4.2 System model

We consider a dynamical system interested in communicating its state transitions to a cloud against an eavesdropping adversary as shown in Fig. 4.1. The linear dynamical system is described as,

$$\begin{aligned} \widetilde{X}_{t+1} &= A\widetilde{X}_t + BU_t + w_t, \\ Y_t &= C\widetilde{X}_t + v_t, \end{aligned} \tag{4.1}$$

where $\widetilde{X}_t \in \mathbb{R}^n$ is the state of the system at time $t \in [1:N]$ with N being the final time instance, $U_t \in \mathbb{R}^m$ is the input to the system at time t, $w_t \in \mathbb{R}^n$ is the process noise, Y_t are the system

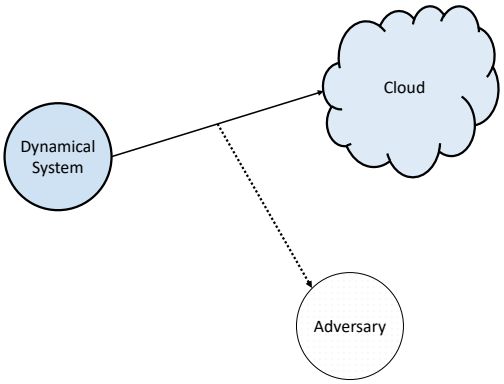


Figure 4.1: Communication in cyber-physical systems.

observations, and $v_t \in \mathbb{R}^p$ is the observation noise. We denote $\widetilde{X}_1^N$ by $\widetilde{X}$, U_1^{N-1} by U and w_1^{N-1} by w. Based on the initial state $\widetilde{X}_1$ and target state $\widetilde{X}_N$, the controller computes a sequence of inputs that moves the state from initial state $\widetilde{X}_1$ to the target state $\widetilde{X}_N$ in N time instances. We assume that the system uses the observations Y_1^N to optimally estimate the states $\widetilde{X}$. The optimal estimates of $\widetilde{X}$ made by the system are denoted by X – in the case of *perfect observation*, *i.e.*, noiseless and observable systems, then $X = \widetilde{X}$.

Communication and adversary models:

At each time instance the system (Alice) transmits information about its state estimate to a legitimate receiver, which is referred to as Bob, via a noiseless link. This situation occurs for example when Bob is remotely monitoring the execution of the system as in Supervisory Control And Data Acquisition (SCADA) systems or in the remote operation of drones.

A malicious receiver, referred to as Eve, is assumed to eavesdrop on the communication between the system and Bob and is able to receive all transmitted signals. The goal of Eve is to make an estimate that is as close to X as possible: since Bob receives X and makes control decisions with this information, Eve is interested in X. Eve is assumed to be passive: she does not actively communicate but is interested in learning the system's states from $t = 1$ to $t = N$. We assume that the System and Bob have a k-bits long pre-shared key K which they use to encode/decode the transmitted messages.

Inputs and states random process model:

We assume that both receivers are only aware of the system model, the matrices A, B, C and the statistics of noises. Therefore, from the perspective of the receivers, the input and output sequences have random distributions which depend on A, B, C and the statistics of the noise. In addition to the process noise w, the joint distribution $f(X, U, w)$ depends on (i) the initial and target states, (ii) the control law of the system, and (iii) the state estimation process. So, even in noiseless systems, X and U possess inherent randomness from a receiver's perspective due to its lack of knowledge about the initial and target states.

Encoding model:

The system encodes and transmits packets Z_1^N to ensure that Bob is able to accurately receive X_1^N, the optimal estimates of the system. To do so, the system transmits a packet Z_t at each time step t. In this work, we use light-weight memoryless encryption schemes. The t-th transmitted packet is a function of only the current state estimate and the pre-shared key, thus, $Z_t := \mathcal{E}_t(X_t, K)$, where $\mathcal{E}_t$ is the encoding function used at time t. We will denote Z_1^N by Z.

Bob/Eve models of decoding:

Bob noiselessly receives the transmitted packets from the system, and decodes them using the pre-shared key. Then, using the decoded information, it generates an estimate of the state of the system at times $t \in [N]$. We require that Bob's estimate is as accurate as Alice's. If we assume that, at time $t \in [N]$, Bob's decoding function is $\Gamma_t\left(Z_1^t, K\right)$, then the previous condition is satisfied by ensuring that $\Gamma_t\left(Z_1^t, K\right) = X_t$ for all $t \in [N]$. Similarly, Eve also receives all transmissions from the system. However, unlike Bob, she does not have the key K. Therefore, Eve's estimate of X_t is $\hat{X}_t := \phi_t\left(Z_1^N\right), t \in [N]$, where ϕ_t is the decoding function used by Eve at time t.

Distortion metrics:

We consider a distortion-based security metric which captures how far an estimate is from the actual value. In particular, our analysis is based on the Euclidean distance as our distance metric. However, our analysis can be extended to any p-norm, since other norms are just a constant factor away, i.e., $\|X\|_p \leq n^{\frac{1}{p}-\frac{1}{q}}\|X\|_q$. We assess the performance of Eve as how far its estimate $\hat{X}$, is from Alice's estimate X. Formally, for a given time instance t and a transmitted codeword Z_1^N, we define the following quantity,

$$D(t, Z_1^N) := \mathbb{E}_{X_t|Z_1^N}\left\|X_t - \hat{X}_t\right\|^2 \overset{(a)}{=} \text{tr}\left(R_{X_t|Z_1^N}\right), \tag{4.2}$$

where (4.2) captures the distortion incurred by Eve while estimating X_t for transmitted symbols Z_1^N. Equality in (a) follows because the best (minimizing) estimates of Eve at time t are,

$$\hat{X}_t = \phi_t\left(Z_1^N\right) = \mathbb{E}\left[X_t|Z_1^N\right].$$

Note that Bob is required to successfully estimate X_t knowing Z_1^t and the key. Therefore, for a given realization of the key, the encoding function can only map one X_t and that key realization to each value of Z_1^N. Therefore Eve realizes that only trajectories from a particular subset can be the true trajectory for a given Z_1^N: those are the ones which correspond to each key realization. Therefore, the expectation in (4.2) is in fact taken over the randomness in the key taking into account posterior probabilities given Z_1^N. If Eve does not have observations, the expectation is taken over X_t with prior distribution and we get $D(t, Z_1^N) = \text{tr}(R_{X_t})$.

As $D(t, Z_1^N)$ is a function of time t and the transmitted sequence Z_1^N, we consider two overall distortion metrics: the "average case" distortion (denoted by D_E) where we take expectation over all possible Z_1^N and average out over time; and the "worst-case" distortion (denoted by D_W) where we take minimum over all possible Z_1^N and time instances.

$$\text{(Average-case distortion) } D_E := \mathbb{E}_{Z_1^N}\left[\frac{1}{N}\sum_{t=1}^{N} D(t, Z_1^N)\right] \tag{4.3}$$

$$\text{(Worst-case distortion) } D_W := \min_{Z_1^N}\left[\min_{t\in[N]} D(t, Z_1^N)\right]. \tag{4.4}$$

It is worth to note that the definitions of D_E and D_W in (4.3) and (4.4) imply that Eve's state estimation must be associated to a time instance. In other words, making a random/constant estimate of the state hoping that it matches the actual state at some time will lead to high distortion values. Further, D_W can be defined even when there is no prior distribution on X_1^N. However, to provide a baseline comparison with the case when the adversary has no observations, we assume that X_1^N always have a known prior distribution.

Design goals:

Our goal is to choose the encoding and decoding functions, $\mathcal{E}_t$ and ϕ_t, so that Bob can decode loselessly while the distortion is maximized for Eve's estimate. In addition, we seek to achieve this with the minimum length of the pre-shared key K. In absence of any observation by Eve, these distortions will be,

$$D_E^{\max} = \frac{1}{N}\sum_{t=1}^{N} \text{tr}(R_{X_t}),$$
$$D_W^{\max} = \min_{t\in[N]} \text{tr}(R_{X_t}).$$

These will serve as upper bounds as,

$$\begin{aligned} D_E &= \frac{1}{N}\mathbb{E}_{Z_1^N}\sum_{t=1}^{N}\text{tr}(R_{X_t|Z_1^N}) \\ &\overset{(a)}{\leq} \frac{1}{N}\sum_{t=1}^{N}\text{tr}(R_{X_t}) \\ &= D_E^{\max}, \end{aligned} \tag{4.5}$$

$$\begin{aligned} D_W &= \min_{Z_1^N}\min_{t\in[N]} \text{tr}(R_{X_t|Z_1^N}) \\ &\leq \min_{t\in[N]} \mathbb{E}_{Z_1^N}\left[\text{tr}(R_{X_t|Z_1^N})\right] \\ &\overset{(b)}{\leq} \min_{t\in[N]} \text{tr}(R_{X_t}) \\ &= D_W^{\max}, \end{aligned} \tag{4.6}$$

where (a) and (b) follows by noting that the trace of the conditional covariance matrix is a quadratic (convex) function in Z_1^N and therefore we can use Jensen's inequality.

4.3 Optimizing average-case distortion D_E

In this section, we will discuss schemes to optimize the average-case distortion (D_E). We will analyze encoding schemes which use *one* bit of the pre-shared key, and characterize their attained level of distortion. We then show that such schemes attain the maximum level of distortion for a family of distributions on X which exhibit a certain class of symmetry.

We now discuss encoding schemes that use one bit of the pre-shared key and show how the achieved distortion compares to the upper bound in (4.5). These encoding schemes work as follows:

$$Z_t = \begin{cases} X_t & \text{if } K = 0, \\ \alpha_t(X_t) & \text{if } K = 1, \end{cases} \quad \forall t \in [N], \tag{4.7}$$

where $K \in \{0, 1\}$ is the shared bit and $\alpha_t(X_t)$ is a transformation of the state vector X_t. We denote by $\alpha_t^{-1}(X_t)$ the inverse transformation of α_t. We will next show the attained distortion of such schemes.

Theorem 19 (Proof in Appendix A.7)**.** *The average-case distortion (D_E) attained by using the scheme in* (4.7) *is,*

$$\frac{1}{2N} \sum_{t=1}^{N} \mathbb{E}_X \left\{ \frac{f_X(\alpha^{-1}(X))}{f_X(X) + f_X(\alpha^{-1}(X))} \left\| X_t - \alpha_t^{-1}(X_t) \right\|^2 \right\}, \tag{4.8}$$

where $\alpha^{-1}(X) := [(\alpha_1(X_1))^T, (\alpha_2(X_2))^T, \cdots (\alpha_N(X_N))^T]^T$. Moreover, if the following condition holds,

$$f_X(x) = f_X(\alpha^{-1}(x)),\ \forall x \in \mathcal{X}, \tag{4.9}$$

then the expression simplifies to

$$D_E = \frac{1}{4N} \sum_{t=1}^{N} \mathbb{E}_X \left\| X_t - \alpha_t(X_t) \right\|^2. \tag{4.10}$$

Condition (4.9) implies a general notion of symmetry in the distribution of $f_X(x)$. In the following, we focus on a particular notion of distribution symmetry, for which we show the corresponding choice of $\alpha_t(X_t)$ and how it can achieve high levels of distortion. Consider a transformation function $\alpha_t(x)$ which reflects a point x across an affine subspace of dimension d, defined

by the equations $S_t x = b_t$ where $S_t \in \mathbb{R}^{d\times n}$ consists of $d \leq n$ orthonormal rows, and $b_t \in \mathbb{R}^d$; the transformation is $\alpha_t(x) = \left(I - 2(S_t)^T S_t\right) x + 2(S_t)^T b_t$. The choice of the dimension d and the subspace (S_t, b_t) depend on the properties we would like the encoded trajectories to have. We refer to encoding schemes that are based on this transformation as *mirroring schemes*. For example, consider $X_t \in \mathbb{R}^2$ where $S_t = \frac{1}{\sqrt{2}}[-1,\ 1]$ and $b_t = 0$. Then $\alpha_t(X_t)$ corresponds to mirroring across a line that passes through the origin with a 45° angle. This is shown in Fig. 4.2. We are interested in *mirroring schemes* as they are light-weight and can be implemented on low-complexity IoT devices. Moreover, such schemes can provide the maximum distortion level for a class of distributions with what we refer to as *point-symmetry*.

Definition 5 (**Point symmetry**). *A random vector X is said to have point symmetry if there exists a point v for which $f_X(x) = f_X(2v - x),\ \forall x \in \mathcal{X}$.*

Lemma 20. *If X has point symmetry across v, then $v = \mu_X$.*

Proof. Since X has point symmetry, then

$$\begin{aligned} & f_X(x) = f_X(2v - x) \\ \Rightarrow\ & f_X(x) = f_{2v-X}(x) \\ \Rightarrow\ & \mu_X = 2v - \mu_X \\ \Rightarrow\ & \mu_X = v. \end{aligned}$$

□

The following result characterizes the performance of the mirroring scheme, and shows that it achieves the maximum distortion for distributions with point symmetry.

Corollary 21. *If $\alpha_t(X_t)$ is based on a **mirroring scheme** along the planes given by $S_t x = b_t,\ t \in [N]$ and the condition* (4.9) *holds, then* (4.10) *becomes,*

$$D_E = \frac{1}{N}\sum_{i=1}^{N} tr\left(S_t R_{X_t}(S_t)^T + (b_t - S_t\mu_{X_t})(b_t - S_t\mu_{X_t})^T\right). \tag{4.11}$$

Moreover, if X has point-symmetry, then $D_E = \frac{1}{N}\sum_{t=1}^{N} tr(R_{X_t})$, the maximum possible distortion.

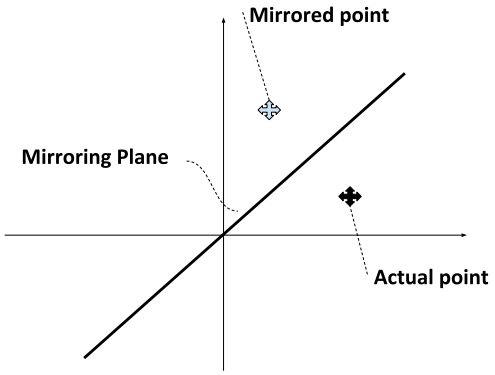


Figure 4.2: Mirroring across the line passing through the origin at 45° angle with the X-axis.

Proof. If condition (4.9) holds, then by simply plugging the expression of $\alpha_t(X_t)$ for the mirroring scheme along $S_t x = b_t$ that is $\alpha_t(X_t) = \left(I - 2(S_t)^T S_t\right) x + 2(S_t)^T b_t$ in (4.10) we get (4.11) (Formal proof in Appendix A.7). Choosing $S_t = I$ and $b_t = \mu_{x_t}$ makes $\alpha^{-1}(X_1^N) = 2\mu_{X_1^N} - X_1^N$ which by point-symmetry satisfies (4.9). Therefore, we get $D^E = \frac{1}{N} \sum_{t=1}^{N} tr(R_{X_t})$. □

Now, we show the implications of our results for mirroring based schemes in the context of a few examples.

Example 10: Consider an example where U is distributed as Gaussian with mean μ_U and covariance matrix R_U. Then for a noiseless system with perfect observation and a zero initial state, X_2^N is also Gaussian distributed. A Gaussian random vector has point-symmetry and therefore, according to Corollary 21, we can get maximum distortion by setting $b_t = \mu_{X_t}$ and $S_t = I$.

The next example is based on a Markov-based model for the dynamical system. For this example, the following lemma is useful.

Lemma 22. *Consider the random vector X_1^N where the following conditions hold: 1) $f_{X_1}(x_1)$ has point-symmetry, and 2) $f_{X_t|X_1^{t-1}}(x_t|x_1^{t-1})$ has point-symmetry, then so does $f_X(X)$, where $X = X_1^N$ and $\mu_X = [(\mu_{X_1})^T, (\mu_{X_2})^T, \cdots, (\mu_{X_N})^T]^T$. Therefore, by virtue of Corollary 21, mirroring schemes can achieve the maximum distortion.*

Lemma 22 allows us to characterize the performance of the following example.

Example 11: Consider the following random walk mobility model. Let $a \in \mathbb{N}^+$, and X_t be its location at time t, then,

$$\begin{aligned} X_1 &\sim \text{Uni}([-a:a]), \\ X_t|X_{t-1} &\sim \text{Uni}([-a:a] \cap \{X_{t-1}-1, X_{t-1}, X_{t-1}+1\}). \end{aligned}$$

One can see that these distributions satisfy the conditions in Lemma 22. Therefore, one can set $b_t = \mu_t = 0$ and $S_t = 1$, which will achieve maximum distortion of D_E.

Example 12: Here we provide a numerical example which shows how our mirroring scheme performs for situations where we compute the state distributions using numerical simulations. In the Section 4.4, we will also show that the controller used in this example falls under the class of controller where we do not need to compute the distribution on states and can directly apply our scheme to achieve the perfect distortion. We consider the quadrotor dynamical system provided in (4) of [KM12]. The quadrotor moves in a 3-dimensional cubed space with a width, length and height of 2 meters, where the origin is the center point of the space. The quadrotor starts its trajectory from an initial point $(-1, y_1, z_1)$ and finishes its trajectory at a target point $(1, y_N, z_N)$ after N time steps, where the points y_1, z_1, y_N, z_N are picked uniformly at random in $[-1, 1]^4$. We assume that $N = 10$ time steps, and that the continuous model in [KM12, (4)] is discretized with a sample time of $N_s = 0.5$ seconds. We assume that the quadrotor encodes and transmits only the states which contain the location information (first three elements of the state vector X_t). The quadrotor is equipped with an LQR controller which designs the input sequence U_1^{N-1} as the solution of the following problem

$$\begin{aligned} \text{minimize} \quad & \|U\|^2 + 10 \left\|X_2^{N-1}\right\|^2 \\ \text{subject to} \quad & X_{t+1} = A^{\text{quad}} X_t + B^{\text{quad}} U_t, \quad \forall t \in [N-1] \\ & X_1 = \begin{bmatrix} -1 & y_1 & z_1 & 0 & \cdots & 0 \end{bmatrix}^T, \\ & X_N = \begin{bmatrix} 1 & y_N & z_N & 0 & \cdots & 0 \end{bmatrix}^T, \end{aligned} \tag{4.12}$$

where A^{quad} and B^{quad} define the quadrotor's discrete-time model. The remaining states of X_1 and X_N are set to zero to allow the drone to hover at the respective locations. We perform numerical simulation of the aforementioned setup: we run 2 millions iterations, where in each iteration a new

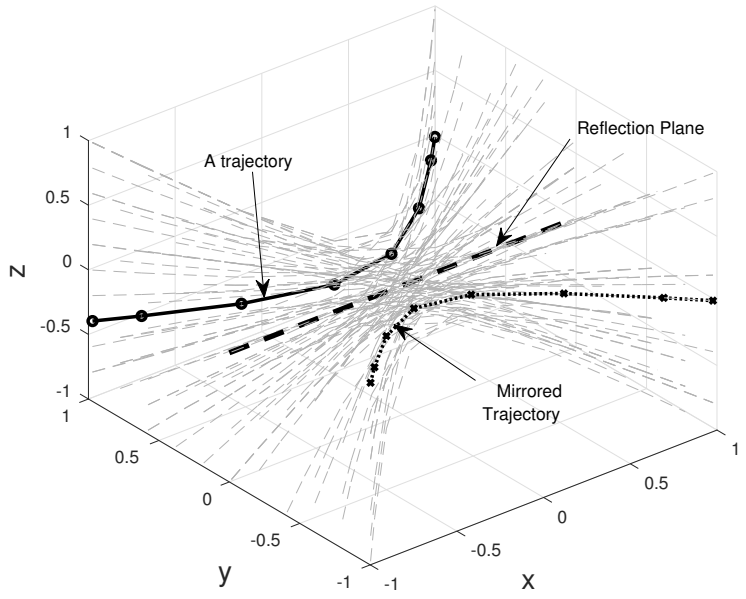


Figure 4.3: An illustration of some trajectories. The reflection plane is shown as a dashed-black line. One trajectory (solid-black) is shown along with its mirrored image (dotted-black).

initial and target points are picked, and the resultant trajectory is recorded. Based on the recorded data, we consider different mirroring schemes and numerically evaluate the attained distortion. To facilitate numerical evaluations, the simulation space is gridded into bins with 0.2 meters of separation, and the location of the drone is approximated to the nearest space bin. Figure 4.3 shows some of the drone trajectories obtained from our numerical simulation. It is clear that not all trajectories are equiprobable, and therefore the distribution of X_t is not uniform across all bins in space. Since the motion of the drone is mainly progressive in the positive x-axis direction, reflection across a fixed point results in mirrored trajectories that are progressing in the opposite direction, and therefore are identified to be fake automatically. Therefore, mirroring across a point here is useless: the numerically computed distortion for this scheme is equal to zero.

Next we consider mirroring across the reflection plane shown in Figure 4.3, where $b_t = 0$ and

$$S_t = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As can be seen from the figure, the reflection plane is indeed an axis of symmetry for the distribution of the drones trajectories, and therefore is expected to provide high distortion values. We numerically evaluate the attained distortion using the scheme by using equation (4.8), which

evaluates to $D_E = 0.3971$. This is slightly less than $D_E^{\max} = 0.3979$.

4.4 Transformations maintaining point symmetry

Encoding and decoding schemes such as the ones mentioned in Section 4.3 can be generally used for any dynamical system with arbitrary distributions on the inputs U, the state vectors $\widetilde{X}_t$ and the state estimates X_t. However, characterizing the attained level of average distortion (using expressions (4.8)) requires the knowledge of the distribution of the state estimate. While a distribution can be obtained for the initial and target state vectors, it may be difficult to incorporate the system dynamics, the estimation method as well as the controller into the process of finding a distribution of the inputs, states and states estimate. In such cases, numerical evaluations can aid into finding the needed distribution, as was shown in Example 12 in Section 4.3. Although it is necessary to find the state distribution in order to characterize the distortion, the knowledge of existing symmetries in the distribution can directly give possible choices for the transformation function $\alpha_t(\cdot)$ which may attain high levels of distortion; for example, if there is a point symmetry in the distribution, mirroring across the symmetry point attains the maximum possible distortion. In this section, we ask the following question: "under which conditions on the dynamical system, does point symmetry in the initial and target states results into point symmetry on the states estimate?"

A general answer to the aforementioned question appears to be difficult. Therefore, we limit our answer in this work to the scope of linear controllers. or a given initial and target state, let $X^{(\text{ref})}$ be the reference trajectory that the control system ideally wishes to follow. We assume that the system controller selects an input vector that is a linear function of $X^{(\text{ref})}$. In many cases, $X^{(\text{ref})}$ is also a linear function of the initial and target states (*e.g.*, when the reference trajectory is the solution of an LQR problem for the noiseless version of the system). Then we can write $U_t = K_t \left(X_t - X_t^{(\text{ref})} \right)$. Moreover, we assume that the optimal estimation function that the system uses is a linear one in the observations, *i.e.*, we assume X_t is a linear function of Y_1^t, X_{init} and X_{target}. By incorporating the controller and estimation equations into the system dynamics, one can arrive at the following relation $X = MQ$, where $Q = \left[(X_{\text{init}})^T, (X_{\text{target}})^T, (w_1^N)^T, (v_1^N)^T \right]^T$, and the matrix M is a function of the matrices A, B, C, K_t and the linear function used in the estimation

of state X_t from the observations. We assume that w_1^N and v_1^N are uncorrelated Gaussian random vectors. We first prove the following lemma.

Lemma 23. *If a random vector $V_1 \in \mathbb{R}^n$ has point-symmetry across μ_{V_1}, and g is an affine function, then the random vector $V_2 = g(V_1)$ has point-symmetry across $g(\mu_{V_1})$.*

Proof. If V_1 has point-symmetry, then the following conditions are equivalent:

$$f_{V_1}(v_1) = f_{V_1}(2\mu_{V_1} - v_1),\ \forall v_1 \in V_1,$$
$$f_{V_1}(v_1) = f_{2\mu_{V_1} - V_1}(v_1),\ \forall v_1 \in V_1.$$

Thus, to prove that V_2 also has point-symmetry, it suffices to prove that the density of V_2 and $2\mu_{V_2} - V_2$ is the same. Consider the two random vectors W_1 and W_2. If they have the same support and the same density function, then $g(W_1)$ and $g(W_2)$ will also have the same density for any function g; we denote this by writing $W_1 \sim W_2$. Thus,

$$\begin{aligned} V_1 &\sim 2\mu_{V_1} - V_1 \\ g(V_1) &\sim g(2\mu_{V_1} - V_1) \\ M_1V_1 + M_2 &\sim 2M_1\mu_{V_1} - M_1V_1 + M_2 \\ V_2 &\sim 2(M_1\mu_{V_1} + M_2) - (M_1V_1 + M_2) \\ V_2 &\sim 2(\mu_{V_2}) - V_2. \end{aligned}$$

Thus, V_2 has a point of symmetry. □

Theorem 24. *If X_{init} and X_{target} are independent of w_1^N and v_1^N, and both have point symmetries, then the vectors X_t as well as X will all have Point Symmetries for any matrix M.*

Proof. First, note that w_1^N and v_1^N are Gaussian random vectors, and therefore have point symmetries across their mean points. Since X_{init} and X_{target} are independent of w_1^N and v_1^N, then the vector $Q = \left[(X_{\text{init}})^T, (X_{\text{target}})^T, (w_1^N)^T, (v_1^N)^T\right]^T$ also has a point-symmetry across the mean point (which is the concatenation of the mean points of the respective components of Q); we denote this point by μ_Q. Then, by virtue of Lemma 23, X (respectively X_t) will also have point-symmetry across the point $M\mu_Q$ (respectively across the point μ_Q left-multiplied by the corresponding section of the matrix M). □

Revisiting Example 12 of Section 4.3: Example 12 in Section 4.3 shows an example where the initial and target points exhibit point-symmetry. In such an example, the LQR controller is a linear function of the previous states (one can find such a controller by applying the KKT conditions). Since the system is noiseless, then the estimated states are equal to the observations. Therefore, the conditions for Theorem 24 are met, and point-symmetry is preserved for X_t and the whole trajectory X. Note, however, that the point of symmetry for X_t changes with t, *i.e.*, it progresses along the x axis as shown in Figure 4.3.

4.5 Optimizing the worst-case distortion D_W

The expected distortion metric might not be well-suited for some applications (for example if an adversary wants to shoot a drone). In this case, the adversary's estimate needs to be far from the actual state *at all* time instances. Therefore, a more appropriate metric would be to consider the worst-case distortion for the adversary. Consider for example the scheme in Fig. 1.4b. Here, the adversary's estimate is always the center point and therefore the maximum expected distortion is achieved. However, when the drone is close to the center, its mirror image will also be close to the center. At this particular time instance, the adversary's distortion will be very small and thus the adversary will essentially know the position.

In this section, we present an encryption scheme that attempts to maximize the worst-case distortion for Eve. The main idea is to obfuscate the initial state in such a way that Eve, even if she optimally uses her knowledge about the dynamics and her observations, her best estimate is close to the maximal distortion. We start by studying the problem of distorting the transmission of a single random variable in Theorems 26 and 27. These results then form the basis for maximizing the worst-case distortion of a trajectory, as described in Theorem 28.

4.5.1 Building step: scalar case:

Consider the case where the system wants to communicate a single scalar random variable X to Bob by transmitting Z. The worst-case distortion D_W for Eve will be $D_W = \min_Z \text{Var}(X|Z)$.

Note that if Eve does not overhear Z, Eve uses the minimum mean square error estimate, (*i.e.*, the mean value) as her estimate, and thus experience a worst-case distortion equal to the var(X).

We first assume that $X \sim \mathcal{N}(0,1)$, and thus, the worst-case distortion can not be larger than 1 by (4.6). We next develop our scheme progressively, from simple to more sophisticated steps. We will also use the following lemma.

Lemma 25. *The variance of two real numbers a and b with probabilities p_a and p_b is given by $p_a p_b (a-b)^2$.*

Mirroring: Reflecting around the origin (as we did for optimizing the average case distortion in Section 4.3) does not work well when X takes small values: indeed Var$(X|Z)$ is $Pr(X = Z|Z)(Pr(X=-Z|Z))(Z-(-Z))^2$ using Lemma 25 and has a worst-case value that goes to zero as Z approaches zero.

Shifting: To avoid this, we could try to use a "shifting" scheme where we add a constant θ to X whenever the shared key bit is one; but now this scheme does not perform well for large values of Z: as Z increases Var$(X|Z)$ goes to zero. This is because using Lemma 25:

$$\begin{aligned}\text{Var}(X|Z) &= Pr(X=Z|Z)(Pr(X=Z-\theta|Z))(Z-(Z-\theta))^2 \\ &= Pr(X=Z|Z)(Pr(X=Z-\theta|Z))(\theta)^2,\end{aligned}$$

and $Pr(X = Z|Z)(Pr(X = Z-\theta|Z))$ goes to zero for large value of Z.

Shifting+Mirroring: We here combine shifting and mirroring, in order to achieve a good performance for both small and large values of X. We start from the case where we have $k = 1$ bit of the pre-shared key and then go to the case $k \geq 1$.

- $k = 1$. We select a $\theta_1 \in \mathbb{R}$ that determines a window size (θ_1 is public and known by Eve). The encoding function is

$$Z = \mathcal{E}(X,K) = \begin{cases} X & \text{if } K=0 \\ -X & \text{if } K=1,\ |X| > \theta_1 \\ X+\theta_1 & \text{if } K=1,\ -\theta_1 \leq X < 0 \\ X-\theta_1 & \text{if } K=1,\ 0 \leq X < \theta_1 \end{cases}$$

We note that there is one particular value of X, $X = \theta_1$, which we do not transmit. Since this is of zero probability measure, it can be safely ignored. Given Z, there are two possibilities for X:

$$X \in \begin{cases} \{Z, -Z\} & \text{if } |Z| > \theta_1 \\ \{Z, Z + \theta_1\} & \text{if } -\theta_1 \leq Z < 0 \\ \{Z, Z - \theta_1\} & \text{if } 0 \leq Z < \theta_1. \end{cases}$$

Using the fact that $X \sim \mathcal{N}(0, 1)$, we can calculate the posterior probabilities $Pr(X|Z)$ and use

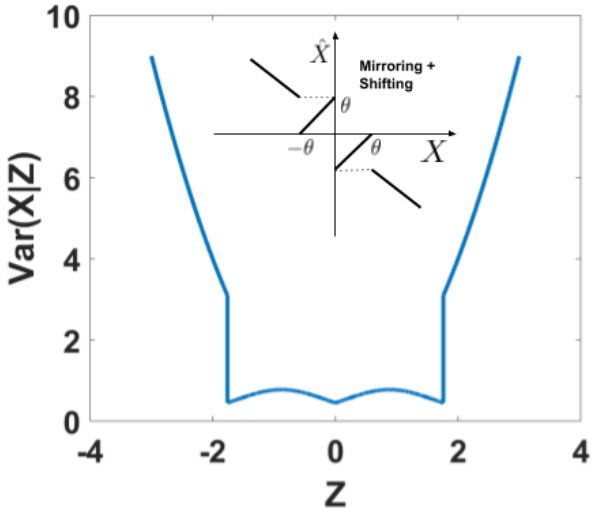


Figure 4.4: Var$(X|Z)$ Vs Z for shifting+mirroring based scheme with $\theta_1 = 1.76$; $D_W = 0.4477$.

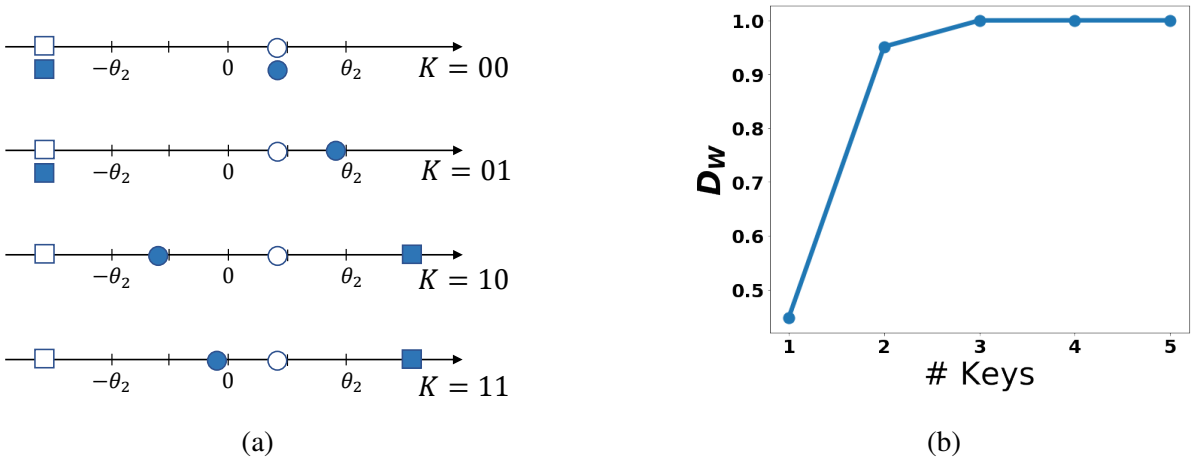
Lemma 25 to compute Var$(X|Z)$. Fig. 4.4 plots Var$(X|Z)$ for $\theta_1 = 1.76$. The worst-case distortion in this case becomes 0.4477, which is the best we can hope for if we have only one bit of the pre-shared key. This follows because for any mapping from X to Z, a transmitted symbol Z can have at most two pre-images (as Bob needs to reliably decode with the one bit of the pre-shared key), and if one of these is $X = 0$, then no matter what the second one is, the distortion corresponding to Z will be at most 0.4477. Equality occurs when the second pre-image of Z is ± 1.76. Note that our scheme also maps 0 to -1.76 (for $\theta_1 = 1.76$).

- $k \geq 1$. For $K \in \{0, 1\}^k$, we use the following encoding:

$$Z = \mathcal{E}(X, K) \tag{4.13}$$

$$= \begin{cases} \begin{cases} X & \text{if } K_d < 2^{k-1} \\ -X & \text{if } K_d \geq 2^{k-1} \end{cases} & |X| > \theta_k \\ X + K_d \frac{2\theta_k}{2^k} \bmod [-\theta_k, \theta_k) & X \in [-\theta_k, \theta_k), \end{cases}$$

where the optimal value of the constant θ_k depends on the length k of the key K we have, K_d is the decimal equivalent of K, and $r \bmod [a, b) = r - i(b - a)$ is such that i is an integer and $r - i(b - a) \in [a, b)$ for $r, a, b \in \mathbb{R}$. Intuitively, if $|X| > \theta_k$ then for half of the times, we reflect across origin and for other half we do nothing; if $|X| < \theta_k$, we divide this window of size $2\theta_k$ into 2^k equal size windows and shift a point from one window to another by jumping K_d windows. An example for $k = 2$ is shown in Fig. 4.5a for the key values $K = 11$ and $K = 10$. Fig. 4.5b plots D_W as a function of the length k of the pre-shared key K. Using $k = 3$ and $\theta_3 = 4.84$ we achieve $D_W = 0.9998$ which is very close to 1, the best we can hope for.



(a) (b)

Figure 4.5: (a) Scheme for $k = 2$: Transparent shapes are true values and solid shapes represent their respective mapping. (b) D_W with the length k of the pre-shared key K for the optimal choice of θ_k.

Theorem 26. *A Gaussian random variable with mean μ and variance σ^2 can be near perfectly (~ 0.9998 times the perfect distortion) distorted in worst-case settings by just using three bits of the pre-shared key.*

Proof. Generate the random variable $V \sim \mathcal{N}(0, 1)$ as $V = (X - \mu)/\sigma$ and encrypt it using $k = 3$

key bits and the previously described scheme. For $c = 0.9998$ we have

$$D_W = \min_Z \text{Var}(X|Z) = \min_Z \text{Var}(\sigma V + \mu|Z) = \sigma^2 \min_Z \text{Var}(V|Z) = c\sigma^2.$$

□

4.5.2 Vector case and time series

Theorem 27 (Proof in Appendix A.8)**.** *For a Gaussian random vector $X \in \mathbb{R}^n$ with mean μ and a diagonal covariance matrix Σ we can achieve D_W within 0.9998 of the optimal by using $3n$ bits of the pre-shared key.*

This theorem uses our 3-bit encryption for each element in the vector. Assume now that this vector captures the probability distribution of the initial state of dynamical system; by encrypting this state we can guarantee the following.

Theorem 28 (Complete Proof in Appendix A.9)**.** *Using $3n$ bits of the pre-shared key we can achieve D_W within 0.9998 of the optimal for the dynamical systems* (4.1) *with $C = I$, $v_t = 0$, singular values of A having absolute value larger than 1, and initial state $X_1 \sim \mathcal{N}(\mu, \Sigma)$, where Σ is diagonal covariance matrix, and U_t and w_t are independent of X_t.*

Remark: Although the independence assumption on the inputs is rather restrictive, the result serves as a stepping stone towards understanding general cases.

Proof. The system transmits $Z_1 = f(Y_1, K) = f(X_1, K)$ where f is the encoding in Theorem 27, and for $t \in [N-1]$,

$$Z_{t+1} = AZ_t + (Y_{t+1} - AY_t) = AZ_t + BU_t + w_t.$$

Bob can decode X_1 using Z_1 and K. Then:

$$\begin{aligned}\hat{X}_{t+1} &= Z_{t+1} - AZ_t + A\hat{X}_t \\ &= (AZ_t + BU_t + w_t) - AZ_t + A\hat{X}_t \\ &= AX_t + BU_t + w_t = X_{t+1}, \quad \forall t \in [N-1].\end{aligned}$$

Eve's distortion is calculated in the Appendix A.9. □

CHAPTER 5

Discussion and Open Directions

In this thesis, we studied three different scenarios of communication in the presence of eavesdropping adversaries. In each of these scenarios, we designed communication schemes requiring either no or a small pre-shared key. However, there are still open problems in each of these scenarios; which we discuss in what follows.

Multiple unicast traffic over wireline networks: We designed a scheme that is capacity-achieving for several different cases, namely: (i) arbitrary networks with two destinations; (ii) arbitrary separable networks with three destinations; and (iii) arbitrary separable networks where only one edge is eavesdropped. However, proving Conjecture 1, which claims the scheme to be capacity achieving for arbitrary separable networks remains open. The capacity characterization for networks which are not separable also remains open.

Characterizing the maximum information flow for networks with an arbitrary number of sources and destinations is a long standing open problem. For secure capacity, its connection to the *edge removal problem* has also been explored in the information theory literature but it remains largely open for networks with an arbitrary number of sources and destinations [LLE13].

Millimeter wave networks (1-2-1 networks): We derived upper and lower bounds on the secure capacity of 1-2-1 networks where all nodes, except the source and the destination, have only one transmitting and one receiving antenna. However, the exact secure capacity characterization remains open. Moreover, secure capacity characterization for multicast and multiple unicast traffic, and also the case of non-uniform edge capacities remain open.

Practical problems in steering antenna arrays and their alignments with the receiver's beam have associated costs, and an analysis with these costs would be an interesting future work.

Distortion based security for cyber-physical systems: For the average-case distortion, we designed a scheme that for several distributions having a point symmetry, distorts the eavesdropper's view with just one bit of shared key. However, designing schemes for other general distributions remains open. For the worst-case distortion, our scheme assumes an open loop controller and a Gaussian distribution on the initial state. Designing schemes tailored for the worst-case distortion without these assumptions remains open.

Moreover, identifying applications specific distortion measures and developing encryption schemes tailored towards them remain open in general. Theoretical analysis of the security aspect of these schemes when eavesdroppers might obtain additional side information is also an important direction for future work.

APPENDIX A

Appendices

A.1 Proof of security: Theorem 3, case 1

We here prove that, for any choice of G, there exists a U in

$$X = \begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix} \begin{bmatrix} W \\ K \end{bmatrix},$$

such that

$$rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) = rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right) \tag{A.1}$$

for every $|\mathcal{Z}| \leq k$. Towards this end, we select a random U and show that the probability of the condition in (A.1) being satisfied is non-zero for a sufficiently large field. This proves the existence of such a matrix U. We have

$$\begin{aligned}
&\Pr\left\{rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) = rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right),\ \forall|\mathcal{Z}| \leq k\right\} \\
&= 1 - \Pr\left\{\bigcup_{|\mathcal{Z}|\leq k}\left[rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) \neq rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right)\right]\right\} \\
&\overset{(a)}{\geq} 1 - \sum_{|\mathcal{Z}|\leq k}\Pr\left\{rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) \neq rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right)\right\} \\
&\overset{(b)}{\geq} 1 - \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^k \max_{|\mathcal{Z}|\leq k}\Pr\left\{rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) \neq rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right)\right\} \\
&= 1 - \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^k \max_{|\mathcal{Z}|\leq k}\left(1 - \Pr\left\{rk\left(\left.\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{bmatrix}\right|_{\mathcal{Z}}\right) = rk\left(\left.\begin{bmatrix} G \\ U \end{bmatrix}\right|_{\mathcal{Z}}\right)\right\}\right)
\end{aligned}$$

$$\overset{\text{(c)}}{\geq} 1 - \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^k \max_{|\mathcal{Z}|\leq k} \left(1 - \left(1 - \frac{1}{q}\right)^k\right)$$
$$\overset{\text{(d)}}{>} 0,$$

where: (i) the inequality in (a) follows by using the union bound; (ii) the inequality in (b) follows since $\binom{n}{t} \leq \left(\frac{\mathrm{e}n}{t}\right)^t$, and (iii) the inequality in (d) holds for sufficiently large q. In order to show the inequality in (c), assume that $\mathcal{Z}$ corresponds to k_1 rows in $\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \end{bmatrix}$ indexed by $\mathcal{Z}_1$ and k_2 rows in $\begin{bmatrix} I_{R_1+R_2} & U \end{bmatrix}$ indexed by $\mathcal{Z}_2$ with $k_1 + k_2 \leq k$. With this, we have that

$$rk\left(\left[\begin{array}{cc} 0_{\ell\times(R_1+R_2)} & G \\ I_{R_1+R_2} & U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = \underbrace{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}_1}\right)}_{\hat{k}_1\leq k_1} + k_2.$$

This follows since, because of the structure of the matrix, the rows in the block $\begin{bmatrix} 0_{\ell\times(R_1+R_2)} & G \end{bmatrix}$ are linearly independent of the rows in the block $\begin{bmatrix} I_{R_1+R_2} & U \end{bmatrix}$. Moreover, we have

$$rk\left(\left[\begin{array}{c} G \\ U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = \hat{k}_1 + k_2,$$

with probability

$$\begin{aligned} p &= \prod_{j=1}^{k_2} \left(1 - \frac{q^{\hat{k}_1+j-1}}{q^k}\right) \\ &\overset{\text{(e)}}{\geq} \prod_{j=1}^{k_2} \left(1 - q^{-1}\right) \\ &= \left(1 - \frac{1}{q}\right)^{k_2} \\ &\overset{\text{(f)}}{\geq} \left(1 - \frac{1}{q}\right)^{k}, \end{aligned}$$

where: (i) the inequality in (e) follows since $\hat{k}_1 + j - 1 - k \leq -1$ for all $j \in [k_2]$, and (ii) the inequality in (f) follows since $k_2 \leq k$. This shows that the inequality in (c) above holds.

A.2 Proof of security: Theorem 3, case 2

We here prove that, for any choice of S and G satisfying $rk\left(\left[\begin{array}{cc} S & G \end{array}\right]\Big|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right)$ for all $|\mathcal{Z}| \leq k$, there exists a U in

$$X = \left[\begin{array}{ccc} S & 0_{\ell\times r} & G \\ 0_{r\times t} & I_r & U \end{array}\right] \left[\begin{array}{c} W' \\ W'' \\ K \end{array}\right],$$

such that

$$rk\left(\left[\begin{array}{ccc} S & 0_{\ell\times r} & G \\ 0_{r\times t} & I_r & U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G \\ U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) \tag{A.2}$$

for every $|\mathcal{Z}| \leq k$. Towards this end, we select a random U and show that the probability of the condition in (A.2) being satisfied is non-zero for a sufficiently large field. This proves the existence of such a matrix U. By following similar steps as in the proof of Case 1 in Appendix A.1, we get

$$\begin{aligned}
&\Pr\left\{rk\left(\left[\begin{array}{ccc} S & 0_{\ell\times r} & G \\ 0_{r\times t} & I_r & U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G \\ U \end{array}\right]\Bigg|_{\mathcal{Z}}\right),\ \forall|\mathcal{Z}| \leq k\right\} \\
&\geq 1 - \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^k \max_{|\mathcal{Z}|\leq k}\left(1 - \Pr\left\{rk\left(\left[\begin{array}{ccc} S & 0_{\ell\times r} & G \\ 0_{r\times t} & I_r & U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G \\ U \end{array}\right]\Bigg|_{\mathcal{Z}}\right)\right\}\right) \\
&\overset{(a)}{\geq} 1 - \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^k \max_{|\mathcal{Z}|\leq k}\left(1 - \left(1 - \frac{1}{q}\right)^k\right) \\
&\overset{(b)}{>} 0,
\end{aligned}$$

where the inequality in (b) holds for sufficiently large q. In order to show the inequality in (a), assume that $\mathcal{Z}$ corresponds to k_1 rows in $\left[\begin{array}{ccc} S & 0_{\ell\times r} & G \end{array}\right]$ indexed by $\mathcal{Z}_1$ and k_2 rows in $\left[\begin{array}{ccc} 0_{r\times t} & I_r & U \end{array}\right]$ indexed by $\mathcal{Z}_2$ with $k_1 + k_2 \leq k$. With this, we have that

$$\begin{aligned}
rk\left(\left[\begin{array}{ccc} S & 0_{\ell\times r} & G \\ 0_{r\times t} & I_r & U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) &= rk\left(\left[\begin{array}{cc} S & G \end{array}\right]\Big|_{\mathcal{Z}_1}\right) + k_2 \\
&= \underbrace{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}_1}\right)}_{\hat{k}_1\leq k_1} + k_2.
\end{aligned}$$

This follows since, because of the structure of the matrix, the rows in the block $\begin{bmatrix} S & 0_{\ell\times r} & G \end{bmatrix}$ are linearly independent of the rows in the block $\begin{bmatrix} 0_{r\times t} & I_r & U \end{bmatrix}$. Moreover, we have

$$rk\left(\left[\begin{array}{c} G \\ U \end{array}\right]\Bigg|_{\mathcal{Z}}\right) = \hat{k}_1 + k_2,$$

with probability $\prod_{j=1}^{k_2}\left(1-\frac{q^{\hat{k}_1+j-1}}{q^k}\right) \geq \left(1-\frac{1}{q}\right)^k$.

A.3 Proof of Lemma 5

In this section, we use an iterative algorithm that, for any permutation $\pi = \{\pi(1), \ldots, \pi(m)\}$ of $[m]$, allows to select $R_{\pi(i)}$ vectors from $N_{\pi(i)}$ (with $R_{\pi(i)}$ being defined in (2.14)) so that all the selected $\sum_{i=1}^{m} R_i$ vectors are linearly independent. We next illustrate the main steps of the proposed algorithm.

1. We select $R_{\pi(1)} = \dim(N_{\pi(1)})$ independent vectors from $N_{\pi(1)}$. Note that one possible choice for this consists of selecting the basis of the subspace $N_{\pi(1)}$.

2. Next we would like to select independent vectors from $N_{\pi(2)}$ that are also independent of the $R_{\pi(1)}$ vectors that we selected in the previous step. Towards this end, we note that a basis of the subspace $N_{\pi(1)} + N_{\pi(2)}$ is a subset of the union between a basis of $N_{\pi(1)}$ and a basis of $N_{\pi(2)}$. Therefore, we can keep selecting vectors from a basis of $N_{\pi(2)}$ as long as we select an independent vector. Since there are $\dim(N_{\pi(1)} + N_{\pi(2)})$ independent vectors in a basis of $N_{\pi(1)} + N_{\pi(2)}$, then we can select

   $$R_{\pi(2)} = \dim(N_{\pi(1)} + N_{\pi(2)}) - \dim(N_{\pi(1)})$$

   independent vectors from $N_{\pi(2)}$ that are also independent of the $R_{\pi(1)}$ vectors that we selected in the previous step.

3. Similar to the above step, we now would like to select independent vectors from $N_{\pi(3)}$ that are also independent of the $R_{\pi(1)} + R_{\pi(2)}$ vectors that we selected in the previous two steps. Towards this end, we note that a basis of the subspace $N_{\pi(1)} + N_{\pi(2)} + N_{\pi(3)}$ is a subset of

the union between a basis of $N_{\pi(1)} + N_{\pi(2)}$ and a basis of $N_{\pi(3)}$. Therefore, we can keep selecting vectors from a basis of $N_{\pi(3)}$ as long as we select an independent vector. Since there are $\dim(N_{\pi(1)} + N_{\pi(2)} + N_{\pi(3)})$ independent vectors in a basis of $N_{\pi(1)} + N_{\pi(2)} + N_{\pi(3)}$, then we can select

$$R_{\pi(3)} = \dim(N_{\pi(1)} + N_{\pi(2)} + N_{\pi(3)}) - \dim(N_{\pi(1)} + N_{\pi(2)})$$

independent vectors from $N_{\pi(3)}$ that are also independent of the $R_{\pi(1)} + R_{\pi(2)}$ vectors that we selected in the previous two steps.

4. We keep using the iterative procedure above for all the elements in π, and we end up with $\sum_{i=1}^{m} R_i$ vectors that are linearly independent.

This concludes the proof of Lemma 5.

A.4 Proof of Lemma 7

In this section, we leverage the result in Lemma 5 to prove Lemma 7. We start by noting that the rate region in (2.15) can be expressed as the following polyhedron

$$P_f := \left\{ R \in \mathbb{R}^{[m]} : R \geq \mathbf{0}, \sum_{i \in \mathcal{A}} R_i \leq f(\mathcal{A}),\ \forall\, \mathcal{A} \subseteq [m] \right\}, \tag{A.3}$$

where $f(\mathcal{A}) := \dim\left(\sum_{i \in \mathcal{A}} N_i\right)$. We now prove the following lemma, which states that this function $f(\cdot)$ is a *n*on-decreasing and *s*ubmodular function over subsets of $[m]$.

Lemma 29. *The set function*

$$f(\mathcal{A}) := \textit{dim}\left(\sum_{i \in \mathcal{A}} N_i\right),\ \forall \mathcal{A} \subseteq [m]$$

is a non-decreasing and submodular function.

Proof. Let $\mathcal{A} \subset \mathcal{B} \subseteq [m]$, then

$$f(\mathcal{B}) = \dim\left(\sum_{i \in \mathcal{B}} N_i\right) = \dim\left(\sum_{i \in \mathcal{A}} N_i + \sum_{j \in \mathcal{B} \setminus \mathcal{A}} N_j\right)$$

$$\geq \dim\left(\sum_{i\in\mathcal{A}} N_i\right) = f(\mathcal{A}),$$

which proves that the function $f(\cdot)$ is non-decreasing. For proving submodularity, consider two subsets $\mathcal{C}, \mathcal{D} \subseteq [m]$. Then, we have

$$\begin{aligned}
f(\mathcal{C}\cup\mathcal{D}) &= \dim\left(\sum_{i\in\mathcal{C}\cup\mathcal{D}} N_i\right) = \dim\left(\sum_{i\in\mathcal{C}} N_i + \sum_{j\in\mathcal{D}} N_j\right) \\
&= \dim\left(\sum_{i\in\mathcal{C}} N_i\right) + \dim\left(\sum_{j\in\mathcal{D}} N_j\right) \\
&\quad - \dim\left(\left(\sum_{i\in\mathcal{C}} N_i\right) \cap \left(\sum_{j\in\mathcal{D}} N_j\right)\right) \\
&\leq \dim\left(\sum_{i\in\mathcal{C}} N_i\right) + \dim\left(\sum_{j\in\mathcal{D}} N_j\right) \\
&\quad - \dim\left(\sum_{k\in\mathcal{C}\cap\mathcal{D}} N_k\right) \\
&= f(\mathcal{C}) + f(\mathcal{D}) - f(\mathcal{C}\cap\mathcal{D}),
\end{aligned}$$

which proves that the function $f(\cdot)$ is submodular. □

Since $f(\cdot)$ is a submodular set function, then the polyhedron defined in (A.3) is the polymatroid associated with $f(\cdot)$. Moreover, since $f(\cdot)$ is also non-decreasing, then the corner points of the polymatroid in (A.3) can be found as follows [Sch03, Corollary 44.3a]. Consider a permutation $\pi = \{\pi(1), \ldots, \pi(m)\}$ of $[m]$. Then, by letting $\mathcal{S}_\ell = \{\pi(1), \ldots, \pi(\ell)\}$ for $1 \leq \ell \leq m$, we get that the corner points of the polymatroid in (A.3) can be written as

$$R_{\pi(\ell)} = f(\mathcal{S}_\ell) - f(\mathcal{S}_{\ell-1}).$$

Note that by using $f(\mathcal{A}) = \dim\left(\sum_{i\in\mathcal{A}} N_i\right)$, the above corner points are precisely those in (2.14) in Lemma 5. Since each rate m-tuple $(R_1, R_2, \ldots, R_m)$, with $R_i, i \in [m]$ being defined in (2.14), can be securely achieved by our proposed scheme, it follows that the secure rate region in (2.15) can also be achieved by our scheme. This concludes the proof of Lemma 7.

A.5 Analysis of the dimension of $(V_1 \cap V_2 \cap V_3)$

From our analysis, we have obtained

$$\dim(V_1 \cap V_2 \cap V_3) \leq k + [k - M_{\cap\{i,j\}}]^+ + [k - M_{\cap\{\ell,\{i,j\}\}}]^+ + t - M_{\{1,2,3\}}. \tag{A.4}$$

We now further consider two cases.

Case 3A: There exists a pair $(i,j) \in [3]^2, i \neq j$, such that $M_{\cap\{i,j\}} \geq k$. In this case, with the permutation (i,j,ℓ), the expression in (A.4) becomes

$$\begin{aligned}\dim(V_1 \cap V_2 \cap V_3) &\leq k + [k - M_{\cap\{\ell,\{i,j\}\}}]^+ + t - M_{\{1,2,3\}} \\ &= t - M_{\{1,2,3\}} + \max\{2k - M_{\cap\{\ell,\{i,j\}\}}, k\}.\end{aligned}$$

From (2.23), this implies that

$$\begin{aligned}\dim(N_1 + N_2 + N_3) &\geq M_{\{1,2,3\}} - \max\{2k - M_{\cap\{\ell,\{i,j\}\}}, k\} \\ &= \min\left\{M_{\{1,2,3\}} - k, M_{\{\ell\}} + M_{\{i,j\}} - 2k\right\},\end{aligned}$$

where the last equality follows since $M_{\{1,2,3\}} = M_{\{i,j\}} + M_{\{\ell\}} - M_{\cap\{\ell,\{i,j\}\}}$. With this, the condition in (2.21) is satisfied.

Case 3B: We have $M_{\cap\{i,j\}} < k, \forall (i,j) \in [3]^2, i \neq j$. In this case, we compute $\dim(V_1 \cap V_2 \cap V_3)$ as follows: we first fill the positions of x indexed by $\mathcal{M}_1$ with k degrees of freedom, and then fill the positions of x indexed by $\mathcal{M}_2$ with $(k - M_{\cap\{1,2\}})$ degrees of freedom as before. Now, we may have fixed more than k positions of x corresponding to indexes in $\mathcal{M}_3$, which is not feasible. If that is the case, we backtrack (i.e., remove excess degrees of freedom) that we have used for filling positions of x indexed by $\mathcal{M}_2$. Thus,

1. If $M_{\cap\{3,\{1,2\}\}} \leq k$, then

   $$\dim(V_1 \cap V_2 \cap V_3) \leq t - M_{\{1,2,3\}} + k + (k - M_{\cap\{1,2\}\}}) + (k - M_{\cap\{3,\{1,2\}\}}).$$

   This, from (2.23), implies

   $$\dim(N_1 + N_2 + N_3) \geq M_{\{1\}} + M_{\{2\}} + M_{\{3\}} - 3k,$$

   which satisfies the condition in (2.21).

2. If $M_{\cap\{3,\{1,2\}\}} > k$, then

$$\begin{aligned}\dim(V_1 \cap V_2 \cap V_3) \leq t - M_{\{1,2,3\}} + k + (k - M_{\cap\{1,2\}}) \\ - \min\{k - M_{\cap\{1,2\}}, M_{\cap\{3,\{1,2\}\}} - k\}.\end{aligned}$$

This, from (2.23), implies

$$\begin{aligned}&\dim(N_1 + N_2 + N_3) \\ &\geq \min\left\{M_{\{1,2,3\}} - k, M_{\{1\}} + M_{\{2\}} + M_{\{3\}} - 3k\right\},\end{aligned}$$

which satisfies the condition in (2.21).

A.6 Proof of security: separable networks

In this section, we show that for any choice of G of size $|\mathcal{E}| \times M_{[m]}$ with $M_{[m]} \geq k$, there exists a $\tilde{V}$ such that $\tilde{V}$ is an MDS matrix (i.e., any k rows of $\tilde{V}$ are linearly independent) and

$$rk\left(\left[\begin{array}{cc} GM & G\tilde{V} \end{array}\right]\Big|_{\mathcal{Z}}\right) = rk\left(\left[\begin{array}{c} G\tilde{V} \end{array}\right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k. \tag{A.5}$$

We start by noting that

$$\begin{aligned}rk\left(\left[\begin{array}{c} G\tilde{V} \end{array}\right]\Big|_{\mathcal{Z}}\right) &= rk\left(G|_{\mathcal{Z}} \cdot \tilde{V}\right) \\ &\leq rk\left(\left[\begin{array}{cc} GM & G\tilde{V} \end{array}\right]\Big|_{\mathcal{Z}}\right) = rk\left(G|_{\mathcal{Z}} \cdot \left[\begin{array}{cc} M & \tilde{V} \end{array}\right]\right) \leq rk\left(G|_{\mathcal{Z}}\right).\end{aligned}$$

Thus, if we prove that, for all $|\mathcal{Z}| \leq k$,

$$rk\left(G|_{\mathcal{Z}} \cdot \tilde{V}\right) = rk\left(G|_{\mathcal{Z}}\right), \tag{A.6}$$

then we also show that (A.5) holds. In what follows, we formally prove that a $\tilde{V}$ such that $\tilde{V}$ is an MDS matrix that satisfies the condition in (A.6) for all $|\mathcal{Z}| \leq k$ can be constructed with a non-zero probability. Towards this end, we let $\hat{k} = rk\left(G|_{\mathcal{Z}}\right)$, where $\hat{k} \leq k$ since $|\mathcal{Z}| \leq k$. We have

$$\begin{aligned}&\Pr\left\{\left\{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\tilde{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k\right\} \cap \left\{\tilde{V} \text{ is MDS}\right\}\right\} \\ &\overset{(a)}{=} 1 - \Pr\left\{\left\{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\tilde{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k\right\}^c \cup \left\{\tilde{V} \text{ is not MDS}\right\}\right\}\end{aligned}$$

$$\overset{\text{(b)}}{\geq} 1 - \underbrace{\Pr\left\{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\tilde{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k\right\}^{c}}_{P_1} - \underbrace{\Pr\left\{\tilde{V} \text{ is not MDS}\right\}}_{P_2},$$

where: (i) the equality in (a) follows by using the De Morgan's laws, and (ii) the inequality in (b) follows since for two events A and B, we have $\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$. We now further upper bound the two probability terms P_1 and P_2. For P_1, we obtain

$$\begin{aligned}
P_1 &= \Pr\left\{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\tilde{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k\right\}^{c} \\
&\overset{\text{(c)}}{=} \Pr\left\{\left(\bigcap_{\mathcal{Z}:|\mathcal{Z}|\leq k} A_{\mathcal{Z}}\right)^{c}\right\} \overset{\text{(d)}}{=} \Pr\left\{\bigcup_{\mathcal{Z}:|\mathcal{Z}|\leq k} (A_{\mathcal{Z}})^{c}\right\} \overset{\text{(e)}}{\leq} \sum_{\mathcal{Z}:|\mathcal{Z}|\leq k} \Pr\left\{(A_{\mathcal{Z}})^{c}\right\} \\
&\overset{\text{(f)}}{\leq} \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^{k} \max_{\mathcal{Z}:|\mathcal{Z}|\leq k} \Pr\left\{(A_{\mathcal{Z}})^{c}\right\} = \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^{k} \max_{\mathcal{Z}:|\mathcal{Z}|\leq k} \left(1 - \Pr\left\{A_{\mathcal{Z}}\right\}\right) \\
&\overset{\text{(g)}}{\leq} \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^{k} \max_{\mathcal{Z}:|\mathcal{Z}|\leq k} \left(1 - \Pr\left\{\hat{A}_{\mathcal{Z}}\right\}\right) \\
&\overset{\text{(h)}}{=} \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^{k} \left(1 - \prod_{i=0}^{\hat{k}-1}\left(1 - \frac{q^{i}}{q^{\hat{k}}}\right)\right) \\
&\overset{\text{(i)}}{\leq} \left(\frac{\mathrm{e}|\mathcal{E}|}{k}\right)^{k} \left(1 - \left(1 - \frac{1}{q}\right)^{k}\right),
\end{aligned}$$

where: (i) the equality in (c) follows by defining, for a given $\mathcal{Z}$ such that $|\mathcal{Z}| \leq k$, the event

$$A_{\mathcal{Z}} = \left\{rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\tilde{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right)\right\},$$

(ii) the equality in (d) follows by using the De Morgan's laws; (iii) the inequality in (e) follows by using the union bound; (iv) the inequality in (f) follows since $\binom{n}{t} \leq \left(\frac{\mathrm{e}n}{t}\right)^{t}$; (v) the inequality in (g) follows by defining the event $\hat{A}_{\mathcal{Z}}$ as

$$\hat{A}_{\mathcal{Z}} = \left\{rk\left(\hat{G}\hat{V}\right) = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right)\right\},$$

where $\hat{G}$ is the matrix formed by the $\hat{k} = rk\left(\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}\right)$ independent rows of $\left[\begin{array}{c} G \end{array}\right]\Big|_{\mathcal{Z}}$, and $\hat{V}$ is formed by the first $\hat{k}$ columns of $\tilde{V}$. Thus, the inequality in (g) then follows since $\hat{A}_{\mathcal{Z}} \subseteq A_{\mathcal{Z}}$; (vi) the equality in (h) follows due to the following computation. We write

$$\hat{V} = \left[\begin{array}{cccc} v_1 & v_2 & \ldots & v_{\hat{k}} \end{array}\right] \implies \hat{G}\hat{V} = \left[\begin{array}{cccc} \hat{G}v_1 & \hat{G}v_2 & \ldots & \hat{G}v_{\hat{k}} \end{array}\right].$$

Note that the matrix $\hat{G}\hat{V}$ is of full rank (equal to $\hat{k}$) if the only solution to $\sum_{i=1}^{\hat{k}} c_i \hat{G} v_i = 0$ is $c_i = 0, \forall i \in [\hat{k}]$. Let $\hat{N}$ be the null space of $\hat{G}$, and $\hat{N}^{\perp}$ be the space such that $\hat{N}^{\perp} \cap \hat{N} = \emptyset$ and $\hat{N}^{\perp} \cup \hat{N} = \mathbb{F}_q^{M_{[m]}}$. Then, we can write each $v_i, i \in [\hat{k}]$, as the sum of its projection on $\hat{N}$ (say $v_i^{(a)}$) and the residual in $\hat{N}^{\perp}$ (say $v_i^{(b)}$). This implies that $\hat{G}\hat{V}$ is of full rank if the only solution to $\sum_{i=1}^{\hat{k}} c_i \hat{G} v_i^{(b)} = 0$ is $c_i = 0, \forall i \in [\hat{k}]$ (because $\hat{G} v_i^{(a)} = 0$). Since a random choice of v_i results in a random choice on $v_i^{(b)}$, then the probability of $\hat{G}\hat{V}$ being of full rank is equal to the probability that all the vectors $v_i^{(b)},\ i \in [\hat{k}]$ are mutually independent in $\hat{N}^{\perp}$. This probability, since $\dim(\hat{N}^{\perp}) = \hat{k}$, is equal to $\prod_{i=0}^{\hat{k}-1} \left(1 - \frac{q^i}{q^{\hat{k}}}\right)$; finally, (vii) the inequality in (i) follows since $i - \hat{k} \leq -1$ for all $i \in [0 : \hat{k} - 1]$ and $\hat{k} \leq k$.

For P_2, we obtain

$$
\begin{aligned}
P_2 = \Pr\left\{\tilde{V} \text{ is not MDS}\right\} &\overset{(j)}{=} \Pr\left\{\left(\bigcap_{\mathcal{S}:|\mathcal{S}|=k} A_s\right)^c\right\} \overset{(k)}{=} \Pr\left\{\bigcup_{\mathcal{S}:|\mathcal{S}|=k} (A_{\mathcal{S}})^c\right\} \\
&\overset{(\ell)}{=} \binom{M_{[m]}}{k} \Pr\left\{(A_{\mathcal{S}})^c\right\} = \binom{M_{[m]}}{k} \left(1 - \Pr\left\{A_{\mathcal{S}}\right\}\right) \overset{(m)}{=} \binom{M_{[m]}}{k} \left(1 - \prod_{i=0}^{k-1} \frac{q^k - q^i}{q^k}\right) \\
&\overset{(n)}{\leq} \binom{M_{[m]}}{k} \left(1 - \prod_{i=0}^{k-1} \left(1 - \frac{1}{q}\right)\right) = \binom{M_{[m]}}{k} \left(1 - \left(1 - \frac{1}{q}\right)^k\right),
\end{aligned}
$$

where: (i) the equality in (j) follows by defining, for a given $\mathcal{S}$ such that $|\mathcal{S}| = k$, the event

$$
A_{\mathcal{S}} = \left\{\tilde{V}|_{\mathcal{S}} \text{ is full rank}\right\},
$$

(ii) the equality in (k) follows by using the De Morgan's laws; (iii) the equality in (ℓ) follows by selecting uniformly at random all the subsets of k rows out of the $M_{[m]}$ rows, (iv) the equality in (m) follows by counting arguments to ensure that the k selected rows are all independent, and (v) the inequality in (n) follows since $i - k \leq -1$ for all $i \in [0 : k - 1]$.

Thus, we obtain

$$
\begin{aligned}
&\Pr\left\{\left\{rk\left(\left[\ G\ \right]\Big|_{\mathcal{Z}} \tilde{V}\right) = rk\left(\left[\ G\ \right]\Big|_{\mathcal{Z}}\right),\ \forall |\mathcal{Z}| \leq k\right\} \cap \left\{\tilde{V} \text{ is MDS}\right\}\right\} \\
&\geq 1 - \left(\frac{e|\mathcal{E}|}{k}\right)^k \left(1 - \left(1 - \frac{1}{q}\right)^k\right) - \binom{M_{[m]}}{k} \left(1 - \left(1 - \frac{1}{q}\right)^k\right) \\
&> 0,
\end{aligned}
$$

where the last inequality holds for sufficiently large values of q.

A.7 Proof of Theorem 19 and Corollary 21

We start by computing $R_{X_t|Z_1^N}$. Note that given a sequence of transmitted symbol Z_1^N there are two possible values of sequence of message symbols X_1^N which are $X_1^N = Z_1^N$ and $X_1^N = \tilde{Z}_1^N$, where $\tilde{Z}_t$ is $\alpha_t^{-1}(Z_t)$.

The posterior probability of $X_t = Z_t$ given Z_1^N i.e., $Pr(X_t = Z_t|Z_1^N)$ will be equal to $Pr(X_1^N = Z_1^N|Z_1^N) := p_Z$. We note that $p_Z = \frac{f(Z)}{f(Z)+f(\tilde{Z})}$, where $\tilde{Z} := [(\tilde{Z}_1)^T, (\tilde{Z}_2)^T, \ldots, (\tilde{Z}_N)^T]^T$. Then, $\mathbb{E}(X_t|Z_1^N) = p_Z Z_t + (1-p_Z)(\tilde{Z}_t)$. With this,

$$\begin{aligned}
R_{X_t|Z_1^N} &= \mathbb{E}_{X_t|Z_1^N}\left[\left(X_t - \mathbb{E}(X_t|Z_1^N)\right)\left(X_t - \mathbb{E}(X_t|Z_1^N)\right)^T\right] \\
&= p_Z(1-p_Z)^2(Z_t - \tilde{Z}_t)(Z_t - \tilde{Z}_t)^T \\
&\quad + (1-p_Z)p_Z^2(Z_t - \tilde{Z}_t)(Z_t - \tilde{Z}_t)^T \\
&= p_Z(1-p_Z)(Z_t - \tilde{Z}_t)(Z_t - \tilde{Z}_t)^T \\
D_E &= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} \text{tr}\left(R_{X_t|Z_1^N}\right) \\
&= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} \text{tr}\left(p_Z(1-p_Z)(Z_t - \tilde{Z}_t)(Z_t - \tilde{Z}_t)^T\right) \\
&= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} p_Z(1-p_Z)\text{tr}\left((Z_t - \tilde{Z}_t)(Z_t - \tilde{Z}_t)^T\right) \\
&= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} p_Z(1-p_Z)\|Z_t - \tilde{Z}_t\|^2 \\
&= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} \frac{f_X(Z)f_X(\tilde{Z})}{(f_X(Z)+f_X(\tilde{Z}))^2}\|Z_t - \tilde{Z}_t\|^2.
\end{aligned}$$

Now, Z_1^N is the transmitted symbols if $X_1^N = Z_1^N$ and key was zero or if $\{X_t = \tilde{Z}_t,\ \forall t \in [N]\}$ and key was one. So $f_Z(Z) = \frac{f_X(Z)+f_X(\tilde{Z})}{2}$. Thus D_E,

$$\begin{aligned}
&= \frac{1}{N}\mathbb{E}_Z \sum_{t=1}^{N} \frac{f_X(Z)f_X(\tilde{Z})}{(f_X(Z)+f_X(\tilde{Z}))^2}\|Z_t - \tilde{Z}_t\|^2 \\
&= \frac{1}{N}\int f_Z(Z) \sum_{t=1}^{N} \frac{f_X(Z)f_X(\tilde{Z})}{(f_X(Z)+f_X(\tilde{Z}))^2}\|Z_t - \tilde{Z}_t\|^2 dZ \\
&= \frac{1}{2N}\int \sum_{t=1}^{N} \frac{f_X(Z)f_X(\tilde{Z})}{f_X(Z)+f_X(\tilde{Z})}\|Z_t - \tilde{Z}_t\|^2\, dZ
\end{aligned}$$

$$= \frac{1}{2N}\mathbb{E}_X \sum_{t=1}^{N} \frac{f_X(\tilde{X})}{f_X(X) + f_X(\tilde{X})} \|Z_t - \tilde{Z}_t\|^2$$

$$= \frac{1}{2N}\mathbb{E}_X \sum_{t=1}^{N} \frac{f_X(\alpha^{-1}(X))}{f_X(X) + f_X(\alpha^{-1}(X))} \|X_t - \alpha_t^{-1}(X_t)\|^2,$$

which proves (4.8). Again, if we can choose S_t's, b_t's where $\alpha_t()$ is mirroring across planes given by $S_t x = b_t$ such that,

$$f_X(X) = f_X(\alpha^{-1}(X)),\ \forall X \in \mathbb{R}^{nN},$$

the distortion D_E becomes,

$$\begin{aligned} D_E &= \frac{1}{4N}\mathbb{E}_X \sum_{t=1}^{N} \|X_t - \alpha_t^{-1}(X_t)\|^2 \overset{(a)}{=} \frac{1}{N}\sum_{t=1}^{N} \mathbb{E}_{X_t}\|S_t X_t - b_t\|^2 \\ &= \frac{1}{N}\sum_{t=1}^{N} \text{tr}\left(S_t R_{X_t}(S_t)^T + (b_t - S_t\mu_{X_t})(b_t - S_t\mu_{X_t})^T\right), \end{aligned}$$

where (a) follows as $\alpha_t(.)$ is mirroring across plane given by $S_t x = b_t$, and thus $\alpha_t(x) = \alpha_t^{-1}(x) = (I - 2(S_t)^T S_t)X_t + 2(S_t)^T b_t$. This proves (4.11).

A.8 Proof for Theorem 27

Let the shared key K is $(K_1, K_2, \ldots, K_n)$ where all K_i's are i.i.d. and uniformly distributed in $\{0,1\}^3$. Let us also assume that $X = (X^{(1)}, X^{(2)}, \ldots, X^{(n)})$, where each $X^{(i)} \in \mathbb{R}$. Similar to the scheme for scalar case, we create a random vector $V = (V^{(1)}, \ldots, V^{(n)})$ where $V^{(i)} = (X^{(i)} - \mu^{(i)})/\sqrt{\Sigma_{ii}}$, and encode $V^{(i)}$ using key K_i as in the case of a scalar for all $i \in [n]$. Thus, the distortion D_W will be,

$$\begin{aligned} D_W &= \min_Z \text{tr}(R_{X|Z}) = \min_Z \sum_{i=1}^{n} \text{Var}(X^{(i)}|Z) \\ &= \min_Z \sum_{i=1}^{n} (\Sigma_{ii})\text{Var}(V^{(i)}|Z) = \sum_{i=1}^{n} (\Sigma_{ii}) \min_Z \text{Var}(V^{(i)}|Z) \\ &= \sum_{i=1}^{n} (\Sigma_{ii}) \min_{Z^{(i)}} \text{Var}(V^{(i)}|Z^{(i)}) = c\sum_{i=1}^{n} (\Sigma_{ii}) = c\ \text{tr}(\Sigma), \end{aligned}$$

where $c = 0.9998$. And since $\text{tr}(\Sigma)$ is the expected distortion even when the adversary has no observations, and as we can not beat this by (4.6), this is optimal.

A.9 Proof for Theorem 28

Distortion at the adversary's end. Based on the coding scheme we can see that the adversary get $BU_t + w_t$ by just subtracting AZ_t from Z_{t+1} for $t \in [1 : N-1]$. So the adversary's information is given by following set:

$$\begin{aligned} E_{\text{info}} &= \{Z_1, BU_t + w_t,\ t \in [1 : N-1]\} \\ &= \{f(X_1, K), BU_t + w_t,\ t \in [1 : N-1]\}. \end{aligned}$$

Thus, $D(t, Z_1^N) = D(t, E_{\text{info}}) = \text{tr}(R_{X_t|E_{\text{info}}})$. Let's first compute $D(t = 1, Z_1^N)$,

$$D(t = 1, Z_1^N) = \text{tr}(R_{X_1|E_{\text{info}}}) \overset{(a)}{=} \text{tr}(R_{X_1|f(X_1,K)}) \overset{(b)}{=} c\,\text{tr}(\Sigma),$$

where (a) is because U_t and w_t are independent on X_t and (b) is due to the encoding used in Theorem 27 with $c = 0.9998$.

Now, for other time instances we can use induction to prove that we will have worst case distortion at least $\text{tr}(\Sigma)$.

$$\begin{aligned} D(t+1, Z_1^N) &= \text{tr}(R_{X_{t+1}|E_{\text{info}}}) = \text{tr}(R_{(AX_t+BU_t+w_t)|E_{\text{info}}}) \\ &= \text{tr}(R_{(AX_t)|E_{\text{info}}}) = \text{tr}(AR_{X_t|E_{\text{info}}}A^T) = \text{tr}(A^TAR_{X_t|E_{\text{info}}}) \\ &\overset{(a)}{=} \text{tr}(V\Lambda V^T R_{X_t|E_{\text{info}}}) = \text{tr}(\Lambda V^T R_{X_t|E_{\text{info}}}V) \\ &\overset{(b)}{=} \sum_{i=1}^{n} \lambda_i d_i(V^T R_{X_t|E_{\text{info}}}V) \overset{(c)}{\geq} \sum_{i\in[n]} d_i(V^T R_{X_t|E_{\text{info}}}V) \\ &\overset{(d)}{=} \sum_{i\in[n]} \nu_i(V^T R_{X_t|E_{\text{info}}}V) = \sum_{i\in[n]} \nu_i(R_{X_t|E_{\text{info}}}) \\ &= \text{tr}(R_{X_t|E_{\text{info}}}) \overset{(e)}{\geq} c\,\text{tr}(\Sigma), \end{aligned}$$

where in (a), we do eigenvalue decomposition of A^TA which is a positive definite matrix and thus will have non negative eigenvalues; in (b) $d_i(V^T R_{X_t|E_{\text{info}}}V)$ is the i-th diagonal entry of

$V^T R_{X_t|E_{\text{info}}} V$; (c) is true because $V^T R_{X_t|E_{\text{info}}} V$ is a positive definite matrix and all the diagonal entries of a positive semi definite matrix are non-negative and because of our assumption that singular values of A, i.e. the square root of eigenvalues of $A^T A$ are all more than one; (d) is because summation of eigenvalues is equal to the sum of all the diagonal entries for any square matrix, where $\nu_i(V^T R_{X_t|E_{\text{info}}} V)$ is the i-th eigenvalue of $V^T R_{X_t|E_{\text{info}}} V$; (e) follows by the induction.

REFERENCES

[ACF16] G. K. Agarwal, M. Cardone, and C. Fragouli. "Coding across unicast sessions can increase the secure message capacity." In *2016 IEEE Int. Symp. on Inf. Theory*, pp. 2134–2138, Jul 2016.

[ACF17] Gaurav Kumar Agarwal, Martina Cardone, and Christina Fragouli. "Secure Network Coding for Multiple Unicast: On the Case of Single Source." In Junji Shikata, editor, *Information Theoretic Security*, pp. 188–207, Cham, 2017. Springer International Publishing.

[ACL00] R. Ahlswede, Ning Cai, S. Y. R. Li, and R. W. Yeung. "Network information flow." *IEEE Transactions on Information Theory*, **46**(4):1204–1216, Jul 2000.

[AHM07] Daniel Andrén, Lars Hellström, and Klas Markström. "On the complexity of matrix reduction over finite fields." *Advances in applied mathematics*, **39**(4):428–452, 2007.

[All15] NGMN Alliance. "5G white paper." *Next generation mobile networks, white paper*, pp. 1–125, 2015.

[Aro19] Jacob Aron. "IBM unveils its first commercial quantum computer.", 2019. https://www.newscientist.com/article/2189909-ibm-unveils-its-first-commercial-quantum-computer/ Last Accessed: May 9th, 2019.

[BN05] Kapil Bhattad, Krishna R Narayanan, et al. "Weakly secure network coding." *NetCod, Apr*, **104**, 2005.

[CDH16] J. Corts, G. E. Dullerud, S. Han, J. L. Ny, S. Mitra, and G. J. Pappas. "Differential privacy in control and network systems." In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 4252–4272, Dec 2016.

[CHK13a] T. Cui, T. Ho, and J. Kliewer. "On Secure Network Coding With Nonuniform or Restricted Wiretap Sets." *IEEE Transactions on Information Theory*, **59**(1):166–176, Jan 2013.

[CHK13b] T. Cui, T. Ho, and J. Kliewer. "On Secure Network Coding With Nonuniform or Restricted Wiretap Sets." *IEEE Transactions on Information Theory*, **59**(1):166–176, January 2013.

[CLR09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009.

[CPD14] L. Czap, V. M. Prabhakaran, S. Diggavi, and C. Fragouli. "Triangle network secrecy." In *2014 IEEE International Symposium on Information Theory*, pp. 781–785, June 2014.

[CPF11] L. Czap, V.M. Prabhakaran, C. Fragouli, and S. Diggavi. "Secret message capacity of erasure broadcast channels with feedback." In *IEEE Inf. Theory Workshop (ITW)*, pp. 65–69, 2011.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, New York, NY, USA, 2006.

[CY02] Ning Cai and R. W. Yeung. "Secure network coding." In *Proceedings IEEE International Symposium on Information Theory (ISIT),*, pp. 323–, July 2002.

[CY07] N. Cai and R. W. Yeung. "A Security Condition for Multi-Source Linear Network Coding." In *2007 IEEE International Symposium on Information Theory*, pp. 561–565, June 2007.

[DCN09] Jing Dong, Reza Curtmola, and Cristina Nita-Rotaru. "Secure network coding for wireless mesh networks: Threats, challenges, and directions." *Computer Communications*, **32**(17):1790–1801, 2009.

[DOM19] DOMO.COM. "Data Never Sleeps 6.0.", 2019. https://www.domo.com/learn/data-never-sleeps-6 Last Accessed: May 9th, 2019.

[ECF18] Y. H. Ezzeldin, M. Cardone, C. Fragouli, and G. Caire. "Gaussian 1-2-1 Networks: Capacity Results for mmWave Communications." *arXiv:1801.02553*, January 2018.

[ES07] Salim Y El Rouayheb and Emina Soljanin. "On wiretap networks II." In *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, pp. 551–555. IEEE, 2007.

[FMS04] Jon Feldman, Tal Malkin, C Stein, and RA Servedio. "On the capacity of secure network coding." In *Proc. 42nd Annual Allerton Conference on Communication, Control, and Computing*, pp. 63–68, 2004.

[HLK04] Tracey Ho, Ben Leong, Ralf Koetter, Muriel Médard, Michelle Effros, and David R Karger. "Byzantine modification detection in multicast networks using randomized network coding." In *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, p. 144. IEEE, 2004.

[JLK07] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, and M. Medard. "Resilient network coding in the presence of Byzantine adversaries." In *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pp. 616–624, May 2007.

[JSC05] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. M. G. M. Tolhuizen. "Polynomial time algorithms for multicast network code construction." *IEEE Transactions on Information Theory*, **51**(6):1973–1982, June 2005.

[KM03] R. Koetter and M. Medard. "An algebraic approach to network coding." *IEEE/ACM Transactions on Networking*, **11**(5):782–795, October 2003.

[KM12] Vijay Kumar and Nathan Michael. "Opportunities and challenges with autonomous micro aerial vehicles." *The International Journal of Robotics Research*, **31**(11):1279–1291, 2012.

[KOK17] Kaoru Kurosawa, Hiroyuki Ohta, and Kenji Kakuta. "How to make a linear network code (strongly) secure." *Designs, Codes and Cryptography*, **82**(3):559–582, 2017.

[KSK09] Azadeh Khaleghi, Danilo Silva, and Frank R Kschischang. "Subspace codes." In *IMA International Conference on Cryptography and Coding*, pp. 1–21. Springer, 2009.

[KTT09] O. Kosut, L. Tong, and D. Tse. "Nonlinear network coding is necessary to combat general Byzantine attacks." In *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 593–599, Sept 2009.

[KTW14] S. Kamath, D. N. C. Tse, and C. C. Wang. "Two-unicast is hard." In *IEEE International Symposium on Information Theory (ISIT)*, pp. 2147–2151, June 2014.

[LLE13] E. J. Lee, M. Langberg, and M. Effros. "Outer bounds and a functional study of the edge removal problem." In *2013 IEEE Information Theory Workshop (ITW)*, pp. 1–5, Sep. 2013.

[Mau93] U. M. Maurer. "Secret key agreement by public discussion from common information." *IEEE Trans. Inf. Theory,*, **39**(3):733–742, 1993.

[MMS13] Waseem A Malik, Nuno C Martins, and Ananthram Swami. "LQ control under security constraints." In *Control of Cyber-Physical Systems*, pp. 101–120. Springer, 2013.

[MS11] A. Mukherjee and A. L. Swindlehurst. "Robust Beamforming for Security in MIMO Wiretap Channels With Imperfect CSI." *IEEE Transactions on Signal Processing*, **59**(1):351–361, Jan 2011.

[MSC08] A. Mills, B. Smith, T. C. Clancy, E. Soljanin, and S. Vishwanath. "On secure communication over wireless erasure networks." In *2008 IEEE International Symposium on Information Theory*, pp. 161–165, July 2008.

[PCF14] Athanasios Papadopoulos, László Czap, and Christina Fragouli. "Secret message capacity of a line network." *CoRR*, **abs/1407.1922**, 2014.

[PCF15] Athanasios Papadopoulos, Laszlo Czap, and Christina Fragouli. "LP formulations for secrecy over erasure networks with feedback." In *2015 IEEE Int. Symp. on Inf. Theory*, pp. 954–958, June 2015.

[RW09] Aditya Ramamoorthy and Richard D Wesel. "The single source two terminal network with network coding." *arXiv:0908.2847*, August 2009.

[SC14] Curt Schieler and Paul Cuff. "Rate-distortion theory for secrecy systems." *IEEE Trans. Inf. Theory*, **60**(12):7584–7605, 2014.

[SCA16] I. Safaka, L. Czap, K. Argyraki, and C. Fragouli. "Creating Secrets Out of Packet Erasures." *IEEE Transactions on Information Forensics and Security*, **11**(6):1177–1191, June 2016.

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume B. Springer Science & Business Media, 2003.

[Sha49] Claude E Shannon. "Communication theory of secrecy systems." *Bell Labs Technical Journal*, **28**(4):656–715, 1949.

[SK09] Danilo Silva and Frank R Kschischang. "Universal weakly secure network coding." In *Networking and Information Theory, 2009. ITW 2009. IEEE Information Theory Workshop on*, pp. 281–285. IEEE, 2009.

[STA19] STATISTA.COM. "Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions).", 2019. https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/ Last Accessed: May 9th, 2019.

[TGP16] Anastasios Tsiamis, Konstantinos Gatsis, and George J. Pappas. "State Estimation with Secrecy against Eavesdroppers." *CoRR*, **abs/1612.04942**, 2016.

[TGP17] Anastasios Tsiamis, Konstantinos Gatsis, and George J. Pappas. "State-Secrecy Codes for Networked Linear Systems." *CoRR*, **abs/1709.04530**, 2017.

[TSS17] T. Tanaka, M. Skoglund, H. Sandberg, and K. H. Johansson. "Directed information and privacy loss in cloud-based control." In *2017 American Control Conference (ACC)*, pp. 1666–1672, May 2017.

[Vai89] P. M. Vaidya. "Speeding-up linear programming using fast matrix multiplication." In *30th Annual Symposium on Foundations of Computer Science*, pp. 332–337, Oct 1989.

[WJO16] M. Wiese, K. H. Johansson, T. J. Oechtering, P. Papadimitratos, H. Sandberg, and M. Skoglund. "Uncertain wiretap channels and secure estimation." In *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 2004–2008, July 2016.

[WYG10] Y. Wei, Z. Yu, and Y. Guan. "Efficient Weakly-Secure Network Coding Schemes against Wiretapping Attacks." In *2010 IEEE International Symposium on Network Coding (NetCod)*, pp. 1–6, June 2010.

[Wyn75] A. D. Wyner. "The wire-tap channel." *The Bell System Technical Journal*, **54**(8):1355–1387, 1975.

[Yam88] Hirosuke Yamamoto. "A rate-distortion problem for a communication system with a secondary decoder to be hindered." *IEEE Trans. Inf. Theory*, **34**(4):835–842, 1988.