2022-01

Contributions to Information Theoretic Multiterminal Secret Key Agreement

Poostindouz, Alireza

UNIVERSITY OF CALGARY

Contributions to Information Theoretic Multiterminal Secret Key Agreement

by

Alireza Poostindouz

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN COMPUTER SCIENCE

CALGARY, ALBERTA

JANUARY, 2022

Abstract

A multiterminal secret key agreement (SKA) protocol is used to establish a shared secret key among a group of terminals. We study SKA protocols with information-theoretic security. In the source model of SKA, each terminal can sample from a correlated random variable. In the channel model of SKA, terminals instead are connected through an underlying noisy channel that is used for distributing the correlated variables. The terminals arrive at a shared secret key by establishing correlation (as per the presumed source/channel model) and communicating over a noiseless authenticated public channel. In the general models of SKA, it is assumed that terminals' variables are partially leaked to the adversary, Eve, in the form of a random variable which we call Eve's wiretap side information. Eve has unlimited computational power and has read access to all public communication messages. The key rate of an SKA protocol is given by the key length divided by the terminals' variables length, and the maximum possible key rate calculated for an SKA model is called the wiretap secret key (WSK) capacity of that model. Finding a general expression for the WSK capacity continues to be one of the hardest open problems within the context of information-theoretic key agreement.

Our contributions include proving the WSK capacity and proposing capacity achieving SKA protocols for *the wiretapped PIN, Tree-PIN, and Polytree-PIN models*, that are special multiterminal SKA models of interest in practice. Also, we introduce a new channel model of SKA that we call *the transceiver model* for which we prove multiple upper and lower bounds on key capacity under various assumptions. Furthermore, we note that traditionally the key capacity was studied and calculated for SKA models, while in the actual implementation of SKA protocols, the achievable key length as a function of terminals' variables length is needed. Compared to calculating WSK capacity, finding the key length requires different information-theoretic techniques for evaluating the protocols. We prove finite-length upper and lower bounds on the maximum achievable key length for some of the models that we have considered. In the concluding sections, we outline directions for future research.

Acknowledgements

Nietzsche said,"He whose life has a why can bear almost any how."

For me, Parisa is the why.

To Parisa

Table of Contents

List of Figures and Illustrations

List of Tables

List of Symbols, Notations, and Abbreviations

Abbreviation	Meaning
RV	Random Variable
Enc	Encoder
Dec	Decoder
DMS	Discrete Memoryless Source Model
DMC	Discrete Memoryless Channel Model
BSC	Binary Symmetric Channel
BEC	Binary Erasure Channel
BMS	Binary-input Memoryless Symmetric Channel
IID	Independent and Identically Distributed
INID	Independent but Not Identically Distributed
SK	Secret Key
SKA	Secret Key Agreement
WSK	Wiretapped Secret Key
PK	Private Key
IR	Information Reconciliation
PA	Privacy Amplification
CR	Common Randomness
CO	Communication for Omniscience
LP	Linear Programming
PIN	Pairwise Independent Network
LLN	Law of Large numbers
LHL	Leftover Hash Lemma
UHF	Universal Hash Function

Symbol/Notation	Meaning
$\mathbb{R}$	The set of all real numbers.
$\mathbb{N}$	The set of all natural numbers.
$\mathcal{A}$	A set $\mathcal{A}$.
$\mathcal{A} \cup \mathcal{B}$	The union of sets $\mathcal{A}$ and $\mathcal{B}$.
$\mathcal{A} \cap \mathcal{B}$	The intersection of sets $\mathcal{A}$ and $\mathcal{B}$.
$\varnothing$	Empty set.

$\lvert \mathcal{A} \rvert$	Cardinality of set $\mathcal{A}$.
$[m]$	The set of natural numbers from 1 to m, $\{1, \ldots, m\}$.
$\mathcal{A} \backslash \mathcal{B}$	The set of all elements in set $\mathcal{A}$ that are not in set $\mathcal{B}$.
$\mathcal{B}^c$	The compliment of set B, i.e., $\mathcal{A} \backslash \mathcal{B}$ if $\mathcal{A}$ is the universal set.
$\Pr\{\cdot\}$	Probability function.
X	Random variable X.
$\mathcal{X}$	The alphabet of the random variable X.
$\mathrm{supp}(X)$	The support of the random variable X.
P_X	The probability distribution of X.
x	A sample from the random variable X.
$P_X(x)$	The probability of realization x, i.e., $\Pr\{X = x\}$.
$X \sim P_X$	Random variable X has distribution P_X.
$\mathrm{Bern}(q)$	Bernoulli distribution with success probability q.
$\mathbb{E}\{X\}$	The expectation value of X.
$\mathbb{E}_{P_X}\{X\}$	The expectation value of X.
$\mathbb{Var}\{X\}$	The variance of X.
$\mathbb{Var}_{P_X}\{X\}$	The variance of X.
P_{XY}	The joint distribution of X and Y.
$P_{X\mid Y}$	The conditional distribution of X given Y.
$\mathbf{V}$	A random vector $\mathbf{V}$.
$\mathcal{X}^n$	The n-th Cartesian product of alphabets $\mathcal{X} \times \mathcal{X} \times \cdots \times \mathcal{X}$.
$X^n = (X_1, \ldots, X_n)$	A sequence of random variables X_j for $j \in [n]$.
$x^n = (x_1, \ldots, x_n)$	A realization of random vector X^n.
P_{X^n}	The joint probability distribution of X^n
$V_{[m]}$	A vector $V_{[m]} = (V_1, \ldots, V_m)$.
$V_{\mathcal{A}}$	A vector $V_{\mathcal{A}} = (V_j \mid \forall j \in \mathcal{A})$.
$\mathrm{sum}(R)$	The sum of all elements of a real vector R.
$X - Y - Z$	A Markov chain relation.
$H(X)$	The Shannon entropy of X.
$H_{\min}(X)$	The min-entropy of X.
$H(X\mid Y)$	The conditional entropy of X given Y.
$I(X;Y)$	The mutual information between X and Y.
$I(X;Y\mid Z)$	The conditional mutual information between X and Y given Z.
$\mathbf{SD}(X, Y)$	The statistical distance between X and Y.
$G = (\mathcal{M}, \mathcal{E})$	A graph with set of vertexes $\mathcal{M}$ and set of edges $\mathcal{E}$.
$G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$	The subgraph of a given $G = (\mathcal{M}, \mathcal{E})$ with the smallest number of edges that spans all vertexes of $\mathcal{A} \subseteq \mathcal{M}$.
$\Gamma(j)$	The set of all vertexes connected to the vertex j.
$A \oplus B$	The bitwise XOR of binary strings (vectors) A and B.
$\mathrm{length}(A)$	The length of a binary string (vector) A.

Epigraph

We may have knowledge of the past but cannot control it; we may control the future but have no knowledge of it.

- Claude Shannon,

"Coding Theorems for a Discrete Source With a Fidelity Criterion,"

IRE International Convention Records, volume 7, pp. 142–163, 1959.

Chapter 1

Introduction

1.1 Why Information-theoretic Security?

How can two distant parties establish a secure communication link? Suppose Alice wants to send a sensitive message to Bob over an insecure communication channel. Alice transforms the message into a ciphertext, using a designed code called *encryption code*, sends the ciphertext over the insecure channel to Bob, and then Bob recovers the original message from the ciphertext using a code called *decryption code*. A passive adversary, Eve, whose objective is to learn Alice's message, can observe the ciphertext as well. The encryption and decryption codes are *information-theoretically secure* if it is proved theoretically that Eve cannot break the cipher, even by using unlimited computational power.

The idea of information-theoretic security (and theory of cryptography as well) originates from the seminal work of Shannon, where he proved that a secure communication link can be established between Alice and Bob if they share a secret key which is perfectly concealed from Eve [1]. The encryption and decryption codes that require the same shared secret key are referred to as *symmetric-key* cryptosystems. Shannon's result reduces the problem of secure communication to a problem of secret key agreement (SKA), as the security of any symmetric-key cryptosystem relies on the secrecy of the key which is shared in advance.

Diffie and Hellman proposed a practical solution to the key agreement problem, by introducing their pioneering secret key agreement protocol [2], and essentially starting the revolution of modern (*asymmetric-key*) cryptography. In their work and other works within the setting of asymmetric-key cryptography, the security definition is relaxed from information-theoretic security to *computational security*. That is, instead of proving the theoretical impossibility of breaking the cryptographic protocol, it is proved that breaking the protocol is computationally infeasible. However, such proofs rely on "an unproven computational difficulty of solving a certain mathematical problem." Currently, the entire cryptographic infrastructure of the Internet is based on primitives that are only computationally secure.

Unfortunately, quantum computers have the potential to break many of the ubiquitous cryptographic algorithms [3], including the Diffie-Hellman SKA [2], its multiterminal variants [4], and the RSA asymmetric-key cryptosystem [5]. Therefore, adversaries who are currently recording any information that is encrypted by today's cryptography standards can decrypt the information, when they have access to a working quantum computer. This is called the "store now, decrypt later" attack. Following recent rapid advancements in quantum technologies [6–9], standardization efforts [10, 11] and research on novel approaches to quantum-resistant cryptography have been escalated – see e.g., [12–15]. One approach, referred to as *post-quantum* cryptography, is to replace current standard cryptographic algorithms with new ones that are proven to be computationally secure. These new algorithms are designed based on mathematical problems that are believed to be computationally hard even for a quantum computer. This direction, however, still suffers from the same peril of not providing an information-theoretical security guarantee.

Alternatively, it is possible to prove information-theoretical security of primitives that utilize private random observations performed by legitimate parties. One limitation of such models is the assumption that random observations are only partially leaked to the adversary. This approach is especially most suitable for wireless network environments [16–18]. We focus on this latter approach.

1.2 Information-theoretic Secret Key Agreement

This thesis is concerned with the information-theoretic treatment of multiterminal secret key agreement by public discussion [19–23]. Suppose there are m terminals denoted by $\mathcal{M} = \{1, 2, , \ldots, m\}$, where a subset $\mathcal{A} \subseteq \mathcal{M}$ of terminals want to agree on a shared secret key. Terminals have unlimited access to a free, authenticated, noiseless public communication channel, where each message that is sent to the public channel can be observed by all terminals and the adversary, Eve. Terminals also establish statistical correlation. We consider two general categories of SKA models:

(i) **Source Model.** Each terminal has access to a different discrete random variable (RV). These variables are correlated and are partially wiretapped by Eve. The wiretap side information of Eve is modeled by a random variable Z, that is correlated with terminals' variables. Terminals and Eve, sample from their RV's for n times, and then terminals engage in a possibly interactive public discussion. The joint probability distribution of Eve's side information and terminals' variables is known publicly.

(ii) **Channel Model.** Terminals are connected to each other through a noisy multi-input multi-output channel. Some/all terminals can send input symbols and some/all terminals will observe the noisy channel outputs. Terminals use the public communication channel, before, in between, and after each symbol transmission over the noisy channel. Terminals use the noisy channel for n times. Eve's wiretap side information is modeled by an output RV of the noisy channel which we denote by Z. The conditional probability distribution of the underlying noisy channel is known publicly.

At the end of the SKA protocol, terminals of $\mathcal{A}$ compute their estimation of the final key. The key should be the same for all terminals in $\mathcal{A}$, and Eve should obtain (almost) no information about the key. The key may or may not be concealed from the *helper* terminals that are not in subset $\mathcal{A}$. The *key rate* is the key length divided by n, and the ***key capacity*** of any model is defined as the largest asymptotic ($n \to \infty$) achievable key rate.

By imposing specific assumptions on a model, special-case key capacities can be defined. For example, with regard to Eve's side information, three notions of key capacity are defined [21]. If Eve is not wiretapping and has no side information (Z = constant,) then the key capacity is called **secret key (SK) capacity**. For this case, the model is called non-wiretapped. If a group of helper terminals are compromised (wiretapped by Eve) then, Z is equal to the collection of all observations made by the compromised terminals; and we assume that (i) the group is known, and (ii) Z (the observations of compromised terminals) is known by all other terminals. For this case, the key capacity is called **private key (PK) capacity**. In the most general sense, if Eve has access to some wiretap side information ($Z \neq$ constant) which is not known by the terminals, then the key capacity is called the **wiretap secret key (WSK) capacity**. For this case, the model is called wiretapped.

> Finding a general expression for the WSK capacity, even for the case of two-party SKA ($m = 2$,) continues to be an important open problem.

See Chapter 2 which gives a more comprehensive review of multiterminal information-theoretic SKA. Next section provides an outline for each subsequent chapter of this thesis.

1.3   Thesis Scope and Contributions

Chapters 3 to 6, each study a specific problem within the context of information-theoretic SKA:

- Chapter 3 considers the problem of key agreement in the two-party wiretapped source model when public communication is one-way (from terminal 1 to terminal 2.)

- Chapter 4 studies the problem of finding the WSK capacity of a special class of multiterminal wiretapped source models.

- Chapter 5 introduces and analyses a new channel model for secret key agreement.

- Chapter 6 investigates whether a certain SKA approach is WSK capacity achieving for a special class of multiterminal wiretapped channel models.

These main technical chapters (summarized below) are closely related to each other, but for better readability, they are written in a way to be mostly self-contained. That is, each chapter starts with a detailed literature review of works related to the particular problem the chapter addresses; and in each chapter, we give a thorough explanation of the definitions, particular new notions and notations, specific SKA model, assumptions, and limitations.

Chapter 3

Consider the two-party source model of SKA. Alice (terminal 1,) Bob (terminal 2,) and Eve have access to n independent and identically distributed (IID) correlated random variables (X^n, Y^n, Z^n); respectively. Alice is allowed to send Bob only a single public message $F = F(X^n)$ that is computed based on her initial random variable. This model is called the *"two-party model of one-way secret key agreement (OW-SKA)"*. The key capacity of OW-SKA model (denoted by $C_{WSK}^{\rightarrow}$) is known [20]; however, for implementing OW-SKA protocols in practice we need to find the largest achievable *key length* as a function of n, which here we denote by $S^{\rightarrow}$. Let $o(g(n)) = \{f(n)|\ \lim_{n\to\infty} f(n)/g(n) = 0\}$. Previous capacity results imply that

$$S^{\rightarrow} = nC_{WSK}^{\rightarrow} - o(n).$$

Our objective in Chapter 3 is to find more accurate finite-length approximations of $S^{\rightarrow}$. To this end, we prove a new finite-length upper bound, and multiple finite-length lower bounds on $S^{\rightarrow}$. For proving the finite-length upper bound we use the information spectrum methods of [24, 25] and introduce a new entropy called the sup-spectral entropy. Then, we utilize the new spectral entropy and the converse techniques of [26] to prove a general upper bound on $S^{\rightarrow}$ for the case when (X^n, Y^n, Z^n) are not necessarily IID. Our new lower bounds are all in

the form of

$$S^{\rightarrow} \geqslant n C^{\rightarrow}_{WSK} - \sqrt{n} G - o(\sqrt{n}),$$

where G (which is different in each lower bound) is a function of the joint probability distribution of the model P_{XYZ}. The lower bounds are proved by analyzing two new OW-SKA protocols we propose: $\boldsymbol{\Pi}_{\text{HH}}$ and $\boldsymbol{\Pi}_{\text{PH}}$. Both of these protocols follow the same approach of performing SKA through 'information reconciliation followed by privacy amplification' [27]. In $\boldsymbol{\Pi}_{\text{HH}}$, reconciliation is implemented by Universal Hashing [28] whereas in $\boldsymbol{\Pi}_{\text{PH}}$, reconciliation is implemented by Polar Coding [29]. Privacy amplification in both protocols is designed based on universal hashing. To prove the achievable finite key length of these SKA protocols, we prove generalized variants of the Leftover Hash Lemma [30]. A by-product of our analysis is that we show $\boldsymbol{\Pi}_{\text{HH}}$ achieves the WSK capacity of the general source model [31] when (X^n, Y^n, Z^n) are independent (over n) but not necessarily IID, and satisfy Markov relation $X^n - Y^n - Z^n$ (see Definition 2.6.) We finish the chapter by comparing our proposed OW-SKA protocols and previous protocols with respect to their finite key rate, public communication costs, and computational complexity.

The contributions of Chapter 3 are also presented in the following papers:

- S. Sharifian, A. Poostindouz, and R. Safavi-Naini, "A capacity-achieving one-way key agreement with improved finite blocklength analysis," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, Oct. 2020, pp. 407–411, Copyright© 2020 IEICE. [Online]. Available: https://ieeexplore.ieee.org/document/9366148

- A. Poostindouz and R. Safavi-Naini, "Second-order asymptotics for one-way secret key agreement," in *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Jul. 2021, pp. 1254–1259. [Online]. Available: https://ieeexplore.ieee.org/document/9518202/

Figure 1.1: A simple wiretapped Tree-PIN. Here $\mathcal{M} = \{1, 2, 3\}$, $\mathcal{E} = \{e_{12}, e_{23}\}$, $X_1 = V_{12}$, $X_2 = (V_{21}, V_{23})$, $X_3 = V_{32}$, and Eve's side information is $Z = (Z_{12}, Z_{23})$. Both Markov relations $V_{12} - V_{21} - Z_{12}$ and $V_{23} - V_{32} - Z_{23}$ hold, that is, the source model distribution is $P_{X_{\mathcal{M}}Z} = P_{V_{12}} P_{V_{21}|V_{12}} P_{Z_{12}|V_{21}} P_{V_{23}} P_{V_{32}|V_{23}} P_{Z_{23}|V_{32}}$.

Chapter 4

In Chapter 4, we introduce the *"wiretapped Pairwise Independent Network (PIN)"* model with independent leakage. Let $\mathcal{M}$ denote the terminal set and $\mathcal{A}$ denote the subset of terminals that seek to agree on the final secret key. Suppose an undirected graph $G = (\mathcal{M}, \mathcal{E})$ is given. The RV of terminal $j \in \mathcal{M}$ is of the form $X_j = (V_{ji}|\ e_{ji} \in \mathcal{E})$ and with respect to each $e_{ij} = e_{ji} \in \mathcal{E}$, the Markov relation $V_{ij} - V_{ji} - Z_{ij}$ holds. Eve's side information is the collection of all wiretapped components Z_{ij}, that is $Z = (Z_{ij}|\ e_{ij} \in \mathcal{E})$. All triplets of RV's (V_{ij}, V_{ji}, Z_{ij}) are mutually independent. If the graph G is a tree then the model is called wiretapped *"Tree-PIN."* Figure 1.1 depicts a simple wiretapped Tree-PIN.

We prove the WSK capacity of wiretapped Tree-PIN for any arbitrary $\mathcal{A} \subseteq \mathcal{M}$, and show that it is equal to the key capacity of the case when Eve's side information Z^n is known. Let $I(V_1; V_2|V_3)$ denote the conditional mutual information between V_1 and V_2 given V_3 – see Definition 2.13. For the special case when $\mathcal{A} = \mathcal{M}$, the capacity is

$$C_{WSK} = \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}).$$

We propose an SKA protocol that achieves this WSK capacity and show that it uses less public communication bits per sample than the SKA protocol of [21].

We then prove a finite-length upper bound and multiple finite-length lower bounds for the largest achievable key length that can be generated for wiretapped Tree-PIN. Moreover, we extend our Tree-PIN model to three more general scenarios and prove the corresponding WSK capacity of those extended models.

The contributions of Chapter 4 appear also in the following papers:

- A. Poostindouz and R. Safavi-Naini, "Wiretap secret key capacity of Tree-PIN," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Jul. 2019, pp. 315–319. [Online]. Available: https://ieeexplore.ieee.org/document/8849553/

- A. Poostindouz and R. Safavi-Naini, "Secret key agreement in wiretapped Tree-PIN," *To be submitted to IEEE Transactions on Information Theory*, 2022. [Online]. Available: http://arxiv.org/abs/1903.06134

Chapter 5

Previous multiterminal channel models, including the single-input multi-output channel model of [22] and the multiaccess channel model of [23], assume that each terminal either controls one input to the channel, or receives one output variable of the channel. In Chapter 5, we propose a new multiterminal channel model for information-theoretic secret key agreement (SKA) that realistically models wireless communication settings and has the channel models of [22] and [23] as special cases. In our channel model, which we call *the transceiver model*, each terminal (transceiver) $j \in \mathcal{M}$ controls an input variable X_j and observes an output variable Y_j of the underlying noisy channel. Let $V_j = (X_j, Y_j)$ denote the collection of (input and output) variables of terminal j. The channel may be wiretapped which is modeled by providing an output variable Z to Eve. See Figure 1.2. Let $\mathcal{M}$ be the terminal set, then the channel model is denoted by $W = (\mathcal{X}_{\mathcal{M}}, P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}, \mathcal{Y}_{\mathcal{M}} \times \mathcal{Z})$, where

$$P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_m \times \mathcal{Z}.$$
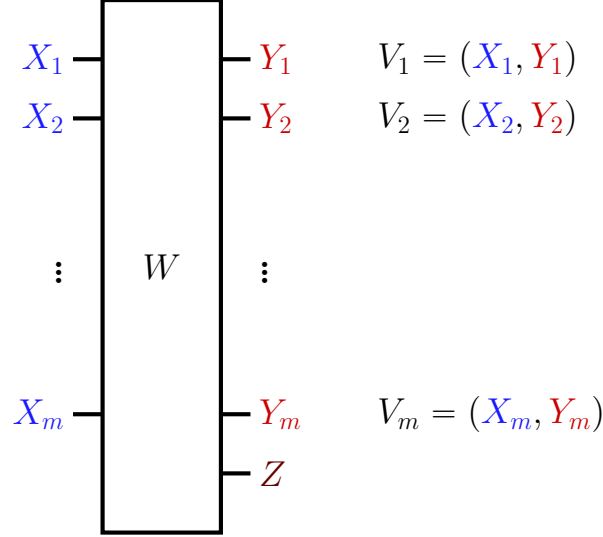
Figure 1.2: The transceiver channel model.

For the transceiver model, we give upper and lower bounds for the SK, PK, and WSK capacities. Our lower bounds are based on the general source emulation approach of [23]. For the proof of our upper bounds, we use a new method that associates any transceiver model with a multiaccess model [23] that has additional "dummy input terminals." We prove that the key capacity of the associated multiaccess model is an upper bound to the key capacity of the original transceiver model, which enables us to employ the converse techniques of [23] to prove our upper bounds for the transceiver model. We then prove the *non-adaptive SK capacity* of non-wiretapped transceiver model under the assumptions that (i) public communication is invoked after all n symbol transmissions over the noisy channel and (ii) input variables of the noisy channel are IID and are generated independently.

The contributions of Chapter 5 appear also in the following papers:

- A. Poostindouz and R. Safavi-Naini, "A channel model of transceivers for multiterminal secret key agreement," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, Oct. 2020, pp. 412–416, Copyright© 2020 IEICE. [Online]. Available: https://ieeexplore.ieee.org/document/9366098

- A. Poostindouz and R. Safavi-Naini, "Secret key agreement in multiterminal channel model of transceivers," *To be submitted to Entropy*, 2022. [Online]. Available: https://arxiv.org/abs/2008.02977

Chapter 6

In Chapter 6, we turn our attention to a special class of multiterminal transceiver channel models, called *wiretapped Polytree-PIN*, in which the underlying noisy channel is given by a collection of independent point-to-point channels such that they define a polytree if we represent each point-to-point channel by a direct edge. A *polytree* is a directed graph whose undirected version is a tree. We consider wiretapped Polytree-PIN with independent leakage, where the output of each point-to-point channel is partially leaked to Eve through a secondary independent noisy channel. Let $\mathcal{M}$ be the terminal set and let $G = (\mathcal{M}, \mathcal{E})$ be a polytree. Each directed edge $e_{ij} \in \mathcal{E}$ represents a point-to-point channel from terminal i to terminal j. The input RV of each terminal j is of the form $X_j = (X_{ji}| \ e_{ji} \in \mathcal{E})$ and its output RV is of the form $Y_j = (Y_{ji}| \ e_{ij} \in \mathcal{E})$. A noisy (wiretapped) version of each Y_{ji}, which we denote by Z_{ij}, is available to the adversary and the collection of all wiretapped components is denoted by $Z = (Z_{ij}| \ e_{ij} \in \mathcal{E})$. Therefore, in a wiretapped Polytree-PIN with independent leakage, with respect to each $e_{ij} \in \mathcal{E}$ the channel model $P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$ satisfies the Markov relation $X_{ij} - Y_{ji} - Z_{ij}$. Polytree-PIN is the channel model counterpart of the Tree-PIN source model. See Figure 1.3.

The main contribution of Chapter 6 is the derivation of WSK capacity for wiretapped Polytree-PIN with independent leakage for any arbitrary $\mathcal{A} \subseteq \mathcal{M}$. When $\mathcal{A} = \mathcal{M}$ the WSK capacity is given by

$$C_{WSK} = \max_{P_{X_{\mathcal{M}}}} \min_{i,j \in \mathcal{M}} I(X_{ij}; Y_{ji}|Z_{ij}).$$

To prove this above result, we first take advantage of the Markov relations that are present in the model, and without assuming non-adaptive input generation, we prove an
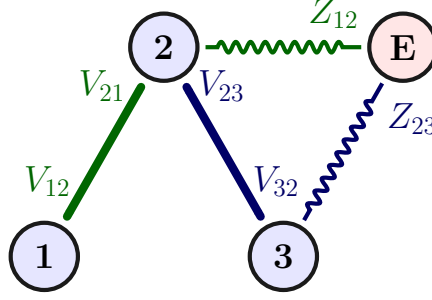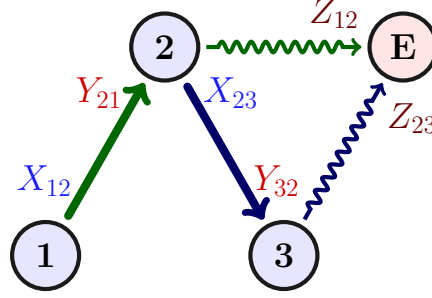
Figure 1.3: A simple wiretapped Polytree-PIN. Here $\mathcal{M} = \{1, 2, 3\}$, $\mathcal{E} = \{e_{12}, e_{23}\}$, $V_1 = X_{12}$, $V_2 = (X_{23}, Y_{21})$, $V_3 = Y_{32}$, and Eve's side information is $Z = (Z_{12}, Z_{23})$. Both Markov relations $X_{12} - Y_{21} - Z_{12}$ and $X_{23} - Y_{32} - Z_{23}$ hold, that is the channel model conditional distribution is $P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} = P_{Y_{21}|X_{12}}P_{Z_{12}|Y_{21}}P_{Y_{32}|X_{23}}P_{Z_{23}|Y_{32}}$.

upper bound on the WSK capacity. Then, we show that the proven upper bound is tight by proposing a WSK capacity achieving channel model SKA protocol. The protocol uses the simple source emulation approach which starts by sending independent IID input symbols through the point-to-point channels, which leads to the realization of a wiretapped Tree-PIN *source* model. Subsequently, by the application of the key capacity achieving source model protocol of Chapter 4, one can achieve key rates that are arbitrarily close to the key capacity. Our result also implies the SK capacity of the non-wiretapped Polytree-PIN model, that is the case when there is no leakage from point-to-point channels to Eve.

The contributions of Chapter 6 appear also in the following paper:

- A. Poostindouz and R. Safavi-Naini, "Secret key capacity of wiretapped Polytree-PIN," in *The 2021 IEEE Information Theory Workshop (ITW2021)*. IEEE, Oct. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9611419

Chapter 2 that follows, provides the sufficient background and intuition on the questions, models, and methods that we explore in this thesis. The final Chapter 7 concludes the thesis by providing suggestions for future research directions.

Chapter 2

Background

Information-theoretic models of key agreement fall under two general categories: the *source model* and the *channel model*. In source model, terminals are given prior access to correlated information sources and are allowed to perform key agreement by public discussion. In channel model, terminals are connected by a noisy information channel, and they can use public discussion while using the noisy channel. In this chapter, we review both of these models in the two-party and multiterminal settings. However, we first review the essential notions, definitions, and theorems; and then recall some important information-theoretic tasks, including "source coding," "information reconciliation," and "privacy amplification."

2.1 Preliminaries

Notations

Sets and Vectors. We use upper-case calligraphic letters (e.g., $\mathcal{M}$, $\mathcal{A}$, etc.) to denote sets, and for any natural number m we define $[m] := \{1, 2, \ldots, m\}$. The notation $|\mathcal{M}|$ is used for denoting the cardinality of a set $\mathcal{M}$. Let $\mathcal{M} = [m]$, then $X_{\mathcal{M}} = X_{[m]} := (X_1, X_2, \ldots, X_m)$ and $X_{\mathcal{A}} = (X_j | \ \forall j \in \mathcal{A})$ for any $\mathcal{A} \subseteq \mathcal{M}$. A set of a finite number of disjoint sets $\mathcal{P} = \{\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{|\mathcal{P}|}\}$ is a partition of a set $\mathcal{M}$ if $\mathcal{P}_j \cap \mathcal{P}_k = \varnothing$ for any two parts $\mathcal{P}_j$ and $\mathcal{P}_k$ in $\mathcal{P}$

and $\bigcup_{j=1}^{|\mathcal{P}|} \mathcal{P}_j = \mathcal{M}$. We show the set of all nontrivial partitions of $\mathcal{M}$ by $\mathcal{P}(\mathcal{M})$. By $\mathcal{A}^n$ we denote the n-th Cartesian product of $\mathcal{A}$ – i.e., $\mathcal{A}^n = \mathcal{A} \times \cdots \mathcal{A}$ for n times. Let $\mathcal{M} = [m]$, then for an arbitrary real vector $R_{\mathcal{M}} = (R_1, \ldots, R_m) \in \mathbb{R}^m$ and for any $\mathcal{A} \subseteq \mathcal{M}$ we define $R_{\mathcal{A}} = (R_j | \; \forall j \in \mathcal{A})$ and $\operatorname{sum}(R_{\mathcal{A}}) := \sum_{j \in \mathcal{A}} R_j$.

Discrete Random Variables. In this study, random variables are discrete variables unless otherwise mentioned. We reserve upper-case letters to denote random variables (RVs) and lower-case letters to denote their realizations. We use calligraphic upper-case letters to show the alphabets of random variables. The probability mass function (PMF) of a discrete random variable X is denoted by $P_X(x) = \Pr\{X = x\}$, and the probability distribution (or probability vector) of X is abbreviated by $P_X = (P_X(x) | \; x \in \mathcal{X})$. The notation $X \sim P_X$ means that RV X has distribution P_X. For example, $X \sim \operatorname{Bern}(q)$ indicates that X is a binary random variable with Bernoulli distribution of success probability q; i.e.,

$$X = \begin{cases} 1 & \text{with probability } q, \\ 0 & \text{with probability } 1 - q. \end{cases}$$

The notations $\mathbb{E}_{P_X}\{X\}$ (or $\mathbb{E}\{X\}$) and $\operatorname{Var}_{P_X}\{X\}$ (or $\operatorname{Var}\{X\}$), are used for expected value and variance of $X \sim P_X$, respectively. The support of a random variable X is denoted by $\operatorname{supp}(X) = \{x \in \mathcal{X} | P_X(x) > 0\}$.

For two random variables X and Y, we use P_{XY} to show their joint probability distribution and $P_{X|Y}$ to represent the conditional probability distribution of X given Y. That is for any $x \in \mathcal{X}$ and any $y \in \mathcal{Y}$ we have $P_{XY}(x, y) = \Pr\{X = x, Y = y\}$, and

$$P_{X|Y}(x|y) = \Pr\{X = x | Y = y\} := \frac{\Pr\{X = x, Y = y\}}{\Pr\{Y = y\}} = \frac{P_{XY}(x, y)}{P_Y(y)}.$$

For a specified dimension n we use $X^n = (X_1, \ldots, X_n)$ to denote a random vector defined over the joint alphabet $\mathcal{X}^n = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$ with joint probability distribution P_{X^n}, and

$x^n = (x_1, \ldots, x_n)$ denotes the corresponding n-fold realization.

Common Functions. All logarithms in this work are in base 2, $\log x = \log_2 x$, unless otherwise mentioned. Also, $\log \log x = \log(\log(x))$ and $\ln x = \log_e x$ denotes the natural logarithm where constant e is the Euler's number defined as $e := \lim_{\tau \to 0} \sqrt[\tau]{1 + \tau}$. Let $h_2(a) := -a \log a - (1 - a) \log(1 - a)$ be the *binary entropy* function. Denoted by $Q(\cdot)$ is the tail probability of the standard Gaussian distribution,

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t.$$

Big-O notations. We use the following notations to describe asymptotic behavior of real valued functions.

$$o(g(n)) = \{f : \mathbb{R}_+ \to \mathbb{R}_+ | \forall c_2 > 0 \quad \exists n_0 > 0 : \quad 0 \leqslant f(n) \leqslant c_2 g(g) \ \forall n \geqslant n_0\},$$

$$\mathcal{O}(g(n)) = \{f : \mathbb{R}_+ \to \mathbb{R}_+ | \exists c_2 > 0 \quad \exists n_0 > 0 : \quad 0 \leqslant f(n) \leqslant c_2 g(g) \ \forall n \geqslant n_0\},$$

$$\Theta(g(n)) = \{f : \mathbb{R}_+ \to \mathbb{R}_+ | \exists c_1, c_2 > 0 \ \exists n_0 > 0 : \ 0 \leqslant c_1 g(n) \leqslant f(n) \leqslant c_2 g(g) \ \forall n \geqslant n_0\}.$$

Note that, $f(n) \in o(g(n))$ is equivalent with

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0,$$

thus, $f(n) \in o(n)$ means $\lim_{n \to \infty} {}^1\!/\!_n f(n) = 0$, and, $f(n) \in o(1)$ means $\lim_{n \to \infty} f(n) = 0$. Also, whenever we write

$$f(n) = h(n) \pm \mathcal{O}(g(n)),$$

that means, for some $g_l, g_u \in \mathcal{O}(g(n))$, we have

$$h(n) - g_l(n) \leqslant f(n) \leqslant h(n) + g_u(n).$$

Random Variables, Independence and Markov Chains

In this section, we briefly review some basic notions of probability theory.

Definition 2.1 (Discrete Probability Space – Chapter 2 of [39]). A Discrete Probability Space is defined by a finite discrete set Ω called the sample space, which is the set of all possible outcomes $\omega \in \Omega$, and a probability function $\Pr\{\cdot\}$, where the following holds:

$$0 \leqslant \Pr\{\omega\} \leqslant 1 \qquad \forall \omega \in \Omega,$$

$$\sum_{\omega \in \Omega} \Pr\{\omega\} = 1.$$

Subsets of Ω are called events and for any $\mathcal{E} \subseteq \Omega$, $\Pr\{\mathcal{E}\} = \sum_{\omega \in \mathcal{E}} \Pr\{\omega\}$.

Definition 2.2 (Discrete Random Variable – Chapter 4 of [39]). A Discrete Random Variable (RV) is essentially a function X from the sample space Ω of a probability space to a subset of the set of real numbers $\mathbb{R}$, where the following properties hold:

- X may be undefined or infinite for an event that has zero probability.

- For every $x \in \mathbb{R}$ then $\{\omega \in \Omega \mid X(\omega) \leqslant x\}$ is an event.

The probability mass function (PMF) of a discrete random variable X is given by $P_X(x) = \Pr\{X = x\}$, and the probability distribution of X is denoted by $P_X = (P_X(x) \mid x \in \mathcal{X})$. The range of RV X is called its alphabet and is denoted by $\mathcal{X}$.

Definition 2.3 (Independence – Definition 2.1 of [40]). Two random variables X and Y are independent if and only if

$$P_{XY}(x, y) = P_X(x)P_Y(y), \qquad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

If two variables X and Y are not independent, they are called *correlated*.

Definition 2.4 (Conditional Independence – Definition 2.4 of [40]). Two random variables X and Z are independent conditioned on Y if and only if

$$P_{XYZ}(x, y, z)P_Y(y) = P_{XY}(x, y)P_{YZ}(y, z) \quad \forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z},$$

or equivalently

$$P_{XZ|Y}(x, z|y) = P_{X|Y}(x|y)P_{Z|Y}(z|y) \quad \text{if } P_Y(y) > 0,$$

otherwise $P_{XYZ}(x, y, z) = 0$.

Definition 2.5 (Mutual Independence – Definition 2.2 of [40]). For $n > 2$, the sequence of RVs $X^n = (X_1, \ldots, X_n)$ are mutually independent if and only if

$$P_{X^n}(x^n) = P_{X_1}(x_1)P_{X_2}(x_2) \cdots P_{X_n}(x_n) \quad \forall x^n \in \mathcal{X}^n.$$

Definition 2.6 (Markov Chain – Definition 2.6 of [40]). For $n > 2$, the sequence of RVs $X^n = (X_1, \ldots, X_n)$ form a Markov chain, denoted by $X_1 - X_2 - \cdots - X_n$, if and only if

$$P_{X^n}(x^n)P_{X_2}(x_2)P_{X_3}(x_3) \cdots P_{X_{n-1}}(x_{n-1}) = P_{X_1 X_2}(x_1, x_2)P_{X_2 X_3}(x_2, x_3) \cdots P_{X_{n-1} X_n}(x_{n-1}, x_n),$$

for all $x^n \in \mathcal{X}^n$. Or equivalently

$$P_{X^n}(x^n) = P_{X_1 X_2}(x_1, x_2)P_{X_3|X_2}(x_3|x_2) \cdots P_{X_n|X_{n-1}}(x_n|x_{n-1}),$$

if $P_{X_j}(x_j) > 0 \ \forall j \in (2, 3, \ldots, n-1)$, otherwise $P_{X^n}(x^n) = 0$.

Remark 2.1. Independence of X and Z conditioned on Y is equivalent to the Markov chain relation of $X - Y - Z$; since $P_{XYZ}(x, y, z)P_Y(y) = P_{XY}(x, y)P_{YZ}(y, z)$ implies both relations.

Statistical Distance

One of the most practical and relevant distance measures in information theoretic security is the *Statistical Distance*.

Definition 2.7 (Statistical Distance – Definition 11.1 and Lemma 11.1 of [41]).
For two random variables $X \sim P_X$ and $Y \sim P_Y$ defined over the same alphabet $\mathcal{W}$, the the statistical distance between X and Y, denoted by $\mathbf{SD}(X, Y)$, is defined as

$$\mathbf{SD}(X, Y) = \frac{1}{2} \sum_{w \in \mathcal{W}} |P_X(w) - P_Y(w)|$$

$$= \max_{\mathcal{T} \subseteq \mathcal{W}} \sum_{w \in \mathcal{T}} P_X(w) - P_Y(w)$$

$$= \sum_{w \in \mathcal{T}^*} P_X(w) - P_Y(w),$$

where $\mathcal{T}^* = \{w \in \mathcal{W} \mid P_X(w) \geqslant P_Y(w)\}$.

Lemmas 2.1, and 2.2 that follow next are direct consequences of Definition 2.7 (Definition 11.1 and Lemma 11.1 of [41].) We present their proofs here for completeness.

Lemma 2.1. *For two random variables X and X' over the same alphabet $\mathcal{X}$ and two random variables Y and Y' over the same alphabet $\mathcal{Y}$ we have $\mathbf{SD}(X, X') \leqslant \mathbf{SD}(XY, X'Y')$.*

Proof: Let $(X, Y) \sim P_{XY}$ and $(X', Y') \sim Q_{X'Y'}$. Then

$$\mathbf{SD}(XY, X'Y') = \max_{\mathcal{T} \subseteq \mathcal{X} \times \mathcal{Y}} \sum_{(x,y) \in \mathcal{T}} P_{XY}(x, y) - Q_{X'Y'}(x, y)$$

$$\geqslant \sum_{(x,y) \in \mathcal{V}} P_{XY}(x, y) - Q_{X'Y'}(x, y)$$

$$= \sum_{x \in \overline{\mathcal{X}}} \sum_{y \in \mathcal{Y}} P_{XY}(x, y) - \sum_{y \in \mathcal{Y}} Q_{X'Y'}(x, y)$$

$$= \sum_{x \in \overline{\mathcal{X}}} P_X(x) - Q_{X'}(x) = \mathbf{SD}(X, X'),$$

where $\mathcal{V} = \overline{\mathcal{X}} \times \mathcal{Y}$, with $\overline{\mathcal{X}} = \{x \in \mathcal{X} \mid P_X(x) \geqslant Q_{X'}(x)\}$. ∎

Corollary 2.1.1. *If* $\mathbf{SD}(XY, X'Y') \leqslant \epsilon$ *then* $\mathbf{SD}(X, X') \leqslant \epsilon$ *and* $\mathbf{SD}(Y, Y') \leqslant \epsilon$.

Corollary 2.1.2. *(i) For any three RVs X, X', and Y we have* $\mathbf{SD}(X, X') \leqslant \mathbf{SD}(XY, X'Y)$.
(ii) If Y is independent from X and X', then $\mathbf{SD}(X, X') = \mathbf{SD}(XY, X'Y)$.

Corollary 2.1.3. *Suppose X and X' are correlated RVs independent from correlated RVs Y and Y'. Then,* $\mathbf{SD}(XY, X'Y') \leqslant \mathbf{SD}(X, X') + \mathbf{SD}(Y, Y')$.

Proof: Using the triangle inequality we have

$$\mathbf{SD}(XY, X'Y') \leqslant \mathbf{SD}(XY, X'Y) + \mathbf{SD}(X'Y, X'Y')$$

$$= \mathbf{SD}(X, X') + \mathbf{SD}(Y, Y'),$$

where the equality is due to preposition (ii) of Corollary 2.1.2. ∎

Lemma 2.2. *Let RV's $X \sim P_X$ and $X' \sim P_{X'}$ be defined over the same alphabet $\mathcal{X}$ and let W be an arbitrary random function characterized by the conditional probability distribution $Q_{Y|X}$, where Y denotes the output of W which takes values over $\mathcal{Y}$. Define $Y \sim P_Y$ and $Y' \sim P_{Y'}$ be the output RV's of W when input RV's are $X \sim P_X$ and $X' \sim P_{X'}$, respectively. Then we have* $\mathbf{SD}(Y, Y') \leqslant \mathbf{SD}(X, X')$.

Proof: Let $(X, Y) \sim P_{XY} = P_X Q_{Y|X}$ and $(X', Y') \sim P_{X'Y'} = P_{X'} Q_{Y|X}$. Then

$$\mathbf{SD}(XY, X'Y') = \frac{1}{2} \sum_{x,y} |P_{XY}(x, y) - P_{X'Y'}(x, y)|$$

$$= \frac{1}{2} \sum_{x,y} Q_{Y|X}(y|x) |P_X(x) - P_{X'}(x)|$$

$$= \frac{1}{2} \sum_{x} |P_X(x) - P_{X'}(x)| \sum_{y} Q_{Y|X}(y|x)$$

$$= \frac{1}{2} \sum_{x} |P_X(x) - P_{X'}(x)|$$

$$= \mathbf{SD}(X, X').$$

Thus, by Lemma 2.1 the proof is complete. ∎

Asymptotic Convergence

Consider a sequence of real-valued variables $\{X_j\}_{j=1}^{\infty}$ that are mutually independent and are drawn from the same identical distribution (IID). We are interested to know about the distribution of the variable $X_n^{\text{sum}} = \frac{1}{n}\sum_{j=1}^{n} X_j$. First, we define the following convergence notions.

Convergence in Distribution. A sequence of real-valued RV's $\{X_j\}_{j=1}^{\infty}$ (not necessarily IID) converges in distribution to RV X if $\lim_{n\to\infty} \Pr\{X_n \leqslant a\} = \Pr\{X \leqslant a\}$ $\forall a \in \mathbb{R}$.

Convergence in Probability. A sequence of real-valued RV's $\{X_j\}_{j=1}^{\infty}$ (not necessarily IID) converges in probability to RV X if $\lim_{n\to\infty} \Pr\{|X_n - X| > \epsilon\} = 0$, $\forall \epsilon > 0$.

Convergence in probability implies convergence in distribution [42, Chapter 1, Section 1.7.5].

Almost Sure Convergence. A sequence of real-valued RV's $\{X_j\}_{j=1}^{\infty}$ (not necessarily IID) almost surely converges to RV X if $\Pr\{\lim_{n\to\infty} X_n = X\} = 1$.

Almost sure convergence implies convergence in probability [42, Chapter 1, Section 1.7.5].

The following theorem (LLN) shows that X_n^{sum} converges to $\mu = \mathbb{E}\{X_1\}$ as $n \to \infty$.

Theorem 2.3 (The Law of Large Numbers (LLN) – Theorem 1.3.2 of [24]). *Let* $\{X_j\}_{j=1}^{\infty}$ *be a sequence of IID real-valued random variables. If* $\mu = \mathbb{E}\{X_1\} < +\infty$, *then* $X_n^{sum} = \frac{1}{n}\sum_{j=1}^{n} X_j$ *almost surely converges to* μ *as* $n \to \infty$.

Note that the above statement which is called the *strong* LLN also implies that X_n^{sum} converges to μ as $n \to \infty$ in probability. The latter statement is called the *weak* LLN.

The prominent central limit theorem states that as $n \to \infty$ not only X_n^{sum} converges to μ but also the distribution of variable $Y_n = \sqrt{n}(X_n^{\text{sum}} - \mu)$ converges to a normal distribution – see Chapter 7, page 283 of [39]. The Berry-Esseen theorem stated below is a non-asymptotic

inequality that explains this phenomenon more qualitatively by even giving the speed of such convergence.

Theorem 2.4 (Berry-Esseen, see [43] Theorem 1, Chapter XVI, Section 5). *Let X^n be an $n-IID$ real-valued variable, and $-\infty < \alpha < \infty$, then*

$$\left| \Pr \left\{ \sum_{j=1}^{n} X_j \geqslant n\mu + \alpha\sqrt{\Delta n} \right\} - Q(\alpha) \right| \leqslant \frac{3\rho}{\Delta^{3/2}\sqrt{n}},$$

where $Q(\cdot)$ is the tail probability of the standard Gaussian distribution, i.e.,

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt,$$

$\mu = \mathbb{E}\{X_1\}, \Delta = \mathrm{Var}\{X_1\}, \text{ and } \rho = \mathbb{E}\{|X_1 - \mu|^3\}.$

2.2 Information Theoretic Concepts

Information Theory, originated from the seminal work of Shannon [44], studies the fundamental problems of transmitting and processing information under various settings and conditions. What follows next is a brief overview of some of the basic information-theoretic concepts. We begin by introducing simple mathematical models of information sources and channels.

Definition 2.8 (Discrete Memoryless Source Model). A Discrete Memoryless Source (DMS) Model is defined by a finite alphabet $\mathcal{X}$ and a discrete probability distribution $P_X(x)$, and informally is denoted by X or P_X. A DMS generates a sequence of independent and identically distributed (IID) random values $X_1, X_2, \ldots$. We denote n consecutive runs of a DMS by $X^n = (X_1, \ldots, X_n)$.

In this work, all sources are discrete memoryless, unless otherwise specified. A $d-$DMS or a Discrete Multiple Memoryless Source (DMMS) Model with d components, is simply a

collection of d correlated DMS's. A DMMS is defined by a finite alphabet $\mathcal{X}_1 \times \cdots \times \mathcal{X}_d$ and a joint probability distribution $P_{X_1 \cdots X_d} = P_{X_{[d]}}$, and is denoted by $X_{[d]} = (X_1, \ldots, X_d)$. The output of $X_{[d]}$ at time t is denoted by $(X_{1,t}, \ldots, X_{d,t})$.

A simple mathematical model of communication channels is the following.

Definition 2.9 (Discrete Memoryless Channel Model). A discrete communication channel model is a random system that given input value x outputs a random value y. Let X_t (and Y_t) denote the input (and random output) variable of the channel at time t. A discrete channel is said to be memoryless (DMC) and time invariant if at any time t

$$\Pr\left\{Y_t = y | X^{(t-1)}, Y^{(t-1)}, X_t = x\right\} = P_{Y|X}(y|x) \quad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}.$$

We denote a DMC by its input and output alphabets and its characterizing conditional probability distribution (transition matrix.) For example a DMC W with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and conditional distribution $P_{Y|X}$ is denoted by $W = (\mathcal{X}, P_{Y|X}, \mathcal{Y})$, and sometimes informally by $W = P_{Y|X}$.

In this work, all communication channels are assumed to be discrete, memoryless, and time invariant; unless otherwise we explicitly mention the type of the channels.

Remark 2.2. Assume that for any $1 \leqslant j < n$ the RV X_j is correlated with the RV X_{j+1} through a [discrete, memoryless, and time invariant] information channel, then the sequence of RVs $X_1, \ldots, X_n$ form a Markov chain $X_1 - X_2 - \cdots - X_n$.

Measures of Information

Here, we present a brief overview on the measures of information. These measures and their interrelations are the essential mathematical toolbox that is needed to prove information theoretical secrecy results. We begin by defining the Shannon Entropy.

The Shannon entropy, is a mathematical function for measuring the amount of uncertainty there is about the outcome of a random process. Equivalently, the Shannon entropy of

a random variable is the average amount of information one receives by observing the random outcome, which is equal to the reduction of uncertainty one had before the observation.

Definition 2.10 (Shannon Entropy). The Shannon entropy of a random variable $X \sim P_X$ over alphabet $\mathcal{X}$ is given by

$$
\begin{aligned}
H(X) &= \mathbb{E}_{P_X} \left\{ \log \frac{1}{P_X(X)} \right\} \\
&= - \sum_{x \in \mathcal{X}} P_X(x) \log P_X(x),
\end{aligned}
$$

where the logarithm is in base 2.

The unit of entropy is bit(s). For example, for a fair coin toss, we can represent the random outcome with a uniform binary distribution $X \sim \text{Bern}(1/2)$. The entropy of X is $H(X) = 1$, that means we get "one bit of information" by observing the outcome of a fair coin toss, or that the amount of uncertainty about the outcome of a fair coin toss is one bit.

The Shannon entropy; however, is not always the best choice for measuring uncertainty. Especially, in most security scenarios, a better measure of uncertainty is the min-entropy.

Definition 2.11 (Min-entropy). The Min-Entropy of a random variable $X \sim P_X$ over alphabet $\mathcal{X}$ is defined by

$$
\begin{aligned}
H_{\min}(X) &= \min_{x \in \mathcal{X}} \left\{ \log \frac{1}{P_X(x)} \right\} \\
&= - \log(\max_{x \in \mathcal{X}} \{ P_X(x) \}).
\end{aligned}
$$

For two random variables X and Y with joint and conditional distributions $P_{XY}(x, y)$ and $P_{X|Y}(x|y)$, respectively, the *joint entropy* $H(X, Y)$ is given by

$$
\begin{aligned}
H(X, Y) &= \mathbb{E}_{P_{XY}} \left\{ \log \frac{1}{P_{XY}(X, Y)} \right\}, \\
&= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{XY}(x, y) \log P_{XY}(x, y),
\end{aligned}
$$

and the *conditional entropy* $H(X|Y)$ is defined as

$$H(X|Y) = \mathbb{E}_{P_{XY}} \left\{ \log \frac{1}{P_{X|Y}(X|Y)} \right\},$$
$$= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{XY}(x,y) \log P_{X|Y}(X|Y).$$

The following relations hold for joint entropy and conditional entropy. For any given RVs X, Y, and Z we have

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y),$$

and

$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z) = H(Y|Z) + H(X|Y,Z).$$

For two independent random variables X and Y we have $H(X,Y) = H(X) + H(Y)$, and $H(X|Y,Z) = H(X|Z)$.

To measure the amount of information one random variable has about another random variable, we use the *mutual information* function.

Definition 2.12 (Mutual Information). The mutual information between two random variables X and Y is

$$I(X;Y) = \mathbb{E}_{P_{XY}} \{i(X,Y)\}$$
$$= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}$$
$$= H(X) - H(X|Y) = H(Y) - H(Y|X),$$

where

$$i(x,y) = \log \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)},$$

is called the information density [24].

23
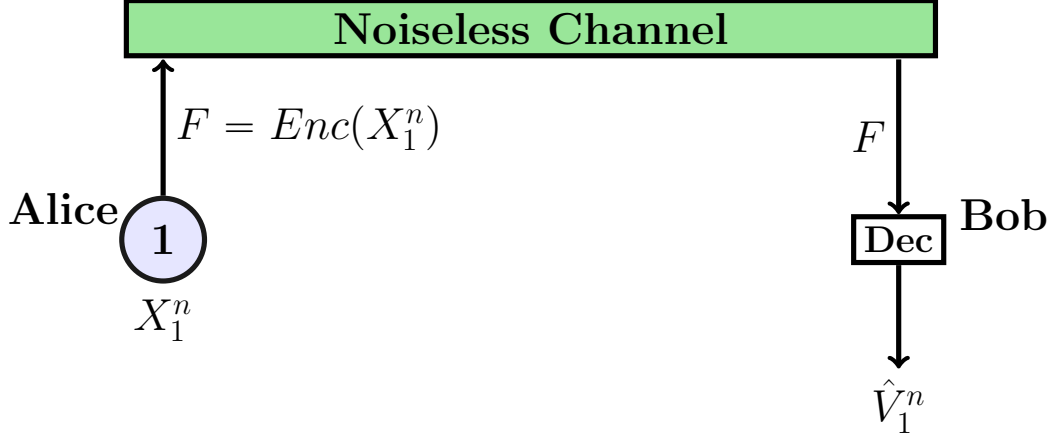
Figure 2.1: The source coding problem.

Definition 2.13 (Conditional Mutual Information). For three random variables X, Y, and Z, the *conditional mutual information* of X and Y given Z is

$$I(X; Y|Z) = \mathbb{E}_{P_{XYZ}} i(X, Y|Z)$$
$$= H(X|Z) - H(X|Y, Z)$$
$$= H(Y|Z) - H(Y|X, Z),$$

where

$$i(x, y|z) = \log \frac{P_{XY|Z}(x, y|z)}{P_{X|Z}(x|z) P_{Y|Z}(y|z)},$$

is the conditional information density.

Remark 2.3. When the Markov relation $X - Y - Z$ holds, we have $I(X; Y|Z) = H(X|Z) - H(X|Y)$, and $I(X; Z|Y) = 0$.

2.2.1 Source Coding

In the source coding (or data compression) problem, Alice (terminal 1), who has a random sequence X_1^n encodes (compresses) her sequence into a message $F = Enc(X_1^n)$ and sends this message F to Bob (terminal 1) using a noiseless communication channel. Bob, then decodes

(decompresses) the message F into an estimate (reconstruction) $\hat{X}_1^n = Dec(F)$ of Alice's sequence. See Figure 2.1. The main objective is to find a pair of algorithms (Enc, Dec) such that for every $n \in \mathbb{N}$ they enable reliable (correct) reconstruction by the shortest message length possible. To be more precise, consider the following source model. A discrete memoryless source (DMS) represented by the random variable X_1, defined over the alphabet $\mathcal{X}_1$ and described by the probability distribution P_{X_1}, is accessible to Alice. Alice samples from X_1 for n times, and observes the independent and identically distributed (IID) sequence X_1^n. Consider a compression code family (Enc, Dec). The asymptotic compression rate of the code is defined as $\limsup_{n\to\infty} 1/n \log|\mathcal{F}|$ and the compression code is called ϵ_n-correct if $\Pr\left\{X_1^n = \hat{X}_1^n\right\} \geqslant 1 - \epsilon_n$. The parameter ϵ_n is called the reliability parameter. Note that here the message length is measured by $\log|\mathcal{F}|$.

The source coding theorem (due to Shannon [44]) gives the minimum asymptotic compression rate of all compression (source) codes with asymptotic perfect reliability.

Theorem 2.5 (Source Coding – see Theorem 3.4 of [45]). *Suppose the probability distribution P_{X_1} of a DMS X_1 is known. Then, the minimum achievable (asymptotic zero error) compression rate is $R^* = H(X_1)$. That is for every $R \in \mathbb{R}$ satisfying,*

$$R \geqslant H(X_1),$$

there exists an ϵ_n-correct compression code family (Enc, Dec) with asymptotic compression rate R and $\lim_{n\to\infty} \epsilon_n = 0$; and every ϵ_n-correct compression code family (Enc, Dec) with $\lim_{n\to\infty} \epsilon_n = 0$ has asymptotic compression rate higher than $H(X_1)$.

A modified version of the source coding problem is for the 2-DMS model in which we assume Bob (terminal 2) has access to a variable X_2 that is correlated with Alice's variable X_1. See Figure 2.2. Next theorem (due to Slepian and Wolf) proves that existence of side information X_2 at the decoder (Bob) allows for lower compression rates.
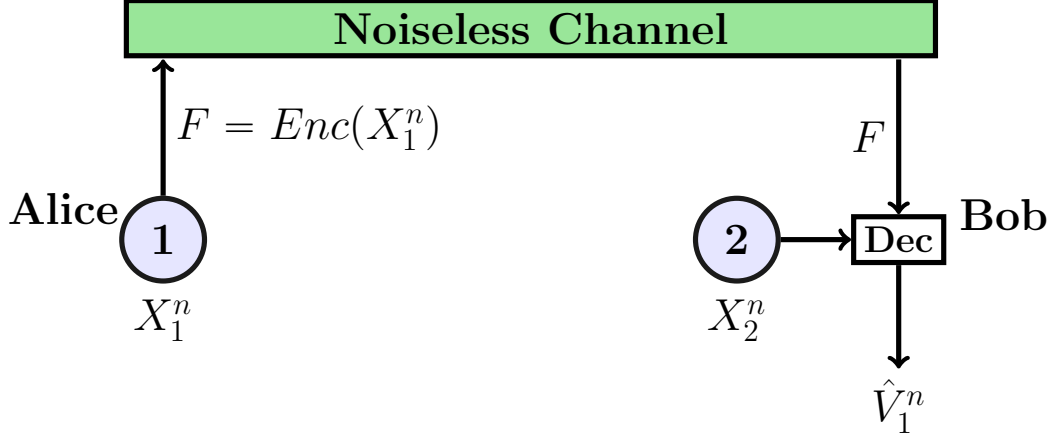
Figure 2.2: The source coding problem with side information at the decoder.

Theorem 2.6 (Source Coding with Side Information at the Decoder – see Section 10.4 of [45]). *Suppose the joint probability distribution $P_{X_1 X_2}$ of a 2-DMS (X_1, X_2) is known. Then, the minimum achievable (asymptotic zero error) compression rate is $R^* = H(X_1|X_2)$. That is for every $R \in \mathbb{R}$ satisfying,*

$$R \geqslant H(X_1|X_2),$$

there exists an ϵ_n-correct compression code family (Enc, Dec) with asymptotic compression rate R and $\lim_{n\to\infty} \epsilon_n = 0$; and every ϵ_n-correct compression code family (Enc, Dec) with $\lim_{n\to\infty} \epsilon_n = 0$ has asymptotic compression rate higher than $H(X_1|X_2)$.

2.2.2 Stochastic Processes and General Sources

We defined the Discrete Memoryless Source (DMS) model that outputs IID sequences of random variables $X_1, X_2, \ldots$. See Definition 2.8. This notion of information source can be generalized by the concept of *stochastic process* that is defined as an arbitrary sequence of random variables that are defined over the same alphabet according to the same probability space (see Definition 2.1), where the sequence of RV's need not to be IID. A stochastic

process X is denoted by the sequence of RV's $\{X_i\}_{i=1}^{\infty}$ and by $X^n = (X_1, X_2, \ldots, X_n)$ we denote the n consecutive runs of the stochastic process X.

A stochastic process $\{X_i\}_{i=1}^{\infty}$ is called *stationary* if for all $n \in \mathbb{N}$ and for every $x^n = (x_1, \ldots, x_n) \in \mathcal{X}^n$ and every integer τ we have

$$P_{X_1 X_2 \cdots X_n}(x^n) = P_{X_{1+\tau} X_{2+\tau} \cdots X_{n+\tau}}(x^n).$$

See Chapter 9, page 387 of [39]. An IID process is an stationary process.

A stochastic process $\{X_i\}_{i=1}^{\infty}$ is called *ergodic* (or mean-ergodic) if for every n-fold realization $x^n = (x_1, \ldots, x_n) \in \mathcal{X}^n$ the time average $\mathrm{avg}(x^n) = \frac{1}{n} \sum_{j=1}^{n} x_j$ approaches the ensemble average defined by $\mu(n) = \mathbb{E}\{X_n\}$ as $n \to \infty$ with probability close to 1. See Chapter 12, page 523 of [39]. An IID process is an stationary ergodic process.

Definition 2.14 (Entropy Rate - Chapter 4 of [47]). The *entropy rate* of a stochastic process $\{X_i\}_{i=1}^{\infty}$ is

$$\begin{aligned}
H_r(X) &= \lim_{n \to \infty} \frac{1}{n} H(X^n) \\
&= \lim_{n \to \infty} \frac{1}{n} \mathbb{E}_{P_{X^n}} \{-\log P_{X^n}(X^n)\}.
\end{aligned}$$

The source coding theorem 2.5 of IID source model can be generalized to the case of stationary ergodic processes.

Theorem 2.7 (Source Coding for Stationary Ergodic Processes – Theorem 2 of [24]). *The minimum achievable (asymptotic zero error) compression rate for a stationary ergodic process X with a finite source alphabet $\mathcal{X}$ is $R^* = H_r(X)$.*

Information Spectrum Methods

Han and Verdu [24, 48] introduced the notion of general sources and channels that are "arbitrarily *nonstationary* and/or arbitrarily *nonergodic* with arbitrarily given abstract source/channel

alphabets." The goal of this generalization is to obtain results similar to the known ones for these general cases and provide a "unified general treatment of a collection of mostly known results in the literature." The main technical approach is *information-spectrum method* and introducing two probabilistic limit operations for a sequence $(J_1, J_2, J_3, \ldots)$ of real-valued random variables, called the *limit superior in probability* given by

$$\text{p-}\limsup_{n \to \infty} J_n = \inf\{\alpha | \lim_{n \to \infty} \Pr\{J_n > \alpha\} = 0\},$$

and the *limit inferior in probability* given by

$$\text{p-}\liminf_{n \to \infty} J_n = \sup\{\beta | \lim_{n \to \infty} \Pr\{J_n < \beta\} = 0\}.$$

Note that J_n converges to a constant c in probability if and only if $\text{p-}\limsup_{n \to \infty} J_n = \text{p-}\liminf_{n \to \infty} J_n = c$.

Information-spectrum methods refers to mathematical tools and approaches that are used for analyzing information-theoretic tasks with regard to general sources and channels. These methods most often employ information measures (such as the spectral sup-entropy rate which we define below) that are not in terms of the expectation operation. (Whereas for example the Shannon entropy (Definition 2.10) and the entropy rate (Definition 2.14) are defined using the expectation operation.)

The concept of "general sources" generalizes the notion of stochastic processes.

Definition 2.15 (General Source – Chapter 1 of [24]). A *general source* X given by

$$\{X^n = (X_1^{(n)}, \ldots, X_n^{(n)})\}_{n=1}^{\infty},$$

outputs X^n for every $n \in \mathbb{N}$ that is the n-fold random output of the source defined over the set $\mathcal{X}^n$ subject to the joint distribution P_{X^n} – see [24, Chapter 1].

Note that random behavior of the jth component $X_j^{(n)}$ may depend on n. General

source model has stochastic processes as a special case by imposing the following consistency condition

$$X_j^{(m)} = X_j^{(n)} \qquad \forall j \leqslant m, \ \forall m < n.$$

In this case, the n-fold output of the general source X is denoted by $X^n = (X_1, \ldots, X_n)$.

An example of a general source is an information source (stochastic process) where output symbols are independent but not identically distributed (INID).

An important quantity that is defined with respect to any general source X is the *spectral sup-entropy rate* of X, which is given by

$$\overline{H}(X) = \text{p-}\limsup_{n \to \infty} -\frac{1}{n} \log P_{X^n}(X^n)$$
$$= \inf \left\{ \alpha \middle| \ \lim_{n \to \infty} \Pr \left\{ -\frac{1}{n} \log P_{X^n}(X^n) > \alpha \right\} = 0 \right\},$$

where the distribution of the real-valued variable

$$-\frac{1}{n} \log P_{X^n}(X^n)$$

is called the *information spectrum* of X.

The source coding theorems of 2.5 and 2.7 are generalized for the case of general sources by the following theorem which proves that for general sources the optimum compression rate is characterized by the spectral sup-entropy rate.

Theorem 2.8 (General Source Coding – see Theorem 1.3.1 of [24]). *The minimum achievable (asymptotic zero error) compression rate for a general source X is $R^* = \overline{H}(X)$.*

Remark 2.4. The above result implies Theorem 2.5 and Theorem 2.7 as special cases. When X is a DMS and X^n is IID then by the LLN Theorem 2.3 we know that $-\frac{1}{n} \log P_{X^n}(X^n) = \frac{1}{n} \sum_{j=1}^{n} -\log P_{X_j}(X_j)$, converges almost surely and in probability to $\mathbb{E}\left\{-\log P_{X_1}(X_1)\right\}$ and thus $R^* = \overline{H}(X) = H(X)$. When X is a stationary ergodic process then by the general AEP Theorem 16.8.1 of [47] (due to Shannon, McMillan, and Breiman) we know that

$-\frac{1}{n}\log P_{X^n}(X^n)$ converges almost surely (and in probability) to entropy rate, $\lim_{n\to\infty}\frac{1}{n}H(X^n)$, and thus $R^* = \overline{H}(X) = H_r(X)$.

In Chapter 3, we introduce a new spectral entropy (which we call sup-spectral entropy) and use it to prove an upper bound for the maximum key length of one-way secret key agreement.

2.2.3 Channel Coding

In many information-theoretic scenarios (e.g., problems we studied earlier) it is assumed that a noiseless channel is available to the terminals; however, this is not a valid assumption in real-life as almost all mediums for communication are noisy. Channel coding is the process of realizing an almost noiseless channel from a noisy channel by incorporating coding techniques. A simple model of a noisy channel is the discrete memoryless channel (DMC) model – see Definition 2.9.

Suppose Alice and Bob are connected through a DMC $W = P_{Y|X}$. Alice wants to send a message M to Bob using W. To make the message transfer reliable Alice uses a function (algorithm) called channel encoder denoted by Enc that maps M into a codeword X^n, and sends this codeword to Bob by using the channel W for n times. Upon receiving the whole output variable Y^n, Bob uses another function (algorithm) called channel decoder denoted by Dec that maps the variable Y^n to an estimate of the message $\hat{M}$. The objective is to have negligible errors $\Pr\left\{M \neq \hat{M}\right\} \approx 0$, and be able to send messages at the largest rate possible.

A channel code for a DMC $W = (\mathcal{X}, P_{Y|X}, \mathcal{Y})$ is given by three components:

- Message set $\mathcal{M}^{(n)} = \{1, 2, \ldots, |\mathcal{M}^{(n)}|\}$

- Encoding function $Enc : \mathcal{M}^{(n)} \to \mathcal{X}^n$

- Decoding function $Dec : \mathcal{Y}^n \to \mathcal{M}^{(n)} \cup \{\flat\}$

where the special character ♭ denotes decoding error. The set $\mathcal{C} = \{Enc(j)\}_{j=1}^{|\mathcal{M}^{(n)}|}$ is called the codebook, where $c_j = Enc(j)$ is the codeword corresponding to message $j \in \mathcal{M}^{(n)}$, and $|\mathcal{M}^{(n)}|$ is referred to as the code size. The message estimate of Bob is denoted by $\hat{M} = Dec(Enc(M))$. It is assumed that messages are uniformly distributed over $\mathcal{M}^{(n)}$.

Let Ψ be a family of channel codes (indexed by n) as defined above. Then, for every n the average error probability of Ψ is given by $\epsilon_n = \Pr\left\{M \neq \hat{M}\right\}$. The asymptotic channel coding rate of Ψ defined by $r_M(\Psi) = \liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{M}^{(n)}|$ is called *achievable* if Ψ has average error probability ϵ_n such that $\lim_{n \to \infty} \epsilon_n = 0$.

Definition 2.16 (Channel Capacity). The channel capacity of a DMC is the largest achievable channel coding rate.

We denote the channel capacity of DMC $W = P_{Y|X}$ by $C_M(P_{Y|X})$. A channel code family Ψ is an $(|\mathcal{M}^{(n)}|, \epsilon_n)$−code if it has error probability ϵ_n and code size $|\mathcal{M}^{(n)}|$. Note that both of these parameters are functions of n and a "good" code has large code size and negligible error. To be precise, an $(|\mathcal{M}^{(n)}|, \epsilon_n)$−code family Ψ is capacity achieving for channel W if and only if $\liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{M}^{(n)}| = C_M(W)$, and $\lim_{n \to \infty} \epsilon_n = 0$.

Next theorem was proved by Shannon [44] and gives a single-letter expression for calculating the channel capacity – see also Chapter 3, Section 1 of [45].

Theorem 2.9 (Channel Coding – see Theorem 3.1 of [45]). *The channel capacity of a DMC $W = P_{Y|X}$ is*

$$C_M(P_{Y|X}) = \max_{P_X} I(X;Y).$$

2.2.4 Finite-length Analysis

The channel coding theorem 2.9 characterizes the asymptotic limit of communication rate for a given DMC; however, in practice the channel codes are designed at fixed finite blocklengths. Therefore, the problem of finite-length analysis of channels and channel codes is of practical relevance and specifically has recently received a lot of attention – see e.g., [49–51].

Consider that a DMC $W = P_{Y|X}$ is given. Recall that a channel code family Ψ (indexed by n) is called an $(|\mathcal{M}^{(n)}|, \epsilon_n)$−code if it has error probability ϵ_n and code size M. For a finite $n \in \mathbb{N}$ and a fixed $\epsilon \in (0,1)$ define [50]

$$M_\epsilon(X^n; Y^n) = \max\{\log(|\mathcal{M}^{(n)}|) \mid \exists \text{ an } (|\mathcal{M}^{(n)}|, \epsilon_n)-\text{code } \Psi \text{ with } \epsilon_n \leqslant \epsilon\}.$$

Note that $M_\epsilon(X^n; Y^n)$ as a function of n describes the maximum possible code size of all codes with error probability less than or equal ϵ. It is desired to find expressions for calculating $M_\epsilon(X^n; Y^n)$. The first theorem about $M_\epsilon(X^n; Y^n)$ is due to Shannon where he proved that [44, Theorem 12] for every $\epsilon \in (0,1)$

$$M_\epsilon(X^n; Y^n) = nC_M - o(n),$$

which means that channel capacity can be achieved asymptotically even when error probability does not approach to zero at large n's.

The objective of finite-length analysis for channel coding is to find more accurate approximations of $M_\epsilon(X^n; Y^n)$. The following is due to Strassen – see also [50] Equation (275).

Theorem 2.10 (Second-order Approximation for Channel Coding [52]). *For a given DMC $W = P_{Y|X}$ and $\epsilon \in (0, {}^1/_2)$ we have*

$$M_\epsilon(X^n; Y^n) = nC_M - \sqrt{n\Delta_{\min}}Q^{-1}(\epsilon) \pm \mathcal{O}(\log n),$$

with $\Delta_{\min} := \min_{P_X \in \widetilde{\mathcal{P}}} \text{Var}\{i(X,Y)\}$ where $\widetilde{\mathcal{P}} := \{P_X \mid I(X;Y) = C_M(P_{Y|X})\}$.

The above finite-length result was extended and improved for more general settings by Polyanskiy et al. and Hayashi. For the same practical reasons finite-length analysis of other information-theoretic tasks have recently gained lots of attention [51, 53].
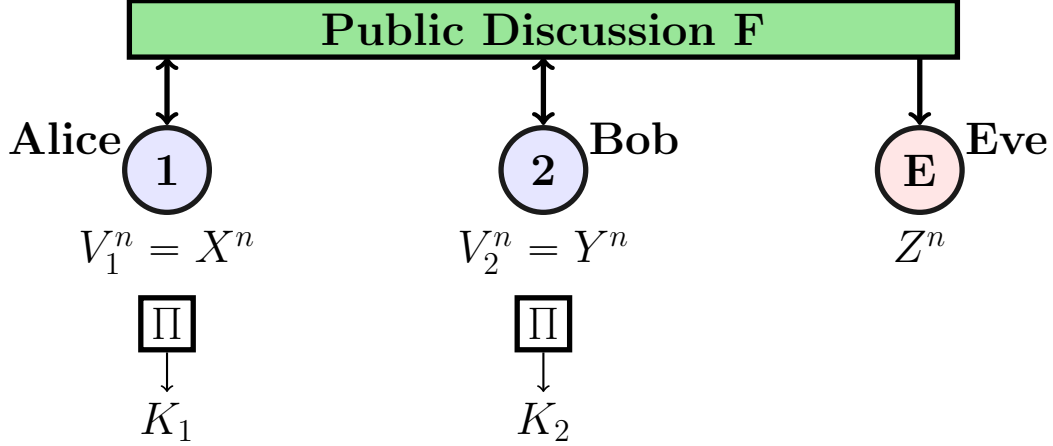
Figure 2.3: Two-party SKA in source model.

2.3   Two-party Secret Key Agreement in Source Model

Information-theoretic secret key agreement was first studied in the two-party scenario by Maurer [19] and independently by Ahlswede and Csiszár [20]. In this section, we review two-party SKA in the source model.

Suppose a 3-DMS[1] (V_1, V_2, Z) with a known probability distribution $P_{V_1 V_2 Z}$ is given. In the two-party source model of SKA, terminal 1 (Alice), terminal 2 (Bob), and the wiretapping adversary Eve, have access to V_1, V_2, and Z, respectively. Variable Z is referred to as the (wiretap) side information of Eve. For the case of two-party SKA, let us use the conventional notations, where $V_1 = X_1 = X$ and $V_2 = Y_2 = Y$ denote the variables of Alice and Bob respectively – i.e., $P_{V_1 V_2 Z} = P_{XYZ}$.

> ***About the Notation of Terminals' Variables***
>
> In this thesis we use V_j when referring to a generic random variable associated with a terminal labeled as j. Sometimes, we change this notation to $V_j = X_j$ or $V_j = Y_j$ or $V_j = (X_j, Y_j)$ depending on the probabilistic model of SKA that we study.

---

[1]A discrete memoryless source with three components.

Some examples of how such correlated variables can be realized in practice are the following.

- **Mutual Channel Estimation.** Alice and Bob have wireless devices that communicate over the same frequency, and they have a "means of estimating their mutual channel." Then the resulting estimations are highly correlated. Also, observations by other wireless devices are essentially uncorrelated if they are at least half a wavelength away from Alice and Bob. See [54, 55].

- **The Satellite Setting.** Consider a radio source that is broadcasting a string of random bits which is received by Alice, Bob, and Eve through independent noisy channels with different error probability. Hence, random samples of Alice, Bob, and Eve are correlated. For example, one can consider that the independent channels from the satellite to Alice, Bob, and Eve to be binary symmetric channels with different bit flip probabilities. See [19, 56].

- **The Wiretap Channel.** In this setting, Alice is broadcasting a sequence of binary values, and Bob and Eve observe noisy versions of Alice's random sequence. The correlation between Alice, Bob, and Eve is then can be modelled by a single-input multiple-out put channel where Alice provides the input symbols, and Bob and Eve observe output variables. See [20, 57].

The goal of Alice and Bob in an SKA protocol is to use correlated variables X and Y, and a noiseless public communication channel, to agree on the largest possible shared secret key. The SKA protocol is desired to be (i) *reliable*, in the sense that the keys obtained by Alice and Bob should be the same; and (ii) *secure*, in the sense that Eve should not learn more than a negligible amount of information about the shared secret key. More thoroughly, a general two-party SKA protocol Π in source model, works as follows. Parties sample from their variables for n times and observe the IID sequences (X^n, Y^n, Z^n). Alice and Bob engage in a (possibly interactive) public discussion $\mathbf{F}$ where Eve observes all of the public

messages. After public discussion, Alice and Bob both compute their estimates of the final key, respectively denoted by K_1 and K_2. See Figure 2.3.

A key generated by an SKA protocol family Π is called an $(\epsilon_n, \sigma_n)-$secret key (or $(\epsilon_n, \sigma_n)-$SK for short) if $\Pr\{K_1 = K_2 = K\} \geqslant 1 - \epsilon_n$ and $\mathbf{SD}(K\mathbf{F}Z^n, U\mathbf{F}Z^n) \leqslant \sigma_n$, where $\mathbf{F}$ denotes the whole transcript of the public messages transfered between Alice and Bob, U is the uniform distribution over alphabet $\mathcal{K}$, and $\mathbf{SD}(\cdot, \cdot)$ is statistical distance[2].

The key rate of an SKA protocol family Π that for every $n \in \mathbb{N}$ generates an $(\epsilon_n, \sigma_n)-$SK K is given by $r_K(\Pi) = \liminf_{n\to\infty} 1/n \log |\mathcal{K}|$. The key rate $r_K(\Pi)$ of the SKA protocol family Π is called "*achievable*" if $\lim_{n\to\infty} \epsilon_n = \lim_{n\to\infty} \sigma_n = 0$.

Definition 2.17. The key capacity of a model is the largest achievable key rate.

In a two-party source model P_{XYZ}, if the adversary is wiretapping (observes the side information Z), the key capacity is called the "*wiretap secret key*" (WSK) capacity, denoted by $C_{WSK}(P_{XYZ})$. If Eve is not wiretapping ($Z = $ constant), then the key capacity is called the "*secret key*" (SK) capacity, denoted by $C_{SK}(P_{XY})$. It is proved that the two-party SK capacity is given by $C_{SK}(P_{XY}) = I(X;Y)$ [20, Proposition 1]. However, the problem of finding an expression for WSK capacity remains open. Next theorem, gives an important upper bound on WSK capacity and proves its tightness under two certain assumptions.

Theorem 2.11 ([20, Theorems 1 and 3]). *Consider a two-party source model P_{XYZ}. Then,*

a) $C_{WSK}(P_{XYZ}) \leqslant I(X;Y|Z)$,

b) *Bound in a is tight if $X - Y - Z$ holds,*

c) *Bound in a is tight if Z is known by the terminals.*

---

[2]In this thesis we use statistical distance to measure secrecy. This is an standard approach in the context of information-theoretic security, see for example [25, 31, 58, 59]. Other definitions that are widely used in the literature are the notions of weak and strong secrecy that employ mutual information between the key and Eve's knowledge (see [19–23, 60]). All of these secrecy definitions are closely related. For a better understanding of the relation between these definitions of secrecy we refer the reader to [21, Lemma 1 and Appendix B], and [61, Section 2.4].

The problem of secret key agreement is tightly related to two other information-theoretic tasks, namely the "privacy amplification" and the "information reconciliation." Before demonstrating that how a shared secret key can be generated between two distant parties, we review these notions in the following subsections.

2.3.1 Privacy Amplification

The Privacy Amplification (or Key Extraction) is the task of extracting a secret key from a variable that is partially leaked to the adversary [62]. Here, we assume that Alice (terminal 1) and Bob (terminal 2) have access to a common randomness V^n. There is a wiretapping adversary Eve who has access to wiretap side information Z^n which is correlated with legitimate parties' common randomness V^n. The goal of a privacy amplification code is to extract the largest possible key $K = f_{PA}(V^n)$ from V^n such that its distribution is close to uniform and it's almost independent from Eve's knowledge (Eve has negligible information about it.)

Next we define universal hash functions, and briefly explain why they are useful in the context of privacy amplification (key extraction) [27].

Definition 2.18 (Universal Hash − [28]). A family of functions $\{h_s : \mathcal{X} \to \mathcal{K}\}_{s \in \mathcal{S}}$ is a *2-Universal Hash Family* if for any $x \neq x'$, $\Pr\{h_S(x) = h_S(x')\} \leqslant \frac{1}{|\mathcal{K}|}$, where the probability is on the uniform choice of $\mathcal{S}$.

Universal hash functions have a lot of applications in cryptography [58, 63, 64], and are well studied [65, 66]. One of their important properties (proved by the following lemma) is that they can be used as randomness extractors [30]. For a survey on the study and detailed analysis of different implementations of universal hash functions and various applications of randomness extractors we refer the reader to [65–67].

Consider the problem of key extraction under the assumption that Eve has no wiretap side information, i.e., $Z = \text{constant}$. Let S be a random value, that Alice and Bob agree on

publicly. Then a function $K = f_{PA}(V, S)$ is considered $\sigma-$secure if $\mathbf{SD}(KS, US) \leqslant \sigma$ where U is the uniform distribution over $\mathcal{K}$.

The Leftover Hash Lemma (LHL) stated below (Lemma 2.12) proves that a 2-Universal Hash Function (2-UHF) is $\sigma-$secure as long as its output key length satisfies

$$\log |\mathcal{K}| \leqslant H_{\min}(V) + \log 4\sigma^2,$$

where $H_{\min}(V) = -\log(\max_v P_V(v))$ is the min-entropy of V.

Lemma 2.12 (Leftover Hash Lemma (LHL) $-$ [30]). *Assume a family of functions* $\{h_s : \mathcal{V} \to \mathcal{K}\}_{s \in \mathcal{S}}$ *is 2-UHF. Then, for any random variable V and uniformly random S*

$$\mathbf{SD}\Big((h_S(V), S), (U, S)\Big) \leqslant \frac{1}{2}\sqrt{|\mathcal{K}|2^{-H_{\min}(V)}},$$

where U is the uniform distribution over $\mathcal{K}$.

In many practical scenarios, though, the common randomness of Alice and Bob is partially leaked to the adversary in the form of a side information Z and a perhaps a public variable F. Therefore, we consider a slightly modified variant of the above privacy amplification problem – see also Lemma 8 of [31]. Suppose the joint probability distribution of a 2-DMS source (V, Z) is known. The variable V (called the common randomness) is accessible to both Alice and Bob, and the variable Z (called the adversary's wiretap side information) is accessible to Eve. Alice, Bob, and Eve, sample from the 2-DMS source for n times and observe IID sequences V^n and Z^n. Moreover, assume that there exists a publicly available variable F, that is correlated with V^n. We call F the public side information. See Figure 2.4.

A privacy amplification function $f_{PA} : \mathcal{V}^n \to \mathcal{K}$ with asymptotic key rate define by $\liminf_{n \to \infty} 1/n \log |\mathcal{K}|$ is called σ_n-secure if $\mathbf{SD}(KFZ^n, UFZ^n) \leqslant \sigma_n$, where U is the uniform distribution over alphabet $\mathcal{K}$. The parameter σ_n is called the secrecy parameter. Note that
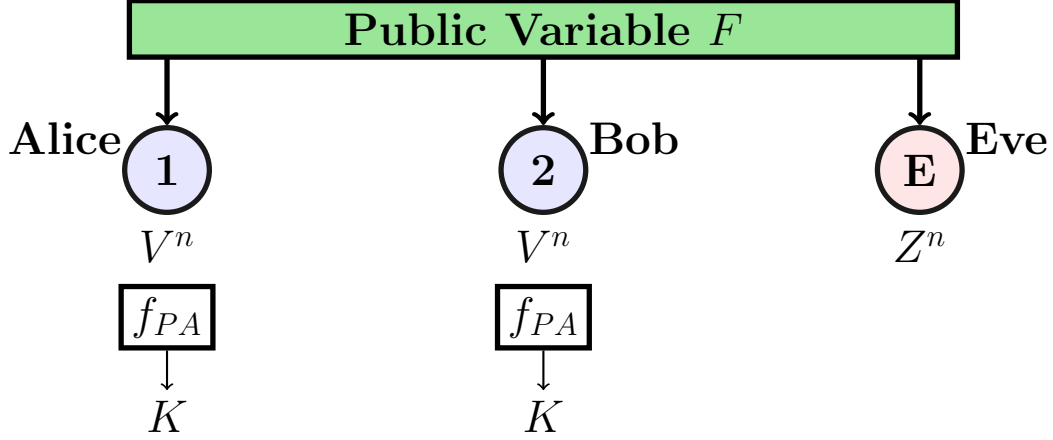
Figure 2.4: The privacy amplification problem with public and wiretap side information available to Eve.

here the key length is measured by $\log |\mathcal{K}|$. Next lemma gives a lower bound on the maximum key rate of all privacy amplification codes with asymptotic perfect secrecy ($\sigma_n \to 0$).

Lemma 2.13 (Privacy Amplification with Public and Wiretap Side information Available to Eve). *Consider the modified privacy amplification problem as stated above. Then, for every $R \in \mathbb{R}$ satisfying*

$$R \leqslant H(V|Z) - \lim_{n \to \infty} \frac{1}{n} \log |\mathcal{F}|,$$

there always exists a σ_n-secure privacy amplification function $f_{PA} : \mathcal{V}^n \to \mathcal{K}$, with asymptotic key rate R and $\lim_{n \to \infty} \sigma_n = 0$.

    Proof sketch: Follows immediately from Lemma 3.14 and Lemma 3.2, by using Universal Hash functions (defined in Definition 2.18) to extract a σ_n-secure key. Full proof is given in the Appendix 2.6 for completeness.   ■

2.3.2   Information Reconciliation

The problem of Information Reconciliation (IR) for the two-party source model is as follows. Suppose a 2-DMS (V_1, V_2) is given. Alice (terminal 1) has access to V_1 and Bob (terminal 2)
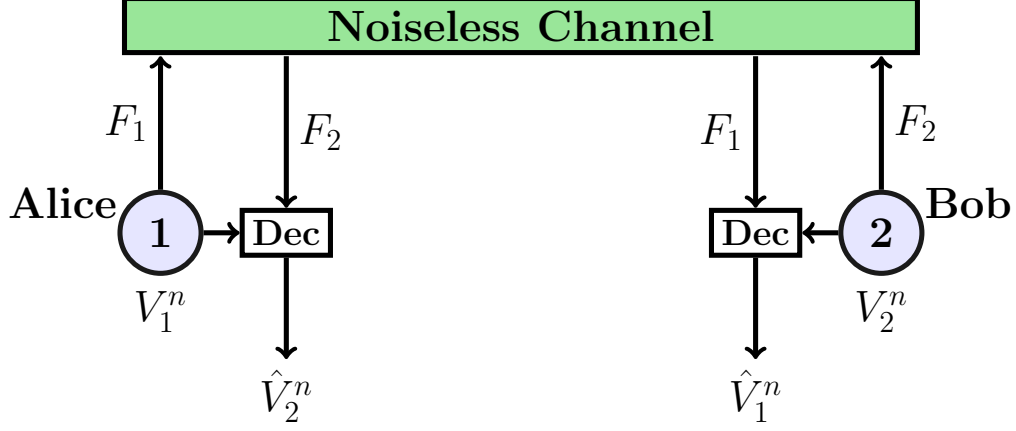
Figure 2.5: Two-party communication for omniscience.

has access to V_2. Alice and Bob sample from their source components for n times and observe the $n-$IID variables (V_1^n, V_2^n) [68, 69]. Through the exchange of a series of (interactive) messages over a noiseless communication channel, Alice and Bob want to arrive at a *common randomness CR* that is a function of (V_1^n, V_2^n), i.e., $CR = CR(V_1^n, V_2^n)$ [68, 69].

This problem is closely related to both the source coding and the channel coding problems and is of particular importance in the context of key agreement [26, 70].

Two notable and well-studied special cases of the IR problem are the following:

- **Communication for Error Correction[3]:** Here, the objective common randomness function is Alice's observation $CR = V_1^n$. This is exactly the same problem as the source coding problem with side information at the decoder. In fact the source coding Theorem 2.6 implies that this task can be done by one-way communication. Let F_1 denote the one-way message of Alice to Bob, which has asymptotic rate $R_1 = \limsup_{n \to \infty} \nicefrac{1}{n} \log |\mathcal{F}_1|$. Then by Theorem 2.6

$$R_1 \geqslant H(V_1|V_2).$$

---

[3]Some works, e.g., [26], define the task of "information reconciliation" as we define this special case when parties want to reconcile on either one's observations.

- **Communication for Omniscience (CO):** The problem of Communication for Omniscience was first introduced in [21]. In this case, terminals send messages to each other (Alice sends F_1 to Bob and Bob sends F_2 to Alice,) to learn each others' observations – i.e., the objective common randomness is $CR = (V_1^n, V_2^n)$. The state of terminals when they learn each other's observation is called "omniscience." It was shown in [21] that omniscience can be achieved noninteractively, i.e., by Alice and Bob each sending each other single independent messages. See Figure 2.5. Let $R_j = \limsup_{n \to \infty} \frac{1}{n} \log |\mathcal{F}_j|$ $j = 1, 2$. Then, by Theorem 2.6

$$R_1 \geqslant H(V_1|V_2),$$
$$R_2 \geqslant H(V_2|V_1).$$

  The problem of Communication for Omniscience can be generalized to case of an m-DMS (multiterminal) source model [21] which we review in Section 2.4.

Both one-way and CO approaches to IR have been well studied and they have direct applications for secret key agreement. In the following section we discuss how these different reconciliation protocols are used for key agreement.

2.3.3 How to Achieve the Key Capacity (When Z is Known)

Consider the two-party SKA model when Z is known to the legitimate terminals. Here we demonstrate how two well known SKA approaches can be proved to be key capacity achieving. We review and compare these important SKA approaches and illustrate how the source coding theorem 2.6 and the privacy amplification lemma 2.13 are used in this context.

Now, let us give the first achievability proof for statement c of Theorem 2.11. Consider the SKA Protocol 1. We show that the asymptotic key rate of Protocol 1 can be as large as $I(X;Y|Z)$, which implies that $C_{WSK}(P_{XYZ}) \geqslant I(X;Y|Z)$ when Z is known.

Protocol 1: A two-party SKA protocol by omniscience for when Z is known ($\mathbf{\Pi_{CO}}$)

Public Information: P_{XYZ}
Input: $n-$IID copies of X, Y, and Z.
Output: Key estimates K_1 and K_2.

1 Alice sends a public message $F_1 = F_1(X^n)$ such that Bob can reliably reconstruct X^n using the knowledge of Y^n and Z^n.

2 Bob sends a public message $F_2 = F_2(Y^n)$ such that Alice can reliably reconstruct Y^n using the knowledge of X^n and Z^n.

3 Alice and Bob extract their respective keys K_1 and K_2 using the common randomness (X^n, Y^n).

SKA by Omniscience. In Protocol 1 terminals first learn each other's initial observations – every legitimate terminal learns all copies of (X, Y). This state of the terminals is called "*omniscience*" and any public discussion protocol that enables omniscience is called a communication for omniscience (CO) protocol.

Alternative Achievability Proof for Theorem 2.11-c: By the privacy amplification Lemma 2.13, as the common randomness is (X^n, Y^n) and the publicly known variables are F_1 and F_2, we get the following lower bound on $C_{WSK}(P_{XYZ})$

$$C_{WSK}(P_{XYZ}) \geqslant H(X, Y|Z) - \lim_{n \to \infty} \frac{1}{n}(\log|\mathcal{F}_1| + \log|\mathcal{F}_2|).$$

The above bound holds for any CO protocol that enables omniscience, thus,

$$C_{WSK}(P_{XYZ}) \geqslant H(X, Y|Z) - \min\left\{\lim_{n \to \infty} \frac{1}{n}(\log|\mathcal{F}_1| + \log|\mathcal{F}_2|)\right\}$$
$$= H(X, Y|Z) - R_{CO}(X, Y|Z),$$

where the minimum on the first line is on all public communication protocols (F_1, F_2) that enable omniscience, and $R_{CO}(X, Y|Z)$ denotes the minimum asymptotic public communication sum rate that is required to achieve omniscience. Theorem 2.6 implies the following

characterization of $R_{CO}(X, Y|Z)$

$$R_{CO}(X, Y|Z) = \min \left\{ R_1 + R_2 \left| \begin{array}{l} R_1 \geqslant H(X|YZ), \\ R_2 \geqslant H(Y|XZ). \end{array} \right. \right\},$$

which, for this case of two-party source model, can be solved as

$$R_{CO}(X, Y|Z) = H(X|YZ) + H(Y|XZ).$$

Therefore, the following lower bound holds on the WSK capacity

$$C_{WSK}(P_{XYZ}) \geqslant H(X, Y|Z) - H(X|YZ) - H(Y|XZ) = I(X; Y|Z). \qquad \blacksquare$$

We note that the original proof given in [20] is based on a different SKA protocol and uses different methods other than Lemma 2.13.

We consider SKA models in which public communication is assumed free; however, it is always desirable to find SKA protocols that use less public communication bits to achieve the key capacity. Let $\mathbf{F}(\Pi)$ denote the public discussion of an SKA protocol Π. Then the asymptotic public communication rate of Π given by $\lim_{n\to\infty} 1/n \log |\mathcal{F}(\Pi)|$ is used to quantify the efficiency of SKA protocol Π with respect to its public communication costs.

In the following discussion we show an alternative approach to achieve the WSK capacity, which asymptotically uses less public communication bits than the SKA by omniscience approach discussed above.

Remark 2.5 (SKA by One-way Public Communication). Protocol 1 ($\mathbf{\Pi_{CO}}$) uses two public messages and (for the two-party model when Z is known) it is not the best SKA approach for achieving the key capacity. A better approach would be to perform key agreement by "*one-way*" public communication. Consider the one-way SKA Protocol 2 ($\mathbf{\Pi_{OW}}$) where only Alice sends a public message F_1 such that Bob can reliably reconstruct

---

Protocol 2: A two-party one-way SKA protocol for when Z is known ($\mathbf{\Pi_{OW}}$)

    Public Information: P_{XYZ}
    Input: $n-$IID copies of X, Y, and Z.
    Output: Key estimates K_1 and K_2.

1 Alice sends a public message $F_1 = F_1(X^n)$ such that Bob can reliably reconstruct X^n using the knowledge of Y^n and Z^n.

2 Alice and Bob extract their keys K_1 and K_2 using the common variable X^n.

X^n using the knowledge of Y^n and Z^n. Bob is silent – does not send public messages. Then, parties use their common randomness X^n for key extraction. Similar to the analysis we did before, one can easily observe that by Theorem 2.6 and Lemma 2.13, the asymptotic key rate of the one-way SKA protocol is, for this model, given by $H(X|Z) - H(X|YZ) = I(X;Y|Z)$; meaning that the one-way SKA also achieves the WSK capacity when Z is known. The one-way SKA protocol is not only capacity achieving (in this case) but also more efficient in terms of public communication cost than the SKA Protocol 1, that uses the omniscience approach. The asymptotic public communication rate of Protocol 1 is given by $R_{CO} = H(X|YZ) + H(Y|XZ)$ while asymptotic public communication rate of the one-way SKA Protocol 2 is equal to $H(X|YZ) \leqslant R_{CO}$.

Moreover, for the case of Theorem 2.11-b where $X - Y - Z$ holds and Z is not know by the legitimate terminals, it is easy to observe that one-way SKA approach achieves the WSK capacity $I(X;Y|Z)$; though, SKA by omniscience is not WSK capacity achieving. In [20], one-way SKA is used to prove Theorem 2.11-b and c. In Chapter 3, we focus our study on two-party one-way secret key agreement protocols for the general case when Z is *not* known.

2.4   Multiterminal SKA in Source Model

In a series of seminal papers, Csiszár and Narayan generalized previous results of Maurer and Ahlswede and Csiszár on two-party SKA to the case of multiterminal SKA [21–23].

Suppose we have a set of m terminals denoted by $\mathcal{M} = \{1, 2, \ldots, m\}$. Consider a discrete

memoryless multiple source (DMMS) denoted by $(V_1, V_2, \ldots, V_m, Z)$ where each component variable $V_j \in \mathcal{M}$ is observed by a different terminal $j \in \mathcal{M}$, and side information Z is observed by Eve. Let us use the conventional notation of $V_j = X_j$ for the case of multiterminal source model. The probability distribution of the DMMS is known and is denoted by $P_{X_{\mathcal{M}}Z}$, where $X_{\mathcal{M}}$ denotes the set of all legitimate terminals' variables. Terminals use a noiseless public communication channel and n IID copies of their variables to establish a shared secret key among terminals of a subset $\mathcal{A} \subseteq M$. Terminals that are not in $\mathcal{A}$ are called helper terminals. See Figure 2.6.

Similar to the case of two-party setting, the key capacity of a multiterminal source model is defined as the largest acheiavle key rate. See Definition 2.17.

The general multiterminal source model was introduced in [21] and three types of key capacities were considered based on the assumptions regarding Eve's side information:

- **Secret Key (SK) capacity**: Eve has no side information, i.e., $Z = \text{constant}$. In this case, the model is called non-wiretapped.

- **Private Key (PK) capacity**: Eve has compromised a subset of helper terminals $\mathcal{D} \subseteq \mathcal{A}^c$ and sees their observations. In this case, the model is called compromised. It is assumed that compromised terminals publicly reveal $Z = X_{\mathcal{D}}$.

- **Wiretap Secret Key (WSK) capacity**: Eve has access to a side information Z. In this case, the model is called wiretapped.

The WSK capacity is the most general key capacity notion and analysis of PK capacity is useful for modeling the special case of WSK capacity when Z is assumed to be known. For example, the two-party scenario when Z is known (statement c of Theorem 2.11) is equivalent with a three-terminals scenario where terminals' variables are $V_1 = X$, $V_2 = Y$ and $V_3 = Z$, $\mathcal{A} = \{1, 2\}$, and terminal 3 is assumed compromised.

Single-letter expressions for the SK and PK capacities of the general multiterminal source model were given in [21], and it was proved that these capacities can be achieved using the

44
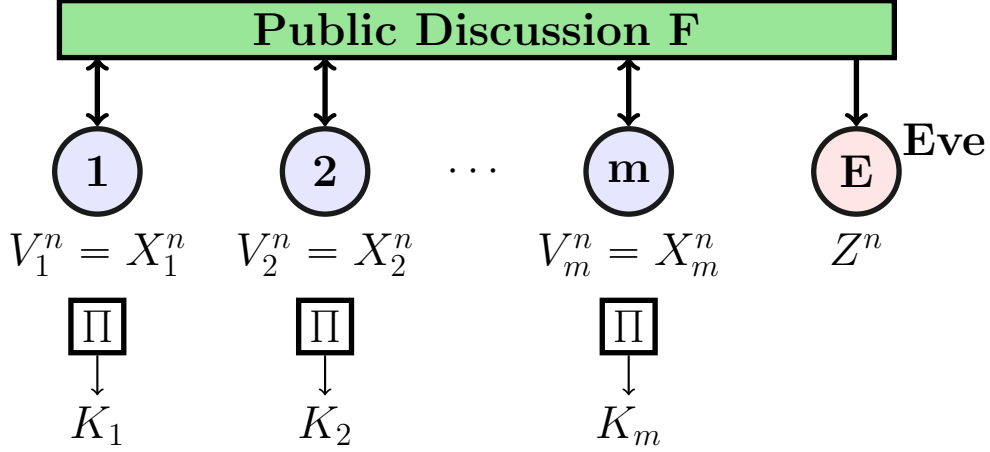
Figure 2.6: Multiterminal SKA in source model.

SKA by omniscience method.

Theorem 2.14 (Multiterminal Source Model SK Capacity – [21, Theorem 1]).
Consider a multiterminal source model $P_{X_{\mathcal{M}}}$. Then, for any $\mathcal{A} \subseteq \mathcal{M}$

$$C_{SK}^{\mathcal{A}}(P_{X_{\mathcal{M}}}) = H(X_{\mathcal{M}}) - R_{CO}^{\mathcal{A}}(X_{\mathcal{M}}),$$

where $R_{CO}^{\mathcal{A}}(X_{\mathcal{M}}) = \min\{\sum_{j=1}^{m} R_j | \sum_{i \in \mathcal{B}} R_i \geqslant H(X_{\mathcal{B}}|X_{\mathcal{B}^c}), \ \forall \mathcal{B} \subset \mathcal{M}, \mathcal{A} \nsubseteq \mathcal{B}\}$ is the minimum asymptotic public communication sum rate that is required for terminals in subset $\mathcal{A}$ to achieve omniscience (learn $X_{\mathcal{M}}$).

The SKA protocol that achieves the SK capacity above, has the following main two steps:

1- Terminals use public discussion so terminals of $\mathcal{A}$ achieve omniscience (learn $X_{\mathcal{M}}^n$).

2- Terminals in $\mathcal{A}$ use their common randomness $X_{\mathcal{M}}^n$ to extract their keys.

Note that Lemma 2.13 immediately implies that $C_{SK}^{\mathcal{A}}(P_{X_{\mathcal{M}}}) \geqslant H(X_{\mathcal{M}}) - R_{CO}^{\mathcal{A}}(X_{\mathcal{M}})$.

Remark 2.6 (When SKA by omniscience is not efficient). While a lot of proposed SK and PK capacity achieving protocols are base on the communication for omniscience

$$V_2 = X_2 = (V_{12}, V_{23})$$

$$V_1 = X_1 = (V_{12}, V_{13})$$

$$V_3 = X_3 = (V_{23}, V_{13})$$
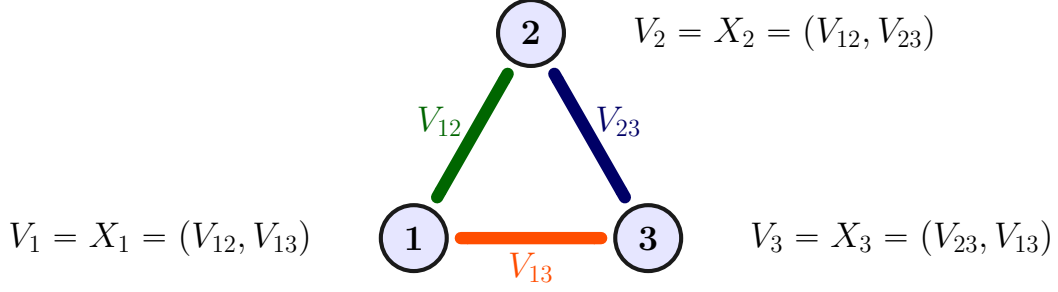
Figure 2.7: A simple multiterminal source model.

approach [21, 71, 72], they are *not* always efficient in terms of public communication. Consider the following example. There are three terminals $\mathcal{M} = \{1, 2, 3\}$, and we want to do key agreement for all $(\mathcal{A} = \mathcal{M})$. The RV's of these terminals have the following specific structure: $X_1 = (V_{12}, V_{13})$, $X_2 = (V_{12}, V_{23})$, $X_3 = (V_{23}, V_{13})$, where each V_{ij} is an independent binary uniform random variable. See Figure 2.7.

Suppose terminals want to achieve omniscience (everyone learns $X_{\mathcal{M}} = (V_{12}, V_{23}, V_{13})$). Then, for example, they can execute these steps:

1- Terminal 2 sends public message $F_2 = V_{12} \oplus V_{23}$,

2- Terminal 3 sends public message $F_3 = V_{23} \oplus V_{13}$.

Here, $\oplus$ denotes the binary XOR operation. Thus, using only 2 bits of communication, terminals could attain omniscience. In fact, the minimum communication rate for omniscience, in this example, is 2 bits. Therefor, for key agreement, terminals can follow the steps above and then agree on $K = V_{12}$. By Theorem 2.14 we can calculate the SK capacity as

$$
\begin{aligned}
C_{SK} &= H(X_1 X_2 X_3) - R_{CO} \\
&= H(V_{12}) + H(V_{23}) + H(V_{13}) - R_{CO} \\
&= 3 - 2 = 1.
\end{aligned}
$$

However, note that to agree on 1 bit of key per sample, terminals do not need to achieve

---

Protocol 3: A three-party SKA protocol for source model of Figure 2.7 ($\mathbf{\Pi_{XOR}}$)

> **Public Information:** $P_{X_1X_2X_3}$
> **Input:** A single copy of $X_1 = (V_{12}, V_{13})$, $X_2 = (V_{12}, V_{23})$, and $X_3 = (V_{23}, V_{13})$.
> **Output:** Key estimates K_1, K_2 and K_3.

---
1   Terminal 2 sends a public message $F_2 = V_{12} \oplus V_{23}$.
2   Terminals 1 and 2 set their final key to $K_j = V_{12} \forall j = \{1, 2\}$ and Terminal 3
    calculates the final key by $K_3 = V_{23} \oplus F_2$.

---

omniscience. For instance, if only Terminal 2 sends $F_2 = V_{12} \oplus V_{23}$, (and other terminals remain silent,) all terminals still can agree on $K = V_{12}$. Terminals 1 and 2 have prior knowledge of $K = V_{12}$ and terminal 3 computes the key by $K = V_{23} \oplus F_2$. See Protocol 3 ($\mathbf{\Pi_{XOR}}$.)

Hence, similar to the case of two-party SKA (see Remark 2.5), we can conclude that the approach of SKA by omniscience is not always optimal from the view point of using public communication. On the other hand; however, we note that SKA by omniscience is the only known approach that achieves the SK and PK capacities of any arbitrary source model.

The multiterminal source model WSK capacity is known when Eve's side information is revealed publicly to all terminals. Next theorem is a generalization of Theorem 2.11-c.

Theorem 2.15 ([21, Theorem 4]). *Consider a multiterminal source model $P_{ZX_\mathcal{M}}$. Then, for any $\mathcal{A} \subseteq \mathcal{M}$, if Z is known by the terminals*

$$C_{WSK}^\mathcal{A}(P_{X_\mathcal{M}Z}) = H(X_\mathcal{M}|Z) - R_{CO}^\mathcal{A}(X_\mathcal{M}|Z),$$

where $R_{CO}^\mathcal{A}(X_\mathcal{M}|Z) = \min\{\sum_{j=1}^m R_j | \sum_{i \in \mathcal{B}} R_i \geqslant H(X_\mathcal{B}|X_{\mathcal{B}^c}Z), \ \forall \mathcal{B} \subset \mathcal{M}, \mathcal{A} \nsubseteq \mathcal{B}\}$ is the minimum asymptotic public communication sum rate that is required for terminals in subset $\mathcal{A}$ to achieve omniscience (learn $X_\mathcal{M}$ in addition to the common variable Z). Moreover, the above expression is an upper bound to WSK capacity when Z is not known by the terminals.

> The problem of finding an expression for the general WSK capacity of multiterminal source model is unsolved.

In Chapter 4, we introduce a large subclass of multiterminal source models, called "*wire-tapped Tree-PIN*," and prove its WSK capacity. This is one of the first WSK capacity results for a large subclass of multiterminal source models. In the special case when Z is known, our proposed capacity achieving SKA protocol is provably more efficient than SKA protocol of [21] that uses the CO approach.

2.5 Secret Key Agreement in Channel Model

In a multiterminal channel model for SKA, terminals are connected by a noisy discrete memoryless channel (DMC) that is used for generating correlation. Terminals also have unlimited access to a noiseless public channel that can be used before, during, and after symbol transmissions over the DMC. The underlying DMC might be wiretapped, which is modeled by providing an output variable Z to the adversary, Eve. Let $\mathcal{M} = \{1, \ldots, m\}$ denote the set of terminals and let $\mathcal{A} \subseteq \mathcal{M}$ be the subset of terminals that want to agree on a shared secret key. Terminals use the DMC for n times and at the end of SKA protocol, each terminal j in $\mathcal{A}$, computes their estimate K_j of the final key.

A key generated by a channel model SKA protocol Π is called an $(\epsilon_n, \sigma_n)-$secret key (or $(\epsilon_n, \sigma_n)-$SK for short) if $\Pr\{K_j = K\} \geqslant 1 - \epsilon_n \quad \forall j \in \mathcal{A}$, and $\mathbf{SD}(K\mathbf{F}Z^n, U\mathbf{F}Z^n) \leqslant \sigma_n$, where $\mathbf{F}$ denotes the whole transcript of the public messages exchanged between the terminals, and U is the uniform distribution over alphabet $\mathcal{K}$.

The key rate of an SKA protocol Π that for every $n \in \mathbb{N}$ generates an $(\epsilon_n, \sigma_n)-$SK K is given by $r_K(\Pi) = \liminf_{n \to \infty} 1/n \log |\mathcal{K}|$. The key rate $r_K(\Pi)$ of the SKA protocol Π is called achievable if $\lim_{n \to \infty} \epsilon_n = \lim_{n \to \infty} \sigma_n = 0$. The key capacity of a channel model is defined as the largest achievable key rate. Same as Definition 2.17.
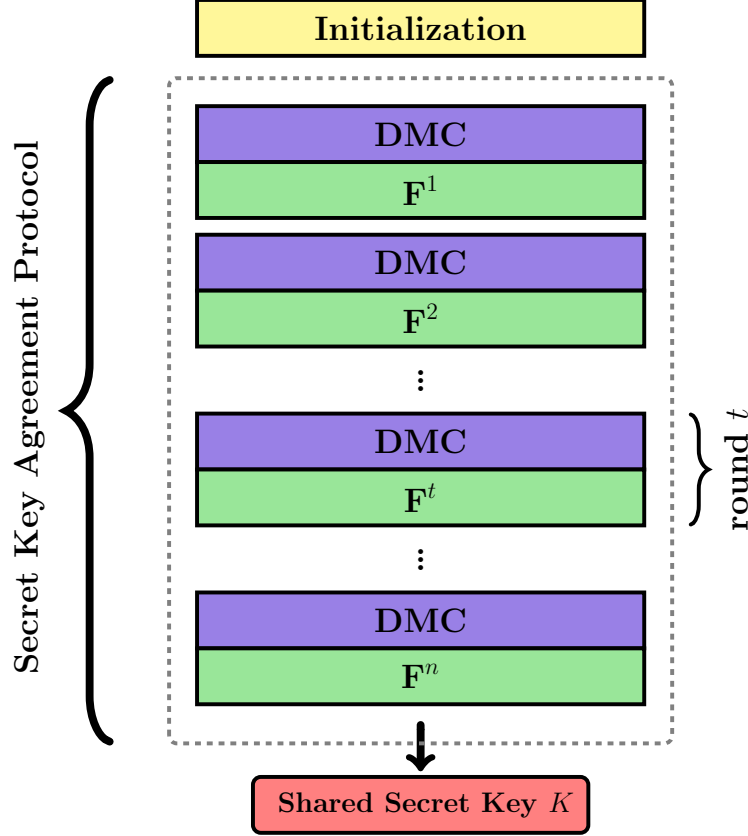
Figure 2.8: General structure of an SKA protocol in channel model.

Similar to the case of multiterminal source model, we define three types of key capacities: SK, PK, and WSK. The main difference between channel model and source model is that in channel model (some) terminals have the privilege to govern the input symbols of the DMC, possibly adaptively based on observing the public feedback messages that other terminals send in between each symbol transmission; whereas in source model, terminals do not have control over the established correlated $n-$fold variables. See Figure 2.8.

The two-party channel model when there is no wiretapping adversary is described by the DMC with conditional probability distribution $W = P_{Y|X}$, where $V_1 = X$ and $V_2 = Y$ denote the input variable of terminal 1 and output variable of terminal 2; respectively.

It is well known that [20, Proposition 1],

$$C_{SK}(P_{Y|X}) = \max_{P_X} I(X;Y),$$

which can be achieved without sending any public messages.

The two-party model with a wiretapper is given by $W = P_{YZ|X}$, where Z is Eve's side information. The two-party WSK capacity is not known; however, similar to Theorem 2.11, an important upper bound and two special case results were given in [20].

Theorem 2.16 ([20, Theorems 2 and 3]). *For any two-party channel model $P_{YZ|X}$,*

a) $C_{WSK}(P_{YZ|X}) \leqslant \max_{P_X} I(X;Y|Z)$,

b) Bound in a is tight if $X - Y - Z$ holds,

c) Bound in a is tight if Z is known by the terminals.

In [22], the above results were extended to a specific multiterminal channel, where the underlying DMC is assumed to be single-input multi-output, where terminal 1 governs the input symbol $V_1 = X_1$ of the DMC, and terminals in $\{2, 3, \ldots, m\}$ observe the output variables $V_2 = Y_2, V_3 = Y_3, \ldots, V_m = Y_m$. See Figure 2.9-(a).

Theorem 2.17 ([22, Theorem 4.1]). *Consider a single-input multi-output channel model $P_{Y_2 Y_3 \ldots Y_m | X_1}$. Then, for any $\mathcal{A} \subseteq \mathcal{M}$*

$$C_{SK}^{\mathcal{A}}(P_{Y_2 Y_3 \ldots Y_m | X_1}) = \max_{P_{X_1}} \{ H(V_{\mathcal{M}}) - R_{CO}^{\mathcal{A}}(V_{\mathcal{M}}) \},$$

where $R_{CO}^{\mathcal{A}}(V_{\mathcal{M}}) = \min\{\sum_{j=1}^{m} R_j | \sum_{i \in \mathcal{B}} R_i \geqslant H(V_{\mathcal{B}} | V_{\mathcal{B}^c}), \ \forall \mathcal{B} \subset \mathcal{M}, \mathcal{A} \nsubseteq \mathcal{B}\}$ is the minimum asymptotic public communication sum rate that is required for terminals in subset $\mathcal{A}$ to achieve omniscience (learn $V_{\mathcal{M}}$).

For the SK capacity above, the proposed capacity achieving SKA protocol is based on the "*source emulation*" approach, that starts by (i) sending IID symbols X_1^n through the
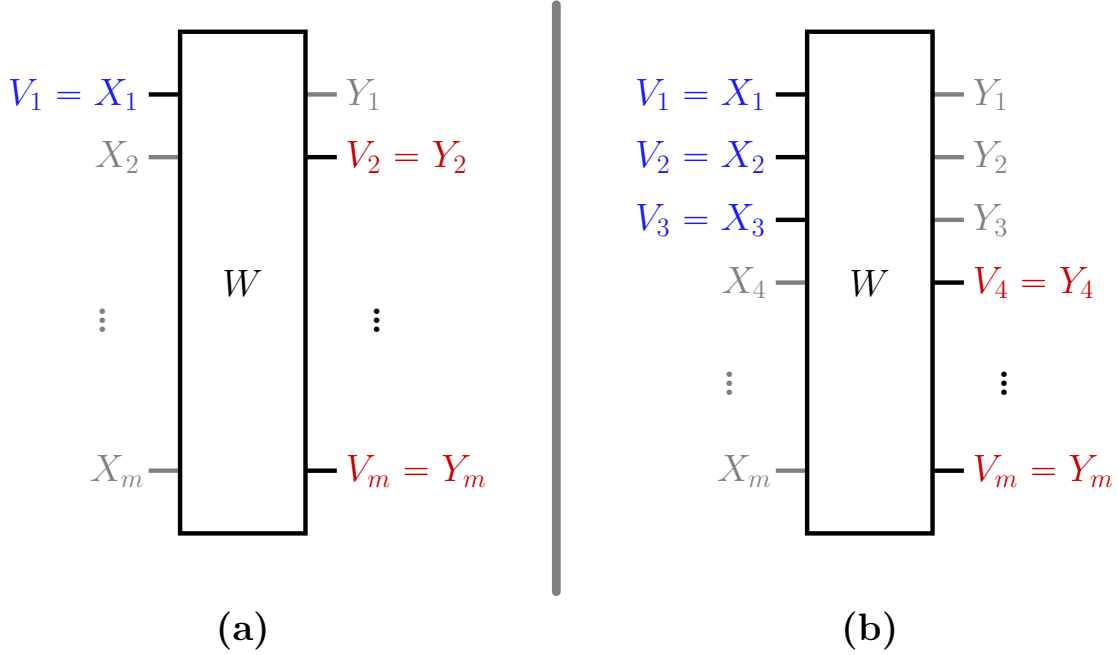
Figure 2.9: (a) The single-input multi-output channel model of [22] and (b) The multiaccess channel model of [23] – here, $\mathcal{I} = \{1, 2, 3\}$.

noisy DMC, which emulates (realizes) a multiterminal source model with joint distribution $P_{V_{\mathcal{M}}} = P_{X_1} P_{Y_2 Y_3 \dots Y_m | X_1}$, and then (ii) employing the source model SKA of [21] (see passage following Theorem 2.14).

The case of secret key agreement in a particular model of multi-input multi-output channels, where a subset $\mathcal{I} \subsetneq \mathcal{M}$ of terminals are input terminals, and the rest of terminals (in $\mathcal{M} \backslash \mathcal{I}$) are output terminals, was considered in [23]. See Figure 2.9-(b). This channel model was called the "*multiaccess*" model, and general upper and lower bounds were proved for the SK and PK capacities. The source emulation approach is proved to be, in general, *not* capacity achieving for the case of multiaccess model [73].

> Finding expressions for the SK, PK, and WSK capacities of the multiaccess channel model remain as open questions.

The multiaccess channel model fails to account for the real-life scenarios where individual

51

terminals can send inputs to *and* receive outputs from the underlying noisy channel. In Chapter 5, we propose a new general channel model that allows the presence of *"transceiver"* terminals that can both send to and receive from the DMC. In Chapter 6, we prove the SK, and WSK capacities of a special class of such transceiver models, called *"Polytree-PIN."*

2.6 Appendix

Lemma 2.13 follows directly from Lemmas 3.14 and 3.2, that we will prove in Chapter 3. A proof is given below for completeness.

Proof of Lemma 2.13: A family of functions $\{h_s : \mathcal{V} \to \mathcal{K}\}_{s \in \mathcal{S}}$ is a *2-Universal Hash Family* if for any $v \neq v'$, $\Pr\{h_S(v) = h_S(v')\} \leqslant \frac{1}{|\mathcal{K}|}$, where the probability is on the uniform choice of $\mathcal{S}$ (see Definition 2.18). To extract the key from common randomness V^n, Alice and Bob first agree on a random seed $s \in \mathcal{S}$. Then they use the hash function $h_s(V^n)$ witch results in a key that has length $\log |\mathcal{K}|$. From Lemma 3.14 and Lemma 3.2, we can conclude that by using $K = h_s(V^n)$ a σ_n−secure key K can be achieved as long as

$$\frac{1}{n}\log |\mathcal{K}| \leqslant H(V|Z) - \frac{1}{n}\log |\mathcal{F}| - \frac{c_1}{\sqrt{n}}Q^{-1}(\sigma_n - \mu_n) + \frac{1}{n}\log 4\mu_n^2 \pm \frac{c_2}{n}$$

where $Q(\cdot)$ is the tail probability of the standard Gaussian distribution, $0 < \mu_n \leqslant \sigma_n$, and c_1 and c_2 are constants that do not depend on n. Let $\sigma_n = 1/n$ and $\mu_n = 1/2n$, and note that $Q^{-1}(x) = \sqrt{2}\mathrm{erfc}^{-1}(2x) \ \forall \ x \in (0,1)$. Here, 'erfc' is the complementary error function, defined as $\mathrm{erfc}(a) = 1 - \frac{2}{\sqrt{\pi}}\int_0^a e^{-t^2}\mathrm{d}t$. Then, we have

$$\frac{1}{n}\log |\mathcal{K}| \leqslant H(V|Z) - \frac{1}{n}\log |\mathcal{F}| - \frac{\sqrt{2}c_1}{\sqrt{n}}\mathrm{erfc}^{-1}(\frac{1}{n}) - \frac{1}{n}4\log n \pm \frac{c_2}{n}.$$

By noting[4] that as $n \to \infty$ we have $\mathrm{erfc}^{-1}(\frac{1}{n}) \approx \sqrt{\frac{\log n}{\log e} - \mathcal{O}(\log \log n)} + \mathcal{O}(n^{-2})$, we have proved that $(1/n)$−secure keys with rates as large as

$$R \leqslant H(V|Z) - \lim_{n \to \infty} \frac{1}{n}\log |\mathcal{F}|$$

can be achieved.   ∎

---

[4]See Wolfram Research's documentation on $\mathrm{erfc}^{-1}(\cdot)$ function at https://functions.wolfram.com/PDF/InverseErfc.pdf.

Chapter 3

Finite-length Lower and Upper Bounds for Two-party One-way Secret Key Agreement

Abstract. Two-party secret key agreement is a fundamental problem in cryptography: Alice wants to share a secret random string called the *key* with Bob, such that a third party, Eve, has negligible information about it. In this chapter, we consider an important information-theoretic model of two-party secret key agreement, in which Alice, Bob, and Eve have access to correlated random variables. In order to make local key generation (at Alice's and Bob's sites) possible, Alice is allowed to send Bob only a single public message that is computed based on her initial random variable. This model is called the *"two-party model of one-way secret key agreement (OW-SKA)"*. In practice, parties sample from their random variables up to a finite number of times, therefore, we aim to give more accurate finite-length expressions for the maximum key length that can be generated. To this end, we prove a new finite-length upper bound, and propose two OW-SKA protocols that imply multiple finite-length lower bounds for the maximum achievable key length. We also compare our OW-SKA protocols with related past results and discuss remaining open research questions.

3.1 Introduction

Information-theoretic secret key agreement (SKA) was first introduced and studied in [19], and independently in [20] and since then has been studied extensively in different models (see [74] and references therein). An important model of SKA that was originally studied in [20] and [19], is the two-party source model with wiretapper in which there are two legitimate parties, Alice (Terminal 1) and Bob (Terminal 2), that want to establish a shared key in the presence of an eavesdropping adversary Eve. Alice, Bob, and Eve have access to variables $V_1 = X, V_2 = Y$, and Z respectively, and the joint public probability distribution P_{XYZ}. The legitimate parties use a reliable public communication channel to arrive at a shared key. This public communication can be interactive and have multiple rounds. The final common secret key of Alice and Bob must be computed in such a way that Eve, who has access to all public communication and the side information Z, learns no information about it.

The primary performance metric for SKA protocols is the achievable length of secret key. Hence, we are interested in finding $S(X; Y|Z)$, the maximum length of secret keys that can be achieved by legitimate parties, given a source model distribution P_{XYZ}. Allowing parties to have access to n independent and identically distributed (IID) repetitions of their respective variables, we can define the wiretap secret key (WSK) capacity, that is the maximum achievable key rate (that is key length divided by n).

A single-letter characterization of the general WSK capacity remains unresolved. However, the key capacity is known under additional assumptions. If we restrict Alice and Bob to only use a single public message (with a finite but unlimited length), from Alice to Bob, the key capacity is known (Theorem 1 of [20]) and is called the one-way wiretap secret key (OW-WSK) capacity, which we denote by $C_{WSK}^{\rightarrow}(P_{XYZ})$. When Markov Relation $X - Y - Z$ holds, the SK capacity can be achieved noninteractively (i.e., by one-way public communication) and it is proved that $C_{WSK}(P_{XYZ}) = C_{WSK}^{\rightarrow}(P_{XYZ}) = I(X; Y|Z)$ (Theorem 2 of [20]). The best known upper and lower bounds on the general WSK capacity are due to [75].

In this chapter, we focus on the "*one-way secret key agreement*" (OW-SKA) source model

in which Alice sends a single message over the public channel to Bob, so they can arrive at a shared key. The OW-SKA problem is important in practice because it avoids the interaction between Alice and Bob that in addition to longer time and more complexity to arrive at an established key, would require stateful protocols which would introduce vulnerabilities in implementation. As noted in [76], the problem is also theoretically interesting because of its relation to circuit polarization and immunization of public-key encryption in complexity theory and cryptography, and its application to oblivious transfer [77], an important cryptographic primitive.

Previous key capacity results prove that $S(X^n; Y^n|Z^n) = nC_{WSK}(P_{XYZ}) + o(n)$ [19, 20]. In practice however, n is finite and a more accurate finite-length approximation of $S(X^n; Y^n|Z^n)$ is needed. For the special case when $X - Y - Z$ holds, Hayashi, Tyagi, and Watanabe [31] proved the second-order asymptotic expansion of $S(X^n; Y^n|Z^n)$ as

$$S(X^n; Y^n|Z^n) = nC_{WSK}(P_{XYZ}) - \sqrt{n}G \pm \mathcal{O}(\log n),$$

where G is a function of probability distribution P_{XYZ} and does not depend on n. They also gave a capacity achieving interactive protocol that asymptotically attains the above second-order key length approximation. The protocol, however, requires $\mathcal{O}(n)$ rounds of interactive public communication, and is not computationally efficient [31].

The finite-length bounds for OW-SKA, however, are less understood. For a finite n for a OW-SKA model with a given source model distribution P_{XYZ}, let $S^{\rightarrow}(X^n; Y^n|Z^n)$ denote the maximum length of secret key that can be achieved via one-way public communication. Finding accurate finite-length approximations of $S^{\rightarrow}(X^n; Y^n|Z^n)$ is also an important problem from the practical point of view. The construction and the bounds in [31] are derived for the general case where the interaction over the public channel is allowed, and is critically used by the optimal protocol. It is not known how tight their upper bound is if communication is restricted to one-way. Known OW-SKA constructions have used random binning

[20], random linear codes and Reed-Solomon codes [76, 78], polar codes [79, 80], and hash functions [26, 81]. The explicit constructions in [26, 78–81] are all capacity achieving one-way constructions, and their finite-length analysis can give a lower bound on achievable SK length, although these bounds have not been explicitly derived, and in some cases deriving the finite-length bound (e.g. in [79, 80]) is not well studied. See Section 3.4.3 for more details and comparison of the achievable key length of these protocols.

In our recent works [32, 33], we have proved new finite-length upper and lower bounds on $S^{\rightarrow}(X^n; Y^n | Z^n)$. However, a tight second-order approximation of $S^{\rightarrow}(X^n; Y^n | Z^n)$ remains unknown. This chapter covers these findings in detail. In the following, we give a brief overview of the claimed results.

3.1.1 Our Work

The contributions of this chapter falls under two parts: (a) Upper Bound, (b) Lower Bounds.

In the first part (Section 3.3), we use smooth Rényi entropy that was introduced in [26], and the information spectrum approach of [24, 25], and define new entropies (Definition 3.4) that allow us to derive a multi-letter upper bound on $S^{\rightarrow}(X^n; Y^n | Z^n)$ (Corollary 3.4.1).

In the second part (Section 3.4), we derive three single-letter finite-length lower bounds on $S^{\rightarrow}(X^n; Y^n | Z^n)$ (Theorems 3.8, 3.9, 3.13). These lower bounds are proved by introducing and analyzing two OW-SKA protocols. Our proposed OW-SKA, both have two main steps: *information reconciliation* (IR) for arriving at a common string, and *privacy amplification* (PA) where the goal is to extract a secret key from the shared string. An initiation phase is included in the protocol during which protocol parameters and public values are determined.

The first protocol ($\mathbf{\Pi_{HH}}$), leads to a finite-length lower bound on the key length, that is tighter than the achievable key length of all the previously known OW-SKA protocols (Theorem 3.9). This OW-SKA protocol uses universal hashing [28, 30] for both of the steps of information reconciliation and privacy amplification. Using universal hash functions for reconciliation has been proposed before in [26, 27, 58, 82]. The novelty of our work is in

using the information spectrum technique of [24] to allow Bob to recover the Alice's variable with high probability. Using the information spectrum approach for reconciliation also decreases the complexity of the proposed protocol in comparison to the similar information reconciliation algorithms of [26, 58, 82] by decreasing the number of hash calculations. An important property of our approach is that it allows us to use the general AEP (Asymptotic Equipartition Property) of [83] that is used for independent experiments (independent but not identically distributed samples), and the Berry–Esseen theorem [84] (first used for channel coding in [50]) for IID (independent and identically distributed) distributions, to obtain two finite-length lower bounds for secret key length, one for independent experiments and one for IID distributions that define the source distribution for our SKA protocol. Both of these lower bounds have the form of $S^{\rightarrow} \geqslant nC_{WSK}^{\rightarrow} - \mathcal{O}(\sqrt{n})$. Thus, our protocol can also be used for the more general case of independent experiments. In fact, we prove that $\mathbf{\Pi_{HH}}$ not only achieves the OW-WSK capacity when parties observations are IID, but also achieves the WSK capacity when parties observations are drawn independently and not necessarily from the same distribution under the assumption that Markov relation $X^n - Y^n - Z^n$ holds. While the OW-SKA protocol $\mathbf{\Pi_{HH}}$ is very efficient in terms of public communication cost, its computational complexity for Alice and Bob are in $\mathcal{O}(n \log n)$ and $\mathcal{O}(2^n)$, respectively. This makes $\mathbf{\Pi_{HH}}$ not computationally efficient. Note that in the context of SKA we call an SKA protocol *efficient* if it has polynomial computational complexity ($\mathcal{O}(n^d)$), and we call an SKA protocol *practically efficient* if it has linear or quasi-linear computational complexity ($\mathcal{O}(n)$ or $\mathcal{O}(n \log n)$) [79].

Observing the computational inefficiency of $\mathbf{\Pi_{HH}}$, we propose an alternative protocol ($\mathbf{\Pi_{PH}}$) that has computational complexity of $\mathcal{O}(n \log n)$ for both Alice and Bob. This protocol uses Polar Coding [29] in the information reconciliation step, and for privacy amplification it employs universal hashing. Polar Coding was previously used for OW-SKA in [79, 80]; however, these results did not give finite-length analysis of their proposed schemes. The novelty of our analysis is in using information spectrum methods to prove a generalized version

of the Leftover Hash Lemma (Spectral LHL 3.14). In Theorem 3.13, we use the Spectral LHL to derive the achievable finite key length of our proposed protocol $\mathbf{\Pi_{PH}}$, which is *near-optimum* as its second-order term is in $\mathcal{O}(\sqrt[\tau]{n^{\tau-1}})$, where $\tau > 2$. Note that in the best known lower bound the second-order term is in $\mathcal{O}(\sqrt{n})$; i.e., $\tau = 2$. The OW-SKA protocol $\mathbf{\Pi_{PH}}$ is very efficient in terms of public communication cost and computational complexity. In Section 3.4.3, we show that our proposed OW-SKA protocols gives way better finite-length performance than the protocols of [76, 78].

3.1.2 Related Works

The analysis of an SKA protocol is in general reduced to two separate phases namely, *information reconciliation* and *privacy amplification*. A lower bound on the key length of SKA protocols can be derived by using an upper bound on the amount of leaked information in the information reconciliation phase, and a matching lower bound on the maximum amount of extractable random bits in the privacy amplification phase. Non-constructive (closely matching) upper and lower bounds for information reconciliation are given in [26]. These bounds are in terms of smooth max entropy that is approximated using Theorem 1 of [83] for finite-length regime. This approximation has been made more precise in [82]. The upper bound on the amount of leaked information during information reconciliation in [26] can be achieved by a protocol that uses universal hashing (see [58, Lemma 6.3.4]). The application of universal hashing for privacy amplification results in an immediate lower bound on the length of the extractable key by using *Leftover Hash Lemma* (LHL) [30] (and its variations – see the bounds given in [25, 26, 65, 85]). The lower bound in [65] is shown to be strictly smaller than the lower bound of [26]. However, the bound of [65] is in terms of entropy quantities for which there is no known finite-length approximation.

The problems of information reconciliation [82] and privacy amplification [27, 62, 65, 85] are of independent interest due to their applications to other information-theoretic tasks such as secure random number generation [86], QKD (quantum key distribution) [26], and

wiretap coding [87, 88]. These areas are out of the scope of this work. Secret key agreement is closely related to the wiretap channel coding [57]. Many works also study the relation of SKA and wiretap codes (e.g., [19, 20, 79, 81]). This relation also has inspired another SKA model called *the channel model* introduced in [20]. Finite-length analysis of wiretap channel coding is given in [88].

SKA problem has also been studied for the case that P_{XYZ} is one of a set of known distributions (so-called *compound* sources) [89], or is assumed to have a given property e.g., the entropy satisfies a lower bound (see [26, Theorem 5] and [81, Section V.B]). In the latter case, the construction is called *universal*. These protocols are capacity achieving but do not have finite-length analysis.

3.1.3 Organization

In Section 3.3, we prove a finite-length upper bound on $S^{\rightarrow}(X^n; Y^n | Z^n)$. In Section 3.4, we propose two new OW-SKA protocols that lead to three finite-length lower bounds for $S^{\rightarrow}(X^n; Y^n | Z^n)$, and compare our proposed protocols with other related SKA protocols in Section 3.4.3. We conclude the chapter in Section 3.5 and discuss interesting future directions. The proof of some of the Lemmas are given in the Appendix 3.6.

3.2 Two-party Secret Key Agreement

A two-party source model with wiretapper is defined as follows. Suppose Alice (Terminal 1), Bob (Terminal 2), and Eve (the adversary), have access to random variables (RV's) $V_1 = X, V_2 = Y$, and Z, respectively. These RV's are correlated and the goal of Alice and Bob by running an SKA protocol is to share a secret key K, utilizing their RV's and a reliable (noiseless), authenticated, and public channel. The public communication may be interactive in general or we might restrict Alice and Bob to only use a single one-way message (say from Alice to Bob). Let F denote the public communication of SKA protocol. Having

access to Z and F, Eve must not be able to learn any information about the secret key K.

Next definition shows how one can information-theoretically measure the quality of a secret key. This definition of information-theoretic SKA is from [76], and is also used in [59] and [31].

Definition 3.1. Consider a source model P_{XYZ}, where Z is Eve's side information about (X, Y). A key K with alphabet $\mathcal{K}$ is an (ϵ, σ)-Secret Key (in short (ϵ, σ)-SK) if there exists an SKA protocol with public communication F and output key estimates $(K_1, K_2) \in \mathcal{K}^2$, such that

$$\text{(reliability)} \quad \Pr\{K_1 = K_2 = K\} \geqslant 1 - \epsilon, \tag{3.1}$$

$$\text{(secrecy)} \quad \mathbf{SD}\,(KFZ, UFZ) \leqslant \sigma, \tag{3.2}$$

where U is the uniform distribution over $\mathcal{K}$.

To increase the length of generated key, parties can sample from their variables for multiple times and observe n independent and identically distributed copies of their RV. Let n be a finite positive integer. Suppose SKA protocol Π establishes an $(\epsilon_n, \sigma_n)-$SK $K^{(n)}$ and let $\ell_\Pi(n) = \log |\mathcal{K}^{(n)}|$ denote the length of $K^{(n)}$. The key rate of Π for $n-$IID observations is given by $1/n\ell_\Pi(n)$, and $r_K(\Pi) = \liminf_{n \to \infty} 1/n\ell_\Pi(n)$ (if exists) is called the *asymptotic key rate* of Π. The asymptotic key rate of Π, i.e., $r_K(\Pi)$, is called *achievable* if $\lim_{n \to \infty} \epsilon_n = \lim_{n \to \infty} \sigma_n = 0$. The key capacity of a source model is defined as the maximum of all achievable asymptotic key rates of SKA protocols for the model. See Definition 3.2. Also, for any integer $n \in \mathbb{N}$, and any given $\epsilon, \sigma \in [0, 1)$, define $S_{\epsilon, \sigma}(X^n; Y^n | Z^n)$ to be the maximum length of all (ϵ, σ)-SK's that can be established for the two-party SKA model given by the probability distribution P_{XYZ}. The public channel is assumed to be available to all parties and is considered "free," that is the cost of establishing it is not considered. In practice, however, communicating over this channel incurs and SKA protocols that use less public communication bits per observed sample are more desirable. Thus, we also evaluate SKA

protocols in terms of their public communication cost. The *asymptotic public communication rate* of Π is defined by $r_{PC}(\Pi) = \limsup_{n \to \infty} {}^1\!/\!n \log(\mathrm{supp}(F(\Pi)))$, where $F(\Pi)$ is the public communication of Π.

Definition 3.2 (Key Capacity – Definition 17.16 of [90]). Consider an IID $P_{X^n Y^n Z^n}$, where Z^n is Eve's side information about (X^n, Y^n). A real number $R \geqslant 0$ is an achievable SK rate if there exists an SKA protocol that for every n establishes an $(\epsilon_n, \sigma_n)-$SK K with alphabet $\mathcal{K}$ where $\lim_{n\to\infty} \epsilon_n = \lim_{n\to\infty} \sigma_n = 0$, and $\lim_{n\to\infty} {}^1\!/\!n \log |\mathcal{K}| = R$. The maximum of all achievable SK rates is called the key capacity of P_{XYZ}.

When $Z =$ constant the key capacity is called the secret key (SK) capacity, and is denoted by $C_{SK}(P_{XY})$. It was proved that [20, Proposition 1]

$$C_{SK}(P_{XY}) = I(X; Y). \tag{3.3}$$

When Z is known by the legitimate parties (Alice and Bob), we call the key capacity private key (PK) capacity, and it was proved [20, Theorem 3] that the two-party PK capacity is given by

$$C_{PK}(P_{XYZ}) = I(X; Y|Z). \tag{3.4}$$

When $Z \neq$ constant the key capacity is called the wiretap secret key (WSK) capacity, and is denoted by $C_{WSK}(P_{XYZ})$. Equivalently the WSK capacity can be defined by [31, Defenition 12]

$$C_{WSK}(P_{XYZ}) = \sup_{\epsilon_n, \sigma_n} \liminf_{n \to \infty} \frac{1}{n} S^{\to}_{\epsilon_n, \sigma_n}(X^n; Y^n | Z^n), \tag{3.5}$$

where the sup is over all (ϵ_n, σ_n)'s with $\lim_{n\to\infty} \epsilon_n = \lim_{n\to\infty} \sigma_n = 0$.

WSK capacity upper bounds. Maurer proved [19] the following upper bound,

$$C_{WSK}(P_{XYZ}) \leqslant \min\{I(X; Y), I(X; Y|Z)\}. \tag{3.6}$$

This bound was later improved in [91] to

$$C_{WSK}(P_{XYZ}) \leqslant I(X;Y \downarrow Z) := \inf_{XY-Z-J} I(X;Y|J), \tag{3.7}$$

where $I(X;Y \downarrow Z)$ is called the *intrinsic mutual information* [91]. Renner and Wolf further improved the above bound by introducing the *double intrinsic mutual information* upper bound [92], that is

$$C_{WSK}(P_{XYZ}) \leqslant I(X;Y \downarrow\downarrow Z) := \inf_{P_{J|XYZ}} H(J) + I(X;Y \downarrow ZJ). \tag{3.8}$$

The best known upper bound on the WSK capacity is due to [93], which is

$$C_{WSK}(P_{XYZ}) \leqslant \inf_{P_{J|XYZ}} \max_{V-U-XY-ZJ} I(X;Y|J) + I(U;J|V) - I(U;Z|V). \tag{3.9}$$

WSK capacity lower bounds.   Maurer proved [19] that

$$C_{WSK}(P_{XYZ}) \geqslant \max\{I(X;Y) - I(X;Z), I(X;Y) - I(Y;Z)\}, \tag{3.10}$$

while Ahlswede and Csiszár proved a tighter lower bound [20], which is

$$C_{WSK}(P_{XYZ}) \geqslant \max_{V-U-X-YZ} I(U;Y|V) - I(U;Z|V). \tag{3.11}$$

Both lower bounds (3.10) and (3.11) are achievable with noninteractive communication. These lower bounds were improved in [93] by utilizing interactive communication. For any given integer θ, let RVs $T_1, T_2, \ldots, T_\theta$ satisfy the following conditions,

$$T_j - XT_{[j-1]} - YZ \text{ for odd } j \tag{3.12}$$

$$T_j - YT_{[j-1]} - XZ \text{ for even } j \tag{3.13}$$

and

$$|\mathcal{T}_j| \leqslant |\mathcal{X}| \prod_{i=1}^{j-1} |\mathcal{T}_i| \quad \text{for odd } j \tag{3.14}$$

$$|\mathcal{T}_j| \leqslant |\mathcal{Y}| \prod_{i=1}^{j-1} |\mathcal{T}_i| \quad \text{for even } j \tag{3.15}$$

where $T_{[j-1]} = (T_1, T_2, \ldots, T_{j-1})$. For any given integer θ, and RVs $T_1, T_2, \ldots, T_\theta$ satisfying conditions (3.12)-(3.15), and for any given integer ζ such that $1 \leqslant \zeta \leqslant \theta$, let

$$L_{\zeta,\theta}(T_{[\theta]}) = \sum_{\substack{j \geqslant \zeta \\ \text{odd } j}} I(T_j; Y|T_{[j-1]}) - I(T_j; Z|T_{[j-1]}) + \sum_{\substack{j \geqslant \zeta \\ \text{even } j}} I(T_j; X|T_{[j-1]}) - I(T_j; Z|T_{[j-1]}).$$

Then it is proved [93] that

$$C_{WSK}(P_{XYZ}) \geqslant \max_{\theta, P_{T_{[\theta]}}, \zeta} L_{\zeta,\theta}(T_{[\theta]}). \tag{3.16}$$

The above bound is the best known lower bound for WSK capacity [94]; however, it is hard to evaluate an equivalent single-letter expression for this bound since the maximization is over any arbitrary θ and ζ and cardinality of $\mathcal{T}_j$ alphabets grow exponentially by θ.

3.2.1 One-way Secret Key Agreement

For SKA protocols that are limited to one-way public communication, let $C_{WSK}^{\rightarrow}(P_{XYZ})$ denote the one-way wiretap secret key (OW-WSK) capacity. Throughout, any quantity with arrow as superscript corresponds to its OW-SKA counterpart. Ahlswede and Csiszár [20] derived the "forward key capacity" (or what we call one-way secret key capacity) of the source model.

Theorem 3.1 (Theorem 1 of [20]). *For any IID $P_{X^n Y^n Z^n}$, the OW-WSK capacity is given by*

$$C_{WSK}^{\rightarrow}(P_{XYZ}) = \max_{P_{VU}} H(U|ZV) - H(U|YV), \tag{3.17}$$

where the maximization is over RV's (V, U) that satisfy $V - U - X - (Y, Z)$.

Remark 3.1. For some source models, the optimizing P_{UV} can be analytically calculated [76, 78], and be used to construct OW-SKA that achieves the OW-WSK capacity [76]. A special case is when the Markov chain $X - Y - Z$ holds. For this case, $U = X$ and $V = $ constant are the maximizing RV's, and the OW-WSK and the WSK capacity are equal, that is

$$C_{WSK} = C_{WSK}^{\rightarrow} = H(X|Z) - H(X|Y) = I(X;Y|Z). \tag{3.18}$$

In general, however, finding the optimizing RV's (V, U) might be difficult [78, 79].

3.2.2 Finite-length Performance

In real-life implementations, bounds on the achievable key length for finite number (n) of samples of the source variables is required. The finite-length performance of an SKA protocol is given by finite-length approximations of its corresponding highest achievable key length. For source models, finite-length upper and lower bounds on highest achievable key length determine finite-length limits associated with the model.

Remark 3.2. When interactive public communication is allowed between Alice and Bob, and if $X - Y - Z$ holds, $C_{WSK} = I(X, Y|Z) = H(X|Z) - H(X|Y) = C_{WSK}^{\rightarrow}$, and it is proved that [31, Theorem 15],

$$S_{\epsilon,\sigma}(X^n; Y^n|Z^n) = nC_{WSK} - \sqrt{n}G_{\epsilon,\sigma} \pm \mathcal{O}(\log n), \tag{3.19}$$

with $G_{\epsilon,\sigma} = \sqrt{\Delta_{XY|Z}}Q^{-1}(\epsilon + \sigma)$, where

$$\Delta_{XY|Z} = \text{Var}\left\{\log P_{XY|Z}(XY|Z) - \log P_{X|Y}(X|Y)P_{X|Z}(X|Z)\right\}, \tag{3.20}$$

and $Q(\cdot)$ is the tail probability of the standard Gaussian distribution, i.e.,

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty \exp\left(-\frac{t^2}{2}\right) dt.$$

It was conjectured in [31] that interaction is necessary to attain the second-order bound of (3.19) (see Remark 16 and Section VII of [31]); this conjecture remains unproved.

Our Objective. Let $S_{\epsilon,\sigma}^{\rightarrow}(X^n; Y^n | Z^n)$ be the maximum key length of all $(\epsilon, \sigma)-$SK's obtained by one-way public communication. Theorem 3.1 implies that for all $\epsilon_n, \sigma_n \to 0$, we have

$$\sup_{\epsilon_n, \sigma_n} S_{\epsilon_n, \sigma_n}^{\rightarrow}(X^n; Y^n | Z^n) = n C_{WSK}^{\rightarrow}(P_{XYZ}) + o(n).$$

However, we are interested in deriving a more refined asymptotic expansion of $S_{\epsilon,\sigma}^{\rightarrow}(X^n; Y^n | Z^n)$.

In the following sections, for fixed reliability and secrecy parameters ϵ, σ, we prove a multi-letter upper bound on $S_{\epsilon,\sigma}^{\rightarrow}(X^n; Y^n | Z^n)$, and propose two OW-SKA protocols based on which we prove multiple finite-length lower bounds.

3.3   Upper Bound

In this section, we prove our upper bound on $S_{\epsilon,\sigma}^{\rightarrow}(X^n; Y^n | Z^n)$. The proof is based on combining the smooth Rényi entropy framework of [26] and the information spectrum methods of [24, 25].

First, we review the smooth min/max entropies [26], introduce the inf/sup-spectral entropies, and prove their important properties.

Definition 3.3 ([25, 26]). For any joint probability distribution $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$, with marginals P_X and P_Y, let

$$H_{\min}^{\epsilon}(X|Y) = \max_{Q_{X'Y'} \in \mathcal{B}^{\epsilon}(P_{XY})} \min_{\substack{y \in \text{supp}(P_Y) \\ x \in \mathcal{X}}} -\log \frac{Q_{X'Y'}(x, y)}{P_Y(y)},$$

and

$$H_{\max}^{\epsilon}(X|Y) = \min_{Q_{X'Y'} \in \mathcal{B}^{\epsilon}(P_{XY})} \max_{y \in \mathcal{Y}} \log |\{x : Q_{X'Y'}(x, y) > 0\}|,$$

66

where $\mathcal{B}^\epsilon(P_{XY}) = \{Q_{X'Y'} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \mathbf{SD}(X'Y', XY) \leqslant \epsilon\}$. Also define,

$$\hat{H}^\epsilon_{\min}(X|Y) = \max_{Q_{X'Y'} \in \hat{\mathcal{B}}^\epsilon(P_{XY})} \min_{\substack{y \in \mathrm{supp}(P_Y) \\ x \in \mathcal{X}}} - \log \frac{Q_{X'Y'}(x,y)}{P_Y(y)},$$

and

$$\hat{H}^\epsilon_{\max}(X|Y) = \min_{Q_{X'Y'} \in \hat{\mathcal{B}}^\epsilon(P_{XY})} \max_{y \in \mathcal{Y}} \log |\{x : Q_{X'Y'}(x,y) > 0\}|,$$

where $\overline{\mathcal{P}}(\mathcal{X})$ denotes the set of all sub-normalized positive distributions on $\mathcal{X}$ and $\hat{\mathcal{B}}^\epsilon(P_{XY}) = \{Q_{X'Y'} \in \overline{\mathcal{P}}(\mathcal{X} \times \mathcal{Y}) : \mathbf{SD}(X'Y', XY) \leqslant \epsilon\}$.

Watanabe and Hayashi introduced the *inf-spectral entropy* in [25]. Similarly, we introduce the *sup-spectral entropy*, and define these entropies below.

Definition 3.4. For $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$, and $0 \leqslant \epsilon \leqslant 1$, let

$$\bar{H}^\epsilon_s(X|Y) = \inf\{r : P_{XY}\{-\log P_{X|Y}(x|y) \geqslant r\} \leqslant \epsilon\},$$

be the sup-spectral entropy, and

$$\underline{H}^\epsilon_s(X|Y) = \sup\{r : P_{XY}\{-\log P_{X|Y}(x|y) \leqslant r\} \leqslant \epsilon\},$$

is the inf-spectral entropy [25].

Remark 3.3. Note that sup-spectral entropy intuitively captures the reliability (IR) aspect of SKA. In [25], inf-spectral entropy is used to bound the output key length of PA (which corresponds to the secrecy aspect of SKA). See also Section 3.6.2 of Appendix.

In the n-IID setting, the following lemma gives the Gaussian approximation of the spectral entropies.

Lemma 3.2. *For IID $P_{X^n Y^n}$, and $0 \leqslant \epsilon \leqslant 1$, we have*

$$\bar{H}^\epsilon_s(X^n|Y^n) = nH(X|Y) + \sqrt{n\Delta_{X|Y}}Q^{-1}(\epsilon) \pm \mathcal{O}(1),$$

and

$$\underline{H}_s^\epsilon(X^n|Y^n) = nH(X|Y) - \sqrt{n\Delta_{X|Y}}Q^{-1}(\epsilon) \pm \mathcal{O}(1).$$

Lemma 3.2 follows from Definition 3.4 and the Berry-Esseen Theorem [84]. First recall the Berry-Esseen Theorem.

Theorem (Berry-Esseen, see Theorem 2.4). *Let W^n be an $n-IID$ real-valued variable, and $-\infty < \alpha < \infty$, then*

$$\left| \Pr\left\{ \sum_{j=1}^n W_j \leqslant n\mu_W - \alpha\sqrt{\Delta_W n} \right\} - Q(\alpha) \right| \leqslant \frac{3\rho_W}{\Delta_W^{3/2}\sqrt{n}},$$

where $Q(\cdot)$ is the tail probability of the standard Gaussian distribution, i.e.,

$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty \exp\left(-\frac{t^2}{2}\right)\mathrm{d}t,$$

$\mu_W = \mathbb{E}\{W\}, \Delta_W = \mathrm{Var}\{W\},$ *and* $\rho_W = \mathbb{E}\{|W-\mu|^3\}.$

Proof of Lemma 3.2: We prove

$$\bar{H}_s^\epsilon(X^n|Y^n) = nH(X|Y) + \sqrt{n\Delta_{X|Y}}Q^{-1}(\epsilon) \pm \mathcal{O}(1).$$

For all j define $W_j = -\log P_{X|Y}(X|Y)$, $\mu_W = \mathbb{E}\{W_1\} = H(X|Y)$, $\Delta_W = \mathrm{Var}\{W_1\}$, and $\rho_W = \mathbb{E}\{|W_1 - \mu_W|^3\}$. Let $r = n\mu_W + Q^{-1}(\epsilon + \theta_n)\sqrt{n\Delta_W}$ with $\theta_n = -3\rho_W/\Delta_W^{3/2}\sqrt{n}$. Then, by Berry-Esseen Theorem we get

$$\left| \Pr\left\{ \sum_{j=1}^n W_j \geqslant r \right\} - (\epsilon + \theta_n) \right| \leqslant \frac{3\rho_W}{\Delta_W^{3/2}\sqrt{n}},$$

which implies

$$\Pr\left\{ \sum_{j=1}^n W_j \geqslant r \right\} \leqslant \epsilon.$$

Let N be such that for all $n \geqslant N$ we have $\epsilon + \theta_n > 0$. Thus, for all $n \geqslant N$ we have

$$\bar{H}_s^\epsilon(X^n|Y^n) = nH(X|Y) + Q^{-1}(\epsilon + \theta_n)\sqrt{n\Delta_W},$$

where by using Taylor expansions we get

$$\bar{H}_s^\epsilon(X^n|Y^n) = nH(X|Y) + Q^{-1}(\epsilon)\sqrt{n\Delta_{X|Y}} \pm \mathcal{O}(1).$$

Similarly we can show that

$$\underline{H}_s^\epsilon(X^n|Y^n) = nH(X|Y) - Q^{-1}(\epsilon)\sqrt{n\Delta_{X|Y}} \pm \mathcal{O}(1). \qquad \blacksquare$$

Lemma 3.3 gives the relation between sup/inf-spectral entropies and smooth min/max entropies.

Lemma 3.3. *For any P_{XY}, $0 \leqslant \epsilon \leqslant 1$, and every $0 < \xi \leqslant 1 - \epsilon$*

$$H_{\min}^\epsilon(X|Y) \leqslant \underline{H}_s^{\epsilon+\xi}(X|Y) - \log \xi,$$

and

$$H_{\max}^\epsilon(X|Y) \geqslant \bar{H}_s^{\epsilon+\xi}(X|Y) + \log \xi.$$

Proof: The first inequality is due to Lemma 4 of [25], and the proof the second inequality is as follows.

Define real number r and $Q'_{\bar{X}\bar{Y}}$ by $r = H_{\max}^\epsilon(X|Y) = \max_{y \in \mathcal{Y}} \log |\{x : Q'_{\bar{X}\bar{Y}}(x,y) > 0\}|$, where $\mathbf{SD}(\bar{X}\bar{Y}, XY) \leqslant \epsilon$ holds. Also, for an arbitrary $\delta > 0$ define the following sets

$$\mathcal{T} = \{(x,y) : P_{XY}(x,y) \leqslant 2^{-r-\delta}P_Y(y)\},$$

$$\mathcal{T}' = \{(x,y) : Q'_{\bar{X}\bar{Y}}(x,y) \leqslant 2^{-r-\delta}P_Y(y)\},$$

$$\mathcal{T}'_y = \{x : (x,y) \in \mathcal{T}'\}.$$

There always exists a distribution $Q'_{\bar{X}\bar{Y}}$ defined as above such that, $P_{XY}\{\mathcal{T}'\} \geqslant P_{XY}\{\mathcal{T}\}$.

Considering such $Q'_{\bar{X}\bar{Y}}$, we have

$$
\begin{aligned}
\epsilon &\overset{(a)}{\geqslant} \mathbf{SD}(\bar{X}\bar{Y}, XY) \\
&\overset{(b)}{\geqslant} P_{XY}\{\mathcal{T}'\} - Q'_{\bar{X}\bar{Y}}\{\mathcal{T}'\} \\
&\overset{(c)}{\geqslant} P_{XY}\{\mathcal{T}\} - \sum_{y\in\mathcal{Y}} P_Y(y) \sum_{x\in\mathcal{T}'_y} 2^{-r-\delta} \\
&\overset{(d)}{\geqslant} P_{XY}\{\mathcal{T}\} - 2^{-\delta},
\end{aligned}
$$

where (a) is due to Definition 3.3, (b) is due to definition of $\mathbf{SD}(\cdot,\cdot)$, (c) is due to definition of $\mathcal{T}'$, and (d) is due to definition of r. Thus,

$$
\begin{aligned}
r + \delta &\geqslant \inf\{r' : P_{XY}\{-\log P_{X|Y}(x|y) \geqslant r'\} \leqslant \epsilon + 2^{-\delta}\} \\
&= \bar{H}_s^{\epsilon+2^{-\delta}}(X|Y),
\end{aligned}
$$

and by choosing $\delta = -\log \xi$, the proof is complete. $\blacksquare$

For the single-shot ($n = 1$) setting, we give the following upper bound on $S_{\epsilon,\sigma}^{\rightarrow}(X; Y|Z)$.

Theorem 3.4 (Single-shot ($n = 1$) upper bound). *For any P_{XYZ},*

$$
S_{\epsilon,\sigma}^{\rightarrow}(X; Y|Z) \leqslant \max_{P_{VU}} \underline{H}_s^{\sigma'}(U|ZV) - \bar{H}_s^{\epsilon'}(U|YV) - \log \mu\nu,
$$

where (V, U) satisfy $V - U - X - (Y, Z)$, $0 \leqslant \epsilon \leqslant 1$, $0 \leqslant \sigma \leqslant 1$, $0 < \sigma' = \sigma + \mu \leqslant 1$, and $0 < \epsilon' = \epsilon + \nu \leqslant 1$.

Proof: From Theorem 3 of [26] we know that, For any P_{XYZ}, we have[1]

$$
S_{\epsilon,\sigma}^{\rightarrow}(X; Y|Z) \leqslant \max_{P_{VU}} H_{\min}^{\sigma}(U|ZV) - H_{\max}^{\epsilon}(U|YV),
$$

where (V, U) are satisfying $V - U - X - (Y, Z)$. Applying Lemma 3.3 completes the proof.

---

[1]The expression of the upper bound here is slightly different that the one given in Theorem 3 of [26]; however, the line of argument used in the proof still applies and the modified upper bound stated here holds.

The next finite-length upper bound on $S_{\epsilon,\sigma}^{\rightarrow}(X^n;Y^n|Z^n)$ follows from Theorem 3.4 for the case of $n-$IID random variables (X^n, Y^n, Z^n).

Corollary 3.4.1 (Finite-length upper bound). *For any IID $P_{X^n Y^n Z^n}$,*

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n;Y^n|Z^n) \leqslant \max_{P_{V^n U^n}} \underline{H}_s^{\sigma'}(U^n|Z^n V^n) - \bar{H}_s^{\epsilon'}(U^n|Y^n V^n) - \log\mu\nu,$$

where (V^n, U^n) satisfy $V^n - U^n - X^n - (Y^n, Z^n)$, $0 \leqslant \epsilon \leqslant 1$, $0 \leqslant \sigma \leqslant 1$, $0 < \sigma' = \sigma + \mu \leqslant 1$, and $0 < \epsilon' = \epsilon + \nu \leqslant 1$.

The tightness of this bound remains an open question. (We note that an error in the proof of this theorem in [33] had lead to the incorrect claim of the tightness of a second-order finite-length upper bound[2].)

3.4   Lower Bounds

In this section, we propose two one-way secret key agreement (OW-SKA) protocols, analyses of which lead to multiple finite-length lower bounds for OW-SKA. Both of these protocols are composed based on the common standard structure of having the two steps (sub-protocols):

- **The information reconciliation (IR),** starts with Alice sending Bob a public message F that is a function of her observation X; i.e., $F = f_{IR}(X)$. Bob, then, uses this public message and his own observation Y, to recover an estimation $\hat{X}$ of Alice's RV X; i.e., $\hat{X} = f_{IR}^{-1}(F, Y)$. The RV X is referred to as the common randomness. An IR protocol (f_{IR}, f_{IR}^{-1}) is called $\epsilon-$correct if $\Pr\left\{f_{IR}^{-1}(f_{IR}(X), Y) \neq X\right\} \leqslant \epsilon$.

- **The privacy amplification (PA),** consists of Alice and Bob extracting their respective secret keys out of their common randomness X; i.e., $K = f_{PA}(X, F)$. A function f_{PA} is called $\sigma-$secure if $\mathbf{SD}\left(KFZ, UFZ\right) \leqslant \sigma$ where U is uniform over $\mathcal{K}$.

---

[2] The proof implicitly assumed that RV's (V^n, U^n) are IID.

When an SKA protocol comprises an $\epsilon-$correct IR and a $\sigma-$secure PA, the final generated shared key is $(\epsilon, \sigma)-$SK.

For the privacy amplification step (in both of our proposed OW-SKA protocols) we use universal hashing [28]. In the following, we prove modified versions of the Leftover Hash Lemma (LHL) (Lemmas 3.7 and 3.14) that will be used for the purpose of analyzing our proposed OW-SKA protocols. See Section 2.3.1 for a review on privacy amplification and LHL.

3.4.1   A OW-SKA Protocol With Highest Finite key length

In the source model setting, suppose Alice, Bob and Eve have their corresponding n components of the source (X^n, Y^n, Z^n). Let the required secrecy and reliability parameters of the key be σ and ϵ, respectively. Alice and Bob use our proposed protocol $\mathbf{\Pi_{HH}}$ for SKA: They choose two 2-UHFs $h_s : \mathcal{X}^n \rightarrow \{0,1\}^t$ and $\hat{h}_{s'} : \mathcal{X}^n \rightarrow \{0,1\}^\ell$, and share (over public channel) two uniformly random seeds $s \in \mathcal{S}$ and $s' \in \mathcal{S}'$ for the two families. The values of t (length of the public message, i.e., the output length of the 2-UHF h_s) and ℓ (the length of the final key, i.e., the output length of the 2-UHF $\hat{h}_{s'}$) are determined according to σ and ϵ and the given source model distribution $P_{X^n Y^n Z^n}$. Alice uses $h_s(\cdot)$ to compute the hash value of her sample vector x^n, and sends it to Bob; Bob uses the received hash value, his sample vector y^n, and the known probability distribution of the source, to recover Alice's sample vector (*information reconciliation*). The main idea behind the reconciliation technique of Protocol 4, used by Bob, is to divide the range of the spectrum of $P_{X^n|y^n}$ into two parts, and search in only the main part to find Alice's vector. This reduces the search complexity of the protocol compared to the similar information reconciliation algorithm of [58] that searches through all x^ns with $\Pr\{x^n|y^n\} > 0$. By choosing an appropriate value for t, Bob can bound the reconciliation error to ϵ. The transmitted hash value will leak information about the key, and so longer public messages will result in shorter keys while reducing error probability. Alice and Bob will estimate the total leaked information about their common

---
Protocol 4: OW-SKA with universal hashing IR and universal hashing PA ($\mathbf{\Pi_{HH}}$)

    Public Information: P_{XYZ}

    Input: n−fold samples $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$, ϵ, σ.

    Output: Key estimates k_A and k_B.

    `// Initialization`

1 Alice and Bob, (i) find and share λ, and ℓ and t for the hash functions
    $h_s : \mathcal{X}^n \to \{0,1\}^t$ and $\hat{h}_{s'} : \mathcal{X}^n \to \{0,1\}^\ell$, (ii) generate and share the seeds $s \in \mathcal{S}$ and
    $s' \in \mathcal{S}'$ for the hash function.

    `// Information Reconciliation`

3 Alice sends the hash value $F = h_s(x^n)$ to Bob.

5 Bob forms a list of guesses for x^n,

$$\mathcal{T}(X^n|y^n) = \{\hat{x^n} : -\log P_{X^n|y^n}(\hat{x^n}|y^n) \leqslant \lambda\}. \tag{3.21}$$

7 Bob finds $\hat{x^n} \in \mathcal{T}(X^n|y^n)$ such that $h_s(\hat{x^n}) = F$.

9 **if** *no $\hat{x^n}$ was found **or** $\hat{x^n}$ was not unique* **then**

10   Abort the protocol.

    `// Privacy Amplification`

12 Alice and Bob find $k_A = \hat{h}_{s'}(x^n)$ and $k_B = \hat{h}_{s'}(\hat{x^n})$.

---

strings, and remove it during the key extraction phase (*privacy amplification*) by using the second 2-UHF $\hat{h}_{s'}$.

Theorem 3.8 proves a finite-length approximation of the maximum achievable key length of Protocol 4 assuming that the source samples are drawn through independent random experiments. However, to prove Theorem 3.8 we first need the following Lemmas.

Lemma 3.5 (Theorem 1 of of [83]). *Let (X^n, Y^n) be an n-independent repetition of (X, Y) with the joint distribution $P_{XY} = P_{X_1Y_1} \times \ldots \times P_{X_nY_n}$. Then for any $\delta \geqslant 0$*

$$\hat{H}^\epsilon_{\min}(X^n|Y^n) \geqslant H(X^n|Y^n) - n\delta,$$

and

$$\hat{H}^\epsilon_{\max}(X^n|Y^n) \leqslant H(X^n|Y^n) + n\delta,$$

where $\hat{H}_{\min}^{\epsilon}(X|Y)$ and $\hat{H}_{\max}^{\epsilon}(X|Y)$ are defined in Definition 3.3, and $2\epsilon = 2^{\frac{-n\delta^2}{2\log^2(|\mathcal{X}|+3)}}$.

Lemma 3.6 (Theorem 2 of [83]). *Let (X^n, Y^n) be an n-independent repetition of (X, Y) with the joint distribution $P_{XY} = P_{X_1 Y_1} \times \ldots \times P_{X_n Y_n}$. Then for any $\delta \in [0, \log(|\mathcal{X}|)]$ and (x, y) representing a random realization according to P_{XY}, we have*

$$\Pr\left\{-\log P_{X|Y}(x|y) \geqslant H(X^n|Y^n) + n\delta\right\} \leqslant \epsilon,$$

and

$$\Pr\left\{-\log P_{X|Y}(x|y) \leqslant H(X^n|Y^n) - n\delta\right\} \leqslant \epsilon,$$

where $\epsilon = 2^{\frac{-n\delta^2}{2\log^2(|\mathcal{X}|+3)}}$.

Lemma 3.7 (Smooth LHL). *For any $P_{XZF} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z} \times \mathcal{F})$, let $f_{PA} = h_s : \mathcal{X} \to \mathcal{K}$ be a 2-universal hash function with seed s drawn uniformly at random from $\mathcal{S}$. Then*

$$\mathbf{SD}(KZFS, UZFS) \leqslant 2\epsilon + \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\hat{H}_{\min}^{\epsilon}(X|Z)}},$$

where $\hat{H}_{\min}^{\epsilon}(X|Z)$ is defined in Definition 3.3, $K = h_S(X)$ and U is uniform over $\mathcal{K}$. This implies that if

$$\log|\mathcal{K}| \leqslant \hat{H}_{\min}^{\frac{\sigma-\eta}{2}}(X|Z) - \log|\mathcal{F}| + \log 4\eta^2,$$

then, function $f_{PA} = h_s$ is $\sigma-secure$ for any $0 < \mu \leqslant \sigma$.

Proof of Lemma 3.7 is given is the Appendix, Section 3.6.1.

Theorem 3.8 is given for the case that the source distribution $P_{X^n Y^n Z^n} = \Pi_j P_{X_j Y_j Z_j}$ is due to independent experiments which are not necessarily identical.

Theorem 3.8. *Let the source model (X^n, Y^n, Z^n) be described by a joint distribution $P_{X^n Y^n Z^n} = \Pi_j P_{X_j Y_j Z_j}$ due to independent experiments (not necessarily identical). For any $2^{-n/2} < \epsilon < 1$*

and any $2^{-n/4} < \sigma < 1$, *Protocol* 4 *(Π_{HH}) generates an* (ϵ, σ)*-SK with maximum key length*

$$\ell_{\mathbf{\Pi_{HH}}}(n) = R_n - \sqrt{n} f_{\epsilon,\sigma}(|\mathcal{X}|) - \log n + \mathcal{O}(1), \qquad (3.22)$$

where $f_{\epsilon,\sigma}(|\mathcal{X}|) = \sqrt{2} \log(|\mathcal{X}| + 3) \left(\sqrt{\log 1/\epsilon} + \sqrt{\log 1/\sigma} \right)$, *and* $R_n = H(X^n|Z^n) - H(X^n|Y^n)$.

Proof of Theorem 3.8: We want to find a maximum final key length (value of ℓ) such that under appropriate choice of t (length of public message) and λ (Bob's search parameter) the final established shared key of $\mathbf{\Pi_{HH}}$ is $(\epsilon, \sigma)-$SK. We first determine expressions for the values of t and λ that guarantee the $\epsilon-$correctness of the IR step.

In $\mathbf{\Pi_{HH}}$, Alice sends Bob a hash value of length t and then Bob chooses a positive λ to define the set

$$\mathcal{T}(X^n|y^n) = \{\hat{x^n} : -\log P_{X^n|y^n}(\hat{x^n}|y^n) \leqslant \lambda\},$$

and searches in the set $\mathcal{T}(X^n|y^n)$ for vector(s) whose hash value is equal to the received hash value, and declares success if a *unique* vector with the required property is found. Bob's search fails in two cases: (i) x is not in the set, and (ii) there are more than one vector in the set whose hash value matches the received hash value v. Bob's failure probability $P_e = \Pr\{K_A \neq K_B\}$ is upper bounded by bounding the probabilities of the above two cases. These probabilities are:

$$\Pr\{\xi_1\} = \Pr\{x^n \notin \mathcal{T}(X^n|y^n)\},$$
$$\Pr\{\xi_2\} = \Pr\{\exists \ \hat{x^n} \in \mathcal{T}(X^n|y^n) \ \text{s.t.} \ h_S(\hat{x^n}) = h_S(x^n)\}.$$

For any ϵ_1, ϵ_2 such that $\epsilon_1 + \epsilon_2 \leqslant \epsilon$, let ϵ_1 determine δ_1 as $\epsilon_1 = 2^{\frac{-n(\delta_1)^2}{2\log^2(|\mathcal{X}|+3)}}$, and let

$$\lambda = H(X^n|Y^n) + n\delta_1.$$

We have $\Pr\{\xi_1\} = \Pr\{-\log P_{X^n|Y^n}(x^n|y^n) > H(X^n|Y^n) + n\delta_1\}$, which by Lemma 3.6 is

bounded as $\Pr\{\xi_1\} \leqslant \epsilon_1$. For $\Pr\{\xi_2\}$, note that the sum of probability of vectors in $\mathcal{T}(X^n|y^n)$ is less than 1, and using (3.21), we have $|\mathcal{T}(X^n|y^n)| \leqslant 2^\lambda$. Using union bound $\Pr\{\xi_2\} \leqslant |\mathcal{T}(X^n|y^n)| \cdot 2^{-t} = 2^{\lambda-t}$, where 2^{-t} is the collision probability of $h_S(\cdot)$ and upper bounds the probability of two x^ns in $\mathcal{T}(X^n|y^n)$ having the same hash value. By letting

$$t = \lambda - \log \epsilon_2$$
$$= H(X^n|Y^n) + n\delta_1 - \log \epsilon_2,$$

we will have $\Pr\{\xi_2\} \leqslant \epsilon_2$. Thus, for $t = H(X^n|Y^n) + n\delta_1 - \log \epsilon_2$, we have $P_e \leqslant \Pr\{\xi_1\} + \Pr\{\xi_2\} \leqslant \epsilon_1 + \epsilon_2 \leqslant \epsilon$, and the reliability condition is satisfied. That is the IR function $f_{IR} = h_S$ is $\epsilon-$correct.

By Smooth LHL 3.7 we know that $f_{PA} = \hat{h}_{S'}$ is $\sigma-$secure if

$$\ell \leqslant \hat{H}_{\min}^{\frac{\epsilon'}{2}}(X^n|Z^n) - t + \log 4(\sigma - \epsilon')^2$$
$$\leqslant H(X^n|Z^n) - n\delta' - t + 2 + 2\log(\sigma - \epsilon')$$
$$= \big(H(X^n|Z^n) - H(X^n|Y^n)\big) - n\delta' - n\delta_1 + 2 + 2\log(\sigma - \epsilon') + \log \epsilon_2,$$

where the second inequality is due to Lemma 3.5 with $0 < \delta'$ and ϵ' satisfying $\epsilon' = 2^{\frac{-n\delta'^2}{2\log^2(|\mathcal{X}|+3)}}$.

If ℓ satisfies the above bound and t is chosen as determined above, then IR protocol is $\epsilon-$correct and the PA is $\sigma-$secure. This guarantees that the final key is an $(\epsilon, \sigma)-$SK.

Finally, the key length approximation of (3.22) is obtained by by choosing $\epsilon_1 = (\sqrt{n}-1)\epsilon/\sqrt{n}$ and $\epsilon_2 = \epsilon - \epsilon_1 = \epsilon/\sqrt{n}$, and $\epsilon' = (\sqrt[4]{n}-1)\sigma/\sqrt[4]{n}$, and by noting that as $n \to \infty$

$$(\log \frac{a\sqrt{n}}{(\sqrt{n}-1)b})^{1/2} = (\log \frac{a}{b})^{1/2} + \frac{1}{2\ln 2\sqrt{n\log\frac{a}{b}}} + \mathcal{O}(\frac{1}{n}).$$

∎

Theorem 3.8 gives the maximum achievable key length $\ell_{\Pi_{\mathbf{HH}}}(n)$ of the protocol and

provides a lower bound on the maximum key length of OW-SKA protocols. That is,

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n | Z^n) \geqslant (H(X^n | Z^n) - H(X^n | Y^n))$$
$$-\sqrt{2n} \log(|\mathcal{X}| + 3) \left(\sqrt{\log \frac{1}{\epsilon}} + \sqrt{\log \frac{1}{\sigma}} \right) - \log n + \mathcal{O}(1) \quad (3.23)$$

Next corollary tightens this lower bound for IID sources, using Berry-Essen inequality [84].

Theorem 3.9. *For any source model described by IID distribution P_{XYZ} we have*

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n | Z^n) \geqslant R_n - \sqrt{n} g_{\epsilon,\sigma} - \log n + \mathcal{O}(1), \quad (3.24)$$

where $R_n = n(H(X|Z) - H(X|Y))$, $g_{\epsilon,\sigma} = Q^{-1}(\epsilon)\sqrt{\Delta_{X|Y}} + Q^{-1}(\sigma)\sqrt{\Delta_{X|Z}}$, and $\Delta_{U|V} = \text{Var}\left\{-\log P_{U|V}\right\}$.

Proof of Theorem 3.9: The proof is along the same lines as for the proof of Theorem 3.8. For reliability we bound the probability of these two events:

$$\xi_1 = \{x^n : -\log P_{X^n | Y^n}(x^n | y^n) > \lambda\}$$
$$\xi_2 = \{x^n \in \mathcal{T}(X^n | y^n) : \exists \, \hat{x}^n \in \mathcal{T}(X^n | y^n) \text{ s.t. } h_S(\hat{x}^n) = h_S(x^n)\}.$$

Let $W_i = -\log P_{X_i | Y_i}$ and let

$$\lambda = nH(X|Y) + \sqrt{n\Delta_{X|Y}} Q^{-1}(\epsilon - \theta_n),$$

where $\Delta_{X|Y} = \text{Var}\left\{-\log P_{X|Y}\right\}$, and $\theta_n = \frac{1}{\sqrt{n}} + \frac{3\rho}{V_{X|Y}^{3/2}\sqrt{n}}$. Then by the Berry-Esseen Theorem 2.4, $\Pr\{\xi_1\} \leqslant \epsilon - \frac{1}{\sqrt{n}}$. By choosing

$$t = \lambda - \log \frac{1}{\sqrt{n}},$$

we get $\Pr\{K_A \neq K_B\} \leqslant \Pr\{\xi_1\} + \Pr\{\xi_2\} \leqslant \epsilon$.

For the secrecy constraint, we use smooth LHL 3.7. For $\eta_n = \frac{1}{2\sqrt[4]{n}}$ we get

$$\ell \leqslant \hat{H}_{\min}^{\frac{\sigma - \eta_n}{2}}(X^n|Z^n) - t + \log 4\eta_n^2.$$

From [25, Lemma 3] we know that for any IID distribution $P_{X^n Z^n}$, $\hat{H}_{\min}^{\epsilon/2}(X^n|Z^n) \geqslant \underline{H}_s^\epsilon(X^n|Z^n)$, and thus, by Lemma 3.2 we have

$$\hat{H}_{\min}^\delta(X^n|Z^n) \geqslant nH(X|Z) - Q^{-1}(2\delta)\sqrt{n\Delta_{X|Z}} + \mathcal{O}(1),$$

where $\Delta_{X|Z} = \mathrm{Var}\left\{-\log P_{X|Z}\right\}$. Thus,

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n) \geqslant n(H(X|Z) - H(X|Y))$$
$$-\sqrt{n}\left(Q^{-1}(\epsilon - \theta_n)\sqrt{\Delta_{X|Y}} + Q^{-1}(\sigma - \eta_n)\sqrt{\Delta_{X|Z}}\right) - \log n + \mathcal{O}(1).$$

And ultimately the proof is complete by using Taylor expansions to remove θ_n and η_n. ∎

Remark 3.4 (Public communication cost of $\mathbf{\Pi_{HH}}$). For the source model described by distribution $P_{X^n Y^n Z^n} = \Pi_j P_{X_j Y_j Z_j}$, let $\log(\mathrm{supp}(F(\mathbf{\Pi_{HH}})))$ denote the public communication cost (in bits) that is used by the OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) to achieve $S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n)$. Then, our results show that

$$\log(\mathrm{supp}(F(\mathbf{\Pi_{HH}}))) = H(X^n|Y^n) + \sqrt{n}B_1^\epsilon + \frac{1}{2}\log n + \mathcal{O}(1), \tag{3.25}$$

where $B_1^\epsilon = \sqrt{2}\log(|\mathcal{X}| + 3)\sqrt{\log\frac{1}{\epsilon}}$. Moreover, for the case when source distribution is $n-$IID, a tighter approximation of $\log(\mathrm{supp}(F(\mathbf{\Pi_{HH}})))$ is given by

$$\log(\mathrm{supp}(F(\mathbf{\Pi_{HH}}))) = nH(X|Y) + \sqrt{n}B_2^\epsilon + \frac{1}{2}\log n + \mathcal{O}(1), \tag{3.26}$$

where $B_2^\epsilon = Q^{-1}(\epsilon)\sqrt{\Delta_{X|Y}}$.

Note that Theorem 3.8 and Theorem 3.9 both prove that the proposed OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) is capacity achieving when parties' observations are $n-$IID and if the Markov relation $X - Y - Z$ holds In this case $C_{WSK} = I(X;Y|Z) = H(X|Z) - H(X|Y)$, and due to both Theorems 3.8 and 3.9 we have

$$r_K(\mathbf{\Pi_{HH}}) = \liminf_{n \to \infty} \frac{1}{n} \ell_{\mathbf{\Pi_{HH}}}(n) = H(X|Z) - H(X|Y) = C_{WSK}.$$

In the following, we prove a lower bound for the general case (when Markov relation $X-Y-Z$ does not necessarily hold) and show that Protocol 4 can also be used for the general case for achieving the OW-WSK capacity. This next finite-length lower bound is the best known (tightest) lower bound known for OW-SKA [3].

Proposition 3.10. *For an IID $P_{X^n Y^n Z^n}$, let (V, U) be the maximizing RV's of Theorem 3.1, and assume they can be calculated, then,*

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n;Y^n|Z^n) \geqslant nC_{WSK}^{\rightarrow} - \sqrt{n}G_{\epsilon,\sigma}^{\rightarrow} - \mathcal{O}(\log n), \tag{3.27}$$

where $G_{\epsilon,\sigma}^{\rightarrow} = \sqrt{\Delta_{U|YV}}Q^{-1}(\epsilon) + \sqrt{\Delta_{U|ZV}}Q^{-1}(\sigma) \neq 0$. The above lower bound holds for any pair of maximizing RV's (V, U) as per Theorem 3.1 (they may not be unique).

    Proof of Proposition 3.10: We show how Alice and Bob can agree on an $(\epsilon, \sigma)-$SK $K \in \mathcal{K}$ with length $\ell = \log |\mathcal{K}|$ that is equal to the RHS of (3.27). Alice who has access to X, first computes (V, U) as per Theorem 3.1 (assuming they can be calculated), then, makes V public. Let $\tilde{X} = U$, $\tilde{Y} = (Y, V)$, and $\tilde{Z} = (Z, V)$. Now, Alice, Bob, and Eve have access to $\tilde{X}, \tilde{Y}$, and $\tilde{Z}$, respectively. By Theorem 3.9 we know that for any source model $P_{\tilde{X}\tilde{Y}\tilde{Z}}$, using OW-SKA Protocol 4 Alice (with input $\tilde{X}$) and Bob (with input $\tilde{Y}$) can agree on an

---

[3]The second-order approximation of the bound given in Equation (3.27) is also mentioned in [31] and can be derived indirectly by the results in [25, 82, 85] .

$(\epsilon, \sigma)-$SK of length

$$\ell = n(H(\tilde{X}|\tilde{Z}) - H(\tilde{X}|\tilde{Y}))$$
$$-\sqrt{n}(\sqrt{\Delta_{\tilde{X}|\tilde{Y}}}Q^{-1}(\epsilon) + \sqrt{\Delta_{\tilde{X}|\tilde{Z}}}Q^{-1}(\sigma)) - \log n - o(\log n), \qquad (3.28)$$

hence, completing the proof. ∎

Note that the above lower bound shows that Protocol 4 achieves the OW-WSK capacity when source samples are $n-$IID. In the remainder of this section we prove that Protocol 4 also achieves the WSK capacity when source samples are independent but not necessarily $n-$IID.

Extension to General n-fold Sources. Hayashi et al., [31], considered the setting of a "general source", a generalized stochastic process (see Section 2.2.2 of Chapter 2, [31] and [24]), and assumed an n-fold (X^n, Y^n, Z^n) source model that satisfies the Markov constraint, $X^n - Y^n - Z^n$. They proved [31, Theorem 14] the wiretap secret key capacity of this general source model is given by

$$C_{WSK}(P_{XYZ}) = \underline{I}(X; Y|Z), \qquad (3.29)$$

where $\underline{I}(X; Y|Z)$ is the inf-conditional information rate [24] of X and Y given Z, defined as

$$\underline{I}(X; Y|Z) = \sup \left\{ \alpha \Big| \lim_{n \to \infty} \Pr \left\{ i(X^n, Y^n|Z^n) < n\alpha \right\} = 0 \right\},$$

with $i(X^n, Y^n \mid Z^n) = \log \frac{P_{X^n Y^n|Z^n}(X^n, Y^n|Z^n)}{P_{X^n|Z^n}(X^n|Z^n)P_{Y^n|Z^n}(Y^n|Z^n)}$.

Now, consider a source model with n-independent but not necessarily identically distributed $(n-$INID$)$ observations (X^n, Y^n, Z^n), and with the associated joint probability distribution $P_{XYZ} = \prod_j P_{X_j Y_j Z_j}$. Further assume that the Markov relation $X_j - Y_j - Z_j$ holds for all $j \leqslant n$. This particular source model is obviously an special case of the general source model of [31], and its capacity is given by Equation (3.29) and is not captured by the OW-WSK capacity result of [20] (see Equations (3.17), and (3.18)). One realization of such a

source model is the wireless network scenario in which there are two independent binary discrete memoryless channels: one from Y to X, and one form Y to Z, without requiring the channels to stay the same over independent transmissions. In practice, wireless channels are time-varying due to multiple signal paths, user mobility, shadowing effects, etc [45]. One example of a more realistic channel model realization of such $n-$INID source models is the fading binary symmetric channels (F-BSC). In this case, if we assume that the broadcast channel $Y \mapsto (X, Z)$ is composed of two independent F-BSCs, then the resulting source distribution is $n-$INID.

We show in the following that our proposed OW-SKA Protocol 4 also achieves the wiretap secret key capacity of such a general source when the Markov relation $X^n - Y^n - Z^n$ holds and the probability distribution of $P_{X^n Y^n Z^n}$ is $n-$INID.

Theorem 3.11. *For a given general source model with n-independent observations (X^n, Y^n, Z^n) with the Markov relation $X^n - Y^n - Z^n$, we have*

$$C_{WSK}(P_{XYZ}) = C_{WSK}^{\rightarrow}(P_{XYZ}) = \liminf_{n \to \infty} \frac{1}{n} I(X^n; Y^n | Z^n),$$

which is achievable by the OW-SKA Protocol 4.

Proof of Theorem 3.11: The proof has two parts: the direct part (achievability) and the converse (upper bound).

The proof of achievability directly follows from Theorem 3.8. Recall that the maximum asymptotic SK length achievable by Protocol 4 is given by $S = H(X^n|Z^n) - H(X^n|Y^n) + o(n)$. When the Markov relation $X^n - Y^n - Z^n$ holds, the maximum achievable asymptotic SK rate is

$$r_K(\mathbf{\Pi_{HH}}) = \liminf_{n \to \infty} \frac{1}{n} I(X^n; Y^n | Z^n),$$

which completes the proof of achievability.

For the proof of converse we first need the following Lemma.

Lemma 3.12 (Adapted from Lemma 3.2.4 of [24], page 185 – also see [95]). *For any given arbitrary sequence of random variables (X^n, Y^n, Z^n) and for all n we have*

$$\mathbb{E}_{P_{XYZ}}\left\{i(X^n, Y^n|Z^n)\mathbb{1}\left[i(X^n, Y^n|Z^n) \leqslant 0\right]\right\} \geqslant \frac{1}{e}\log\frac{1}{e},$$

where $\mathbb{1}[\cdot]$ denotes the indicator function defined by

$$\mathbb{1}[\mathtt{statement}] = \begin{cases} 1 & \textit{if } \mathtt{statement} \textit{ is True}, \\ 0 & \textit{otherwise}. \end{cases}$$

Proof of Lemma 3.12: Recall that

$$i(X^n, Y^n \mid Z^n) = \log\frac{P_{X^n Y^n Z^n}(X^n, Y^n, Z^n)}{P_{Z^n}(Z^n)P_{X^n|Z^n}(X^n \mid Z^n)P_{Y^n|Z^n}(Y^n \mid Z^n)}$$

and let

$$\rho(x^n, y^n, z^n) = \frac{P_{X^n Y^n Z^n}(x^n, y^n, z^n)}{P_{Z^n}(z^n)P_{X^n|Z^n}(x^n \mid z^n)P_{Y^n|Z^n}(y^n \mid z^n)}.$$

Then we have

$$\mathbb{E}_{P_{XYZ}}\left\{i(X^n, Y^n|Z^n)\mathbb{1}\left[i(X^n, Y^n|Z^n) \leqslant 0\right]\right\}$$

$$= \sum_{\substack{x^n, y^n, z^n \\ \text{s.t. } \rho(x^n, y^n, z^n) \leqslant 1}} P_{X^n Y^n Z^n}(x^n, y^n, z^n) \log \rho(x^n, y^n, z^n)$$

$$= \sum_{\substack{x^n, y^n, z^n \\ \text{s.t. } \rho(x^n, y^n, z^n) \leqslant 1}} P_{Z^n}(z^n)P_{X^n|Z^n}(x^n \mid z^n)P_{Y^n|Z^n}(y^n \mid z^n)\rho(x^n, y^n, z^n) \log \rho(x^n, y^n, z^n)$$

$$\geqslant \sum_{\substack{x^n, y^n, z^n \\ \text{s.t. } \rho(x^n, y^n, z^n) \leqslant 1}} P_{Z^n}(z^n)P_{X^n|Z^n}(x^n \mid z^n)P_{Y^n|Z^n}(y^n \mid z^n)\frac{1}{e}\log\frac{1}{e}$$

$$\geqslant \frac{1}{e}\log\frac{1}{e},$$

where the last two inequalities hold since $\rho \log \rho \geqslant \frac{1}{e}\log\frac{1}{e}$ for any $0 \leqslant \rho \leqslant 1$. ∎

Now, for the converse we prove the following upper bound

$$C_{WSK}(P_{XYZ}) = \underline{I}(X;Y|Z) \leqslant \liminf_{n \to \infty} \frac{1}{n} I(X^n; Y^n|Z^n).$$

For any arbitrary $\delta > 0$ we have

$$
\begin{aligned}
\frac{1}{n} I(X^n; Y^n|Z^n) &= \frac{1}{n} \mathbb{E}_{P_{X^nY^nZ^n}} \left\{ i(X^n, Y^n|Z^n) \right\} \\
&= \frac{1}{n} \sum_{i(x^n, y^n|z^n) \leqslant 0} P_{X^nY^n|Z^n}(x^n, y^n|z^n) i(x^n, y^n|z^n) \\
&\quad + \frac{1}{n} \sum_{0 < i(x^n, y^n|z^n) \leqslant n\underline{I}(X;Y|Z) - n\delta} P_{X^nY^n|Z^n}(x^n, y^n|z^n) i(x^n, y^n|z^n) \\
&\quad + \frac{1}{n} \sum_{i(x^n, y^n|z^n) > n\underline{I}(X;Y|Z) - n\delta} P_{X^nY^n|Z^n}(x^n, y^n|z^n) i(x^n, y^n|z^n) \\
&\geqslant \frac{1}{n} \sum_{i(x^n, y^n|z^n) \leqslant 0} P_{X^nY^nZ^n}(x^n, y^n, z^n) i(x^n, y^n|z^n) \\
&\quad + \frac{1}{n} \sum_{i(x^n, y^n|z^n) > n\underline{I}(X;Y|Z) - n\delta} P_{X^nY^nZ^n}(x^n, y^n, z^n) i(x^n, y^n|z^n) \\
&\geqslant \frac{1}{n} \left(\frac{1}{e} \log \frac{1}{e} \right) + (\underline{I}(X;Y|Z) - \delta) \Pr\left\{ \mathcal{E}^n \right\},
\end{aligned}
$$

where $\mathcal{E}^n = \{(x^n, y^n, z^n) | i(x^n, y^n|z^n) > n\underline{I}(X;Y|Z) - n\delta\}$ and the second inequality is due to Lemma 3.12. According to the definition of $\underline{I}(X;Y|Z)$ we know that $\lim_{n \to \infty} \Pr\left\{ \mathcal{E}^n \right\} = 1$. Thus, by taking the limit we get

$$\underline{I}(X;Y|Z) \leqslant \liminf_{n \to \infty} \frac{1}{n} I(X^n; Y^n|Z^n) + \delta,$$

which holds for any arbitrary δ. Then by $\delta \to 0$ the proof is complete. ∎

3.4.2 A Practically Efficient One-way SKA Protocol

The OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) that achieves the best known finite-length lower bound given in Proposition 3.10, is not computationally efficient – it has computational complexity

---

Protocol 5: OW-SKA with polar coding IR and universal hashing PA ($\mathbf{\Pi_{PH}}$)

    Public Information: P_{XYZ}
    Input: $n-$fold samples $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$, ϵ, σ
    Output: Key estimates k_A and k_B

    `// Initialization`

1 Alice and Bob use one-way public communication to agree on (i) a polar coding
    scheme (with parity check matrix H and syndrome decoding protocol $f_{IR}^{-1}(\cdot, \cdot)$),
    and (ii) a random seed $s \in \mathcal{S}$, and 2-UHF $f_{PA} = h_s : \mathcal{X}^n \to \{0,1\}^\ell$.

    `// Information Reconciliation`

2 Alice sends Bob $F = f_{IR}(x^n) = H \cdot (x^n)^\intercal$.
3 Bob recovers x^n by $\hat{x^n} = f_{IR}^{-1}(F, y^n)$.

    `// Privacy Amplification`

4 Alice and Bob find $k_A = f_{PA}(x^n) = h_s(x^n)$, and $k_B = f_{PA}(\hat{x^n}) = h_s(\hat{x^n})$.

---

of $\mathcal{O}(n \log n)$ for Alice, and computational complexity of $\mathcal{O}(2^{nH(X|Y)})$ for Bob (see Section 3.4.3). This lower bound can also be indirectly attained using the results of [25, 82, 85]; and, non of the above approaches are computationally efficient. Therefore, to design an efficient OW-SKA protocol, we consider a protocol that runs a practical (computationally efficient) IR protocol that is based on Polar Codes [29] [96–98], followed by PA via universal hashing [30]. Recall that a family $\mathcal{H}$ of functions $h_s : \mathcal{X} \to \mathcal{K}$, with $s \in \mathcal{S}$, is called 2-universal if, for all $x \neq x' \in \mathcal{X}$, we have $\Pr\{h_s(x) = h_s(x')\} \leqslant {}^1/_{|\mathcal{K}|}$.

We consider IID sources in which $X - Y - Z$ holds, P_X is binary uniform, and the relation between X and Y is given by $Y = W(X)$, where W is a binary-input memoryless symmetric (BMS) channel with transition probability matrix $P_{Y|X}$. For this case, there is no need to find RV's (V, U) (see Remark 3.1). Let H be the parity check matrix of a polar code designed for W. Define $f_{IR}(X^n) = H \cdot (X^n)^\intercal$, and let f_{IR}^{-1} be a corresponding efficient syndrome decoding [99–102] for f_{IR}. Also, let $f_{PA} = h_s$ be a 2-universal hash function with seed s drawn uniformly at random from $\mathcal{S}$. Before, starting the IR and PA phases, parties agree on f_{IR}, f_{IR}^{-1}, and $f_{PA} = h_s$. The construction of the OW-SKA protocol is given in Protocol 5.

Designing one-way IR protocols based on channel coding schemes is in fact a common approach [76, 78–80, 82, 103]. The length of the public message generated by such IR protocols can be calculated from the channel coding rate. The optimum finite-length behavior of channel coding rate is due to [50], however, the corresponding coding scheme is not computationally efficient. Polar codes [29], on the other hand, are gaining a lot of attention in many real-life applications as they have efficient computational complexity of $\mathcal{O}(n \log n)$, and they can achieve the capacity of any BMS. The finite-length behavior of polar codes has been studied extensively over the past years [96–98]. It is known that (for a fixed error bound of ϵ) the polar coding rate is given by $r = C(W) - \sqrt[\tau]{\lambda(\epsilon)/n} \pm o(\sqrt{1/n})$, where $C(W) = I(X;Y)$ is channel capacity of W, and $\lambda(\epsilon)$ is a positive number that is a function of ϵ [96–98]. Therefore, we can prove the following.

Theorem 3.13. *In the $n-IID$ regime, the one-way SKA Protocol 5 achieves an $(\epsilon, \sigma)-SK$ of length*

$$\ell_{\mathbf{\Pi_{PH}}}(n) = nC_{WSK}^{\rightarrow} - \sqrt[\tau]{n^{\tau-1}}G_{IR}(\epsilon) - \sqrt{n}G_{PA}(\sigma) \pm o(\sqrt{n}), \tag{3.30}$$

where, $G_{PA}(\sigma) = \sqrt{\Delta_{X|Z}}Q^{-1}(\sigma)$, $G_{IR}(\epsilon) = \sqrt[\tau]{\lambda(\epsilon)}$, $C_{WSK}^{\rightarrow} = C_{WSK} = H(X|Z) - H(X|Y)$, and $\lambda(\epsilon)$ is a positive number which is a function of ϵ.

To prove Theorem 3.13 we need the following Lemma.

Lemma 3.14 (Spectral LHL). *For any $P_{XZF} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z} \times \mathcal{F})$, let $f_{PA} = h_s : \mathcal{X} \to \mathcal{K}$ be a 2-universal hash function with seed s drawn uniformly at random from $\mathcal{S}$. Then*

$$\mathbf{SD}(KZFS, UZFS) \leqslant \epsilon + \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\underline{H}_s^\epsilon(X|Z)}},$$

where $\underline{H}_s^\epsilon(X|Z)$ is the inf-spectral entropy of X given Z (Definition 3.4), $K = h_S(X)$ and U is uniform over $\mathcal{K}$. This implies that if

$$\log |\mathcal{K}| \leqslant \underline{H}_s^{\sigma-\mu}(X|Z) - \log |\mathcal{F}| + \log 4\mu^2,$$

85

then, function $f_{PA} = h_s$ is $\sigma-$secure for any $0 < \mu \leqslant \sigma$.

The proof is in the Appendix (see Section 3.6.3).

Proof of Theorem 3.13: If the IR protocol is $\epsilon-$correct and the PA protocol is $\sigma-$secure, then the generated key is an $(\epsilon, \sigma)-$SK $K \in \mathcal{K}$. For observation length n, we design a polar code so that the error probability is bounded by ϵ [96–98]. Let $t = \log |\mathcal{F}|$ be the length of the public message F generated by the IR protocol. Then, $t = \log |\mathcal{F}|$ is equal to the length of syndrome vector which is given by $t = n - nr = n(H(X) - C(W)) + n\sqrt[\tau]{\lambda(\epsilon)/n} \pm o(\sqrt{n})$. Note that $H(X) - C(W) = H(X|Y)$.

By Lemma 3.14

$$\ell = \underline{H}_s^{\sigma-\mu}(X^n|Z^n) - t + \log 4\mu^2, \tag{3.31}$$

is an achievable key length, where $0 < \mu \leqslant \sigma$. Applying Lemma 3.2, with $\mu = 1/\sqrt{n}$ and noting that $Q^{-1}(\sigma - 1/\sqrt{n}) = Q^{-1}(\sigma) + \mathcal{O}(1/\sqrt{n})$, completes the proof. ∎

For conventional Polar Codes (that use the original 2×2 polarization kernel of Arikan [29]), it is known that τ (the scaling factor) is around 4, and more precisely $3.579 \leqslant \tau \leqslant 4.714$ [96, 97], where the exact value of τ depends on the underlying channel W. The function $\lambda(\epsilon)$ is not know in general, and finding a universal expression for $\lambda(\epsilon)$ requires furtherer research [97]. Recently, by using large polarization kernels, the scaling factor of Polar Codes, for Binary Erasure Channel (BEC) was improved to $\tau = 2 + \delta$ where $\delta \in (0, 1]$ is an arbitrary parameter [98]. For this specific coding scheme $\lambda(\epsilon) = l(\delta)(1 + 2/\epsilon)^3$ where, $l(\delta) \in \mathcal{O}(\exp(\delta^{-1.01}))$ is the kernel size. Computational complexity is $\mathcal{O}(n \log n)$, as long δ and the kernel size l are fixed (do not depend on n).

Remark 3.5 (Public communication cost of $\mathbf{\Pi_{PH}}$). For the source model described by $n-$IID distribution P_{XYZ}, let $\log(\text{supp}(F(\mathbf{\Pi_{PH}})))$ denote the public communication cost (in bits) that is used by the OW-SKA Protocol 5 ($\mathbf{\Pi_{PH}}$) to achieve $S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n)$. Then, our

results show that

$$\log(\mathrm{supp}(F(\mathbf{\Pi}_{\mathrm{PH}}))) = nH(X|Y) + \sqrt[\tau]{n^{\tau-1}}G_{IR}(\epsilon) + o(\sqrt{n}), \qquad (3.32)$$

where $G_{IR}(\epsilon) = \sqrt[\tau]{\lambda(\epsilon)}$.

3.4.3 Comparing $\mathbf{\Pi_{HH}}$ and $\mathbf{\Pi_{PH}}$ with other related protocols

We compare our proposed OW-SKA protocols, $\mathbf{\Pi_{HH}}$ and $\mathbf{\Pi_{PH}}$, with other known capacity achieving OW-SKA protocols and the interactive protocol of [31]. The comparison is based on the type of information reconciliation step, SK length, public communication cost, and computation complexity of Alice and Bob. The comparison is summarized in Table 3.1. Specifically, at the end, through two numerical examples, we compare the finite key performance of $\mathbf{\Pi_{HH}}$ and $\mathbf{\Pi_{PH}}$ against the protocols of [31, 76, 78].

HR05. Holenstein and Renner proposed a protocol in [76] that achieves the OW-WSK capacity of a general distribution. The reconciliation message uses linear codes. We derive two finite-length lower bounds for the two variations of their SKA [78] (one uses a random linear code, and the second one uses the concatenation of a linear code with a Reed-Solomon code). The bounds given in Theorem 3.13, and Theorem 3.15 of [78], are expressed in the form of $n(C_{WSK}^{\rightarrow} - \delta(\kappa_1, \kappa_2))$, where $n\kappa_1 = \log(1/\epsilon)$ and $n\kappa_2 = \log(1/\sigma)$. We re-derived these bounds as functions of ϵ and σ in the following proposition.

Proposition 3.15. *For any source model with IID distribution P_{XYZ}, let $R_n = H(X^n|Z^n) - H(X^n|Y^n)$. Then for large enough n and any $\epsilon, \sigma < 1/4$, we have*

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n) \geqslant [R_n - \sqrt{n}f_{\epsilon,\sigma}']^+, \qquad (3.33)$$

$$S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n) \geqslant [R_n - \sqrt[4]{n^3}g_{\epsilon,\sigma}'' - \sqrt{n}f_{\epsilon,\sigma}'']^+, \qquad (3.34)$$

where $[a]^+ = \max\{0, a\}$, *and*

$$f'_{\epsilon,\sigma} = 90 \log(|\mathcal{X}||\mathcal{Y}|)(\sqrt{\log 1/\epsilon} + \sqrt{\log 1/\sigma}),$$

$$g''_{\epsilon,\sigma} = \sqrt[4]{2^{22} \log(1/\epsilon) \log^2(|\mathcal{X}|) \log^2(|\mathcal{X}||\mathcal{Y}|)},$$

$$f''_{\epsilon,\sigma} = 8 \log(|\mathcal{X}|)\sqrt{\log(1/\sigma)}.$$

Bound (3.33) corresponds to random linear codes and bound (3.34) is due to concatenated codes. For both lower bounds the SKA protocol uses $\mathcal{O}(n^2)$ bits of communication. The computation complexity of Alice corresponding to both bounds is in $\mathcal{O}(n^2)$. The computation complexity of Bob is in $\mathcal{O}(n^2)|\mathcal{X}|^n$ and $\mathcal{O}(n^2)$, respectively. As it is mentioned in [78] and [79], the computation complexity of (3.34) for Alice and Bob is *efficient* (i.e., in $\mathcal{O}(n^d)$) but it is not *practically efficient* (i.e., it is not in $\mathcal{O}(n)$ or $\mathcal{O}(n \log n)$).

RRS13. Renes et. al proposed an SKA protocol that uses polar codes for both reconciliation and privacy amplification [79]. The implementation cost of the protocol is $\mathcal{O}(n \log n)$ for both Alice and Bob, but the code construction for any given distribution might not be straightforward [79, Section III.C]. The protocol uses a message of length $\mathcal{O}(n)$. Their analysis of the protocol does not provide finite-length approximation of the key length.

CBA15. In [80], authors proposed an SKA protocol using polar codes. The reconciliation and privacy amplification are combined in a single step polar coding. The protocol requires a small pre-shared secret seed of length $\mathcal{O}(2^{-a.n})$. It uses $\mathcal{O}(n)$ bits of public communication and its analysis does not give any finite-length approximations for the key length.

HTW16. The interactive protocol of [31] gives the tightest known bounds for two-party SKA; however, their protocol is interactive with $\mathcal{O}(n)$ rounds of public communication and its IR step is based on random binning and spectrum slicing[4] that implies exponential com-

---

[4]Spectrum slicing is a well-established spectrum technique for information theoretic tasks [24]

putation complexity ($\mathcal{O}(2^n)$) for Alice and Bob. The public communication cost of this SKA protocol is in $\mathcal{O}(nH(X|Y) + n)$.

$\mathbf{\Pi_{HH}}$ (Protocol 4). This protocol is very efficient in terms of public communication, as it uses a single message of length $\mathcal{O}(nH(X|Y))$ (See Remark 3.4). This protocol gives achievability finite-length bounds as given in (3.22) and (3.24) (that are far closer to the capacity upper bound than the finite-length bounds of HR05 and $\mathbf{\Pi_{PH}}$). The computation cost of Alice is practically efficient; i.e., $\mathcal{O}(n \log n)$ (computing a single hash value [66]). But, unfortunately for Bob, the computation cost is in $\mathcal{O}(2^{n.H(X|Y)})$; i.e., the implementation is not efficient.

$\mathbf{\Pi_{PH}}$ (Protocol 5). This OW-SKA protocol is efficient in terms of computation complexity and public communication. It's computation complexity for Alice and Bob is in $\mathcal{O}(2^l n \log n)$ and its public communication cost is $\mathcal{O}(nH(X|Y))$ (See Remark 3.5). We gave finite-length analysis of $\mathbf{\Pi_{PH}}$ and our following numerical example shows that $\mathbf{\Pi_{PH}}$ has far better finite key performance that the OW-SKA of HR05 [76].

Numerical Examples

Here, we give numerical examples to compare our proposed OW-SKA protocols 4 and 5 ($\mathbf{\Pi_{HH}}$ and $\mathbf{\Pi_{PH}}$) with previous results of [31, 76, 78]. The OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) exhibits the best known finite-length performance among all OW-SKAs. The downfall of Protocol 4 ($\mathbf{\Pi_{HH}}$) however, is its high computation complexity for Bob. The OW-SKA Protocol 5 ($\mathbf{\Pi_{PH}}$), has lower computational complexity; but in comparison to $\mathbf{\Pi_{HH}}$ it requires more samples to generate the key. In Example 3.3 we observe that, in a source model with WSK capacity of 0.5 bits per sample, the OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) requires $n = 590$ source samples (one bit each) to generate a key of length $\ell = 256$ bits, and hence the key rate is $\ell/n \approx 0.43$. However, we show that for the same source model, the OW-SKA Protocol 5

Table 3.1: The comparison of Protocols 4 and 5 ($\mathbf{\Pi_{HH}}$ and $\mathbf{\Pi_{PH}}$) with other protocols.

Protocol	HR05	RRS13 & CBA15	HTW16	$\mathbf{\Pi_{HH}}$	$\mathbf{\Pi_{PH}}$
One-way or Interactive	One-way	One-way	Interactive with $\mathcal{O}(n)$ rounds	One-way	One-way
IR Approach	linear codes	polar codes	random binning	universal hashing	polar codes
PA Approach	universal hashing	polar codes	universal hashing	universal hashing	universal hashing
Communication Cost	$\mathcal{O}(n^2)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$	$\mathcal{O}(n)$
Computation Cost Alice	$\mathcal{O}(n^2)$	$\mathcal{O}(n \log n)$	$\mathcal{O}(2^n)$	$\mathcal{O}(n \log n)$	$\mathcal{O}(n \log n)$
Computation Cost Bob	$\mathcal{O}(n^2)$ [a]	$\mathcal{O}(n \log n)$	$\mathcal{O}(2^n)$	$\mathcal{O}(2^n)$	$\mathcal{O}(n \log n)$
Finite-length Analysis	(Prp. 3.15)	N/A	(Rem. 3.2)	(Thm. 3.8 & 3.9)	(Thm. 3.13)

(a) For the lower bound given in (3.33) the computation complexity is in $\mathcal{O}(2^n)$.

($\mathbf{\Pi_{PH}}$) needs $n = 2477 \times 10^3$ source samples to generate a key of the same length ($\ell = 256$ bits,) and thus the key rate is $\ell/n \approx 10^{-4}$.

Example 3.1. Here, we compare finite-length behavior of OW-SKA Protocol 4 ($\mathbf{\Pi_{HH}}$) and interactive SKA of [31]. Consider an IID source model P_{XYZ}, where $X - Y - Z$ holds. P_X is uniform, $Y = BSC_a(X)$, and $Z = BSC_b(Y)$. Here, BSC_a denotes a Binary Symmetric Channel with bit flip probability a. For this case the WSK (and OW-WSK) capacity is given by $C_{WSK} = C_{WSK}^{\rightarrow} = h_2(a * b) - h_2(a)$, where $a * b = a(1 - b) + b(1 - a)$, and h_2 is the binary entropy. Let, $a = 0.02$, and $b = 0.15$, then $C_{WSK}^{\rightarrow} = 0.502$. Set $\epsilon = \sigma = 0.05$, and let observation length $n \in [2000, 50000]$. Figure 3.1 shows the increase in the key rate (ℓ/n) when the observation length grows. The depicted one-way lower bounds are from Proposition 3.10 (with $U = X$ and $V = $ constant) and Theorem 3.8; and the interactive
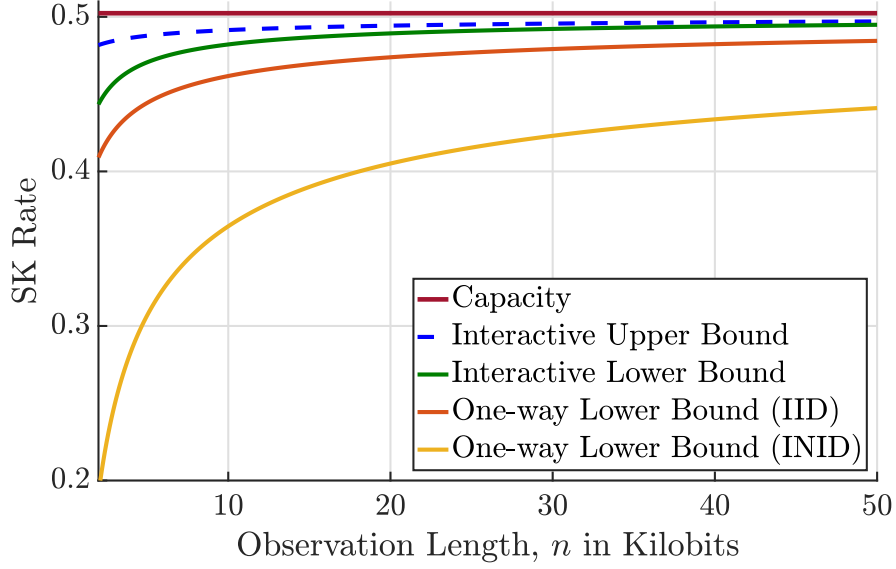
Figure 3.1: Optimum finite-length bounds of interactive SKA (Theorem 15 of [31]), and the finite-length lower bounds of one-way SKA (Proposition 3.10 and Theorem 3.8). Here $\epsilon = \sigma = 0.05$, P_X is uniform, $Y = BSC_a(X)$, and $Z = BSC_b(Y)$, where $a = 0.02$, and $b = 0.15$. Note that in this example, as $X - Y - Z$ holds, both interactive and one-way bounds achieve the WSK capacity.

upper and lower bounds are from Theorem 15 of [31].

The one-way bounds we derived are useful in the sense that for $\mathbf{\Pi_{HH}}$, they give the finite-length gap to capacity (i.e., the difference between one-way lower bound and capacity) and the finite-length gap for not using interaction (i.e., the difference between the one-way and interactive lower bounds).

Note that the one-way lower bound of Proposition 3.10 that was derived for IID sources gives a tighter (and more appropriate) approximation of the key rate, whereas for this IID example source the lower bound of Theorem 3.8 that holds for INID (and IID) sources is less tight. Also we observe that in the observation length domain of this figure, the lower bounds of protocols of [76, 78] are 0 (the bounds are given Preposition 3.15).

Example 3.2. Let us now observe the finite-length behavior of $\mathbf{\Pi_{HH}}$ when parties observations are from variables that are independent but not identically distributed (INID). Note
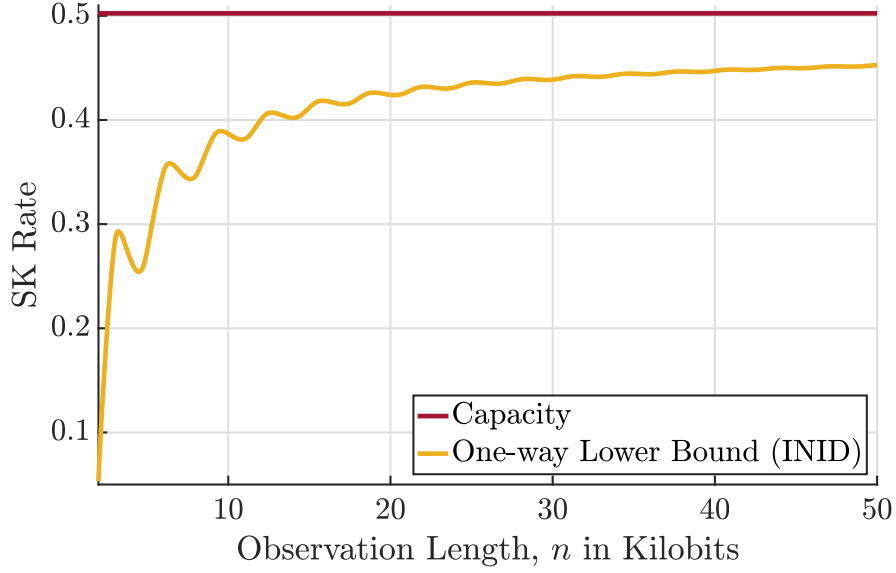
Figure 3.2: Finite-length performance of $\mathbf{\Pi_{HH}}$ for an INID source. The one-way lower bound is given in Theorem 3.8. Here $\epsilon = \sigma = 0.05$, P_X is uniform IID, $Y_n = BSC_{a_n}(X_n)$, and $Z_n = BSC_{b_n}(Y_n)$, where $a_n = 0.02 + \frac{500}{n}\sin\left(\frac{n}{500}\right)$, and $b_n = 0.15$. Here $X_n - Y_n - Z_n$ holds for all n, and both interactive and one-way SKA approaches achieve the WSK capacity.

that in this case the only SKA protocol with finite-length analysis is $\mathbf{\Pi_{HH}}$ and only the one-way $\mathbf{\Pi_{HH}}$ and interactive SKA protocol of [31] can achieve the WSK capacity of such INID sources. (The analysis of [31] does not give finite key length.)

Consider an INID source model $P_{X^n Y^n Z^n} = \prod_{j=1}^{n} P_{X_j Y_j Z_j}$, where $X_n - Y_n - Z_n$ holds for all n. $P_{X_n} = P_X$ is IID uniform, $Y_n = BSC_{a_n}(X_n)$, and $Z_n = BSC_{b_n}(Y_n)$, where $a_n = 0.02 + \frac{500}{n}\sin\left(\frac{n}{500}\right)$ for all $n \geqslant 1$, and $b_n = b = 0.15$. Here, BSC_a denotes a Binary Symmetric Channel with bit flip probability a. In this INID source model, the BSC channel from Alice to Bob varies over time. For this case the WSK (and OW-WSK) capacity is given
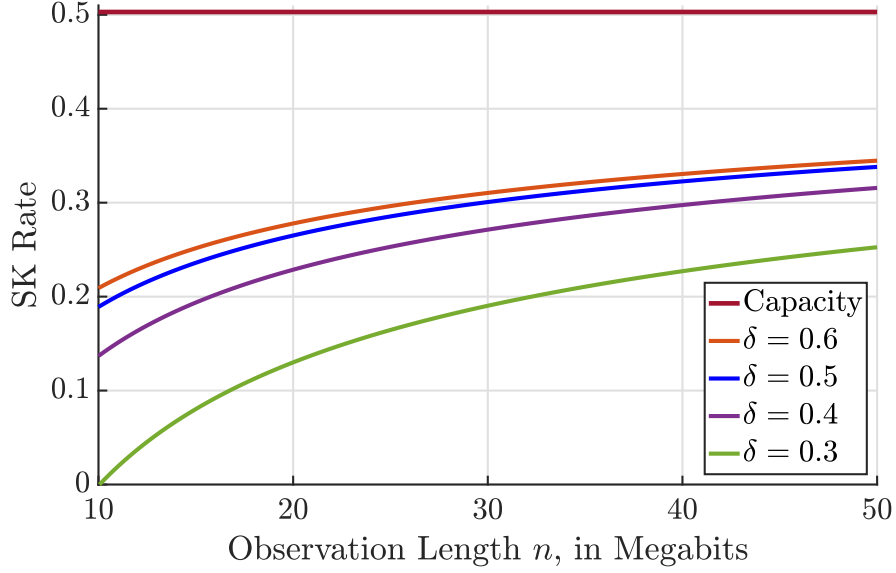
Figure 3.3: The finite-length lower bounds given in Theorem 3.13 for different δ's in, $(0.3, 0.4, 0.5, 0.6)$. These values correspond to polarization kernel sizes of $(30, 13, 8, 6)$ (in the same order). Here $\epsilon = \sigma = 0.05$, P_X is uniform, $Y = BEC_a(X)$, and $Z = BEC_b(Y)$, where $a = 0.1$, and $b = 0.67$.

by Theorem 3.11, then

$$
\begin{aligned}
C_{WSK} &= \liminf_{n \to \infty} \frac{1}{n} I(X^n; Y^n | Z^n) \\
&= \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} I(X_j; Y_j | Z_j) \\
&= \lim_{n \to \infty} I(X_n; Y_n | Z_n) \\
&= h_2(\bar{a} * \bar{b}) - h_2(\bar{a}) \\
&= 0.502
\end{aligned}
$$

where $\bar{a} = \lim_{n \to \infty} a_n = 0.02$ and $\bar{b} = \lim_{n \to \infty} b_n = 0.15$.

Set $\epsilon = \sigma = 0.05$, and let observation length $n \in [2000, 50000]$. In Figure 3.2, that shows the increase in the key rate (ℓ/n) when n grows, we depict the finite-length performance of the capacity achieving OW-SKA protocol 4 ($\mathbf{\Pi_{HH}}$).

93

Example 3.3. Now we compare the finite-length behavior of one-way SKA Protocol 5 $\mathbf{\Pi_{PH}}$ with the one-way SKA protocols of [76, 78]. Remember that $\mathbf{\Pi_{PH}}$ can be used when source model is IID and P_X is uniform.

Consider an IID source model P_{XYZ}, where $X - Y - Z$ holds. P_X is uniform, $Y = BEC_a(X)$, and $Z = BEC_b(Y)$. Here, BEC_a denotes a Binary Erasure Channel with erasure probability a. For this case the OW-WSK capacity is given by $C_{WSK}^{\rightarrow} = b(1-a) - a$. Let, $a = 0.1$, and $b = 0.67$, then $C_{WSK}^{\rightarrow} = 0.503$. Set $\epsilon = \sigma = 0.05$, and let observation length $n \in [10 \times 10^6, 50 \times 10^6]$. We compare our finite-length approximations of achievable key rates that use Polar Coding scheme of [98] for their IR protocol, and use 2-universal hashing for PA. For this case, the SK length is given by ℓ in Equation (3.30). The scaling factor of polarization is $\tau + \delta$, and we set constant $\lambda(\epsilon) = \exp(\delta^{-1.01})(1 + 2/\epsilon)^3$. In Figure 3.3, we plot ℓ/n for different values of δ. Note that we fix the kernel size of IR coding to $l = \lceil \exp(\delta^{-1.01}) \rceil$, and this SKA protocol has computation complexity of $\mathcal{O}(n \log n)$. It is interesting to note that due to the dependence of $\lambda(\epsilon)$ on δ, we observe better finite-length performance, for larger values of δ (smaller sizes of kernel) in this finite-length domain. This pattern does not continue for larger values of n. In the observation length domain of this figure, the interactive and one-way second-order bounds are indistinguishable from the Capacity and the lower bounds of protocols in [76, 78] (that have complexity $\mathcal{O}(n^2)$) are 0.

3.5 Conclusion

We studied OW-SKA protocols in source model. These protocols are important for practical reasons as they do not require any interaction. They are also important from the theoretical viewpoint as they can achieve the OW-WSK capacity, and also are related to problems in computational cryptography. Inspired by the inf-spectral entropy of [25], we introduced the sup-spectral entropy, and utilized these two spectral entropies to prove a new multi-letter finite-length upper bound on the key length. We then proposed two new OW-SKA

constructions that are capacity achieving. Our first construction (Protocol 4) uses a reconciliation method that is inspired by information spectrum analysis of [31]. Our analysis of this protocol led to two finite-length lower bounds for the maximum achievable key length of OW-SKA. This protocol can also be used in the more general case when parties' observations are drawn from independent experiments. The second proposed construction (Protocol 5) employs Polar coding for reconciliation and thus is computationally efficient. We derived the maximum achievable key length of this OW-SKA which constitutes a lower bound on $S_{\epsilon,\sigma}^{\rightarrow}(X^n, Y^n|Z^n)$. Both of our protocols are very efficient in terms of public communication cost. A detailed comparison of these proposed OW-SKA protocols with other related protocols, including numerical examples, were given at the end. An interesting future work is to find a tight second-order finite-length converse (upper bound) for OW-SKA. It is also of practical importance to find efficient OW-SKA protocols that have better finite-length performance than our proposed Protocol 5. Another intriguing research avenue is to investigate the application of spectrum techniques in the multiterminal model of SKA [21].

3.6 Appendix

3.6.1 Proof of Smooth LHL

In order to prove Lemma 3.7 we use the framework of average min-entropy [64] and its relationship with smooth min-entropy.

Definition 3.5 (Average min-entropies [64]). For any joint probability distribution $P_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ the average conditional min-entropy of X given Y is defined by

$$\tilde{H}_{\min}(X|Y) = -\log \mathbb{E}_{P_Y} \max_{x \in \mathcal{X}} P_{X|Y}(x|y),$$

and the smoothed average conditional min-entropy of X given Y is defined by

$$\tilde{H}_{\min}^{\epsilon}(X|Y) = \max_{Q_{\hat{X}\hat{Y}} \in \hat{\mathcal{B}}^{\epsilon}(P_{XY})} \tilde{H}_{\min}(\hat{X}|\hat{Y}),$$

where $\hat{\mathcal{B}}^{\epsilon}(P_{XY}) = \{Q_{\hat{X}\hat{Y}} \in \overline{\mathcal{P}}(\mathcal{X} \times \mathcal{Y}) : \mathbf{SD}(\hat{X}\hat{Y}, XY) \leqslant \epsilon\}$, with $\overline{\mathcal{P}}(\mathcal{X})$ denoting the set of all sub-normalized positive distributions on $\mathcal{X}$.

Lemma 3.16. *For any $\epsilon > 0$ and RVs X and Y, we have $\tilde{H}_{\min}^{\epsilon}(X|Y) \geqslant \hat{H}_{\min}^{\epsilon}(X|Y)$.*

Proof of Lemma 3.16: We start with the left hand side of the above relation:

$$
\begin{aligned}
\tilde{H}_{\infty}^{\epsilon}(X|Y) &= \max_{Q_{X'Y'} \in \hat{\mathcal{B}}^{\epsilon}(P_{XY})} -\log \mathbb{E}_{P_{Y'}} \max_{x} Q_{X'|Y'}(x|y) \\
&\geqslant \max_{Q_{X'Y'} \in \hat{\mathcal{B}}^{\epsilon}(P_{XY})} -\log \max_{y} \max_{x} Q_{X'|Y'}(x|y) \\
&= \max_{Q_{X'Y'} \in \hat{\mathcal{B}}^{\epsilon}(P_{XY})} \min_{x,y} \log \frac{Q_{Y'}(y)}{Q_{X'Y'}(x,y)} \\
&= \max_{Q_{X'Y'} \in \hat{\mathcal{B}}^{\epsilon}(P_{XY})} \min_{x,y} \left(\log \frac{P_Y(y)}{Q_{X'Y'}(x,y)} + \log \frac{Q_{Y'}(y)}{P_Y(y)} \right) \\
&\geqslant \hat{H}_{\min}^{\epsilon}(X|Y) + \max_{Q_{Y'}} \min_{y} \log \frac{Q_{Y'}(y)}{P_Y(y)} \\
&\geqslant \hat{H}_{\min}^{\epsilon}(X|Y),
\end{aligned}
$$

where the last inequality is due to the choice of $Q_{Y'}(y) = P_Y(y)$.   ■

Proof of Lemma 3.7: The privacy amplification (key extraction) function $f_{PA} = h_S$ is $\sigma-$secure if

$$\mathbf{SD}(KZFS, UZFS) \leqslant \sigma,$$

where $K = h_S(X)$. The generalized Leftover Hash Lemma of [64, Lemma 2.4] states that for any probability distribution $P_{\bar{X}\bar{Z}F}$ if we let S be the uniform seed of a 2-universal hash function h_S, and define $\bar{K} = h_S(\bar{X})$, then

$$\mathbf{SD}(\bar{K}S\bar{Z}F, US\bar{Z}F) \leqslant \frac{1}{2}\sqrt{|\mathcal{K}|2^{-\tilde{H}_{\min}(\bar{X}|\bar{Z}F)}}.$$

By Lemma 2.2 of [64] we know that

$$\tilde{H}_{\min}(\bar{X}|\bar{Z}F) \geqslant \tilde{H}_{\min}(\bar{X}|\bar{Z}) - t$$

where $t = \log|\mathcal{F}|$. Then we get

$$\mathbf{SD}((\bar{K}S\bar{Z}F), (US\bar{Z}F)) \leqslant \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\tilde{H}_{\min}(\bar{X}|\bar{Z})}}.$$

Now, let $\bar{X}$ and $\bar{Z}$ be such that for the given X and Z we have

$$\tilde{H}_{\min}^{\epsilon}(X|Z) = \tilde{H}_{\min}(\bar{X}|\bar{Z}).$$

Since $\mathbf{SD}(XZ, \bar{X}\bar{Z}) \leqslant \epsilon$, we have $\mathbf{SD}(KZ, \bar{K}\bar{Z}) \leqslant \epsilon$ and $\mathbf{SD}(Z, \bar{Z}) \leqslant \epsilon$ (See Corollary 2.1.1, and Lemma 2.2). By triangle inequality we have

$$\mathbf{SD}(KZFS, UZFS) \leqslant \mathbf{SD}(KZFS, \bar{K}\bar{Z}FS) + \mathbf{SD}(\bar{K}S\bar{Z}F, US\bar{Z}F) + \mathbf{SD}(U\bar{Z}FS, UZFS)$$

$$= \mathbf{SD}(KZ, \bar{K}\bar{Z}) + \mathbf{SD}(\bar{K}S\bar{Z}F, US\bar{Z}F) + \mathbf{SD}(\bar{Z}, Z)$$

Therefore,

$$\mathbf{SD}(KZFS, UZFS) \leqslant 2\epsilon + \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\tilde{H}^\epsilon_{\min}(X|Z)}}.$$

As from Lemma 3.16 (see also [64, Appendix B]) we know that $\tilde{H}^\epsilon_{\min}(X|Z) \geqslant H^\epsilon_{\min}(X|Z)$, and hence we get

$$\mathbf{SD}(KZFS, UZFS) \leqslant 2\epsilon + \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\hat{H}^\epsilon_{\min}(X|Z)}},$$

which implies that the 2-universal hash function is σ-secure as long as its output key length satisfies

$$\log|\mathcal{K}| \leqslant \hat{H}^{\frac{\sigma-\eta}{2}}_{\min}(X|Z) - \log|\mathcal{F}| + \log 4\eta^2,$$

for any $0 < \eta \leqslant \sigma$. ∎

3.6.2   A Fano-like inequality for sup-spectral entropy

Intuitively speaking, sup-spectral entropy captures the reliability (IR) aspect of SKA. Thus, we believe, this new spectral entropy is of independent interest for future work in the context of lossy single-shot Slepian-Wolf source coding. In the following we prove a Fano-like inequality for sup-spectral entropy.

Proposition 3.17. *Suppose RV's X and Y are such that $\Pr\{X \neq Y\} \leqslant \epsilon$. Then for any $\nu \in (0, 1 - \epsilon)$ we have $\bar{H}^{\epsilon+\nu}_s(X|Y) \leqslant -\log\nu$.*

Proof: Let $r = -\log\nu$ and define $\mathcal{T}_y = \{x|\ P_{X|Y}(x|y) \leqslant \nu\}$. Then

$$\begin{aligned}
P_{XY}\{-\log P_{X|Y}(x|y) \geqslant r\} &= P_{XY}\{P_{X|Y}(x|y) \leqslant \nu\} \\
&= \sum_{y\in\mathcal{Y}} P_Y(y) \sum_{x\in\mathcal{T}_y} P_{X|Y}(x|y) \\
&= \sum_{y\in\mathcal{Y}} P_Y(y)\Big(P_{X|Y}(x|y)\Big|_{\substack{x\in\mathcal{T}_y\\x=y}} + \sum_{\substack{x\in\mathcal{T}_y\\x\neq y}} P_{X|Y}(x|y)\Big) \\
&\leqslant \nu + \Pr\{X \neq Y\} \leqslant \nu + \epsilon.
\end{aligned}$$

∎

3.6.3   Proof of Spectral LHL

For a set $\mathcal{X}$, let $\mathcal{P}(\mathcal{X})$ be the set of all probability distributions on $\mathcal{X}$, and let $\overline{\mathcal{P}}(\mathcal{X})$ be the set of all sub-normalized positive distributions on $\mathcal{X}$. We can define the statistical distance and conditional Rényi Entropies for sub-normalized functions. For $\overline{P}_{\bar{X}\bar{Z}} \in \overline{\mathcal{P}}(\mathcal{X} \times \mathcal{Z})$ and any $P_Z \in \mathcal{P}(\mathcal{Z})$, and by denoting $(\bar{X}, \bar{Z}) \sim \overline{P}_{\bar{X}\bar{Z}}$, and $Z \sim P_Z$, we define

$$H_2(\bar{X}|Z) = -\log \sum_{x,z} \frac{(\overline{P}_{\bar{X}\bar{Z}}(x,z))^2}{P_Z(z)},$$

and

$$H_{\min}(\bar{X}|Z) = \min_{x,z} -\log \frac{\overline{P}_{\bar{X}\bar{Z}}(x,z)}{P_Z(z)}.$$

Even though such definitions do not have information theoretic operational meanings, they are useful tools for deriving bounds and/or characterizing meaningful information theoretic tasks that are defined over probability functions.

   Proof of Lemma 3.14:   The privacy amplification (key extraction) function $f_{PA} = h_s$ is $\sigma-$secure if

$$\mathbf{SD}(KZFS, UZFS) \leqslant \sigma.$$

The generalized Leftover Hash Lemma of [31, Lemma 8] states that, for any $\overline{P}_{\bar{X}\bar{Z}F} \in \overline{\mathcal{P}}(\mathcal{X} \times \mathcal{Z} \times \mathcal{F})$, let $\overline{P}_{\bar{Z}}$ be the marginal of $\overline{P}_{\bar{X}\bar{Z}F}$ for $\bar{Z}$, and let S be the uniform seed of a 2-universal hash function h_S. Then, for $\bar{K} = h_S(\bar{X})$ and any $P_Z \in \mathcal{P}(\mathcal{Z})$, where $\mathrm{supp}(\overline{P}_{\bar{Z}}) \subset \mathrm{supp}(P_Z)$ holds, we have

$$\mathbf{SD}(\bar{K}\bar{Z}FS, U\bar{Z}FS) \leqslant \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-H_2(\bar{X}|Z)}},$$

where U is the uniform distribution over $\mathcal{K}$.

   Now consider $P_{XZF} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z} \times \mathcal{F})$ and its marginals, P_{XZ}, P_X, and P_Z. For $X \sim P_X$ let $K = h_S(X)$, where h_S is a 2-universal hash function. Let $r = \underline{H}^\epsilon_s(X|Z)$ and defined the

following sub-normalized function

$$\overline{P}_{\bar{X}\bar{Z}}(x,z) = P_{XZ}(x,z)\mathbb{1}\{-\log P_{X|Z}(x|z) > r\},$$

which implies

$$\mathbf{SD}(XZ,\bar{X}\bar{Z}) \leqslant \frac{\epsilon}{2}.$$

Also, we have

$$H_2(\bar{X}|Z) \geqslant H_{\min}(\bar{X}|Z)$$

$$\geqslant r = \underline{H}_s^{\epsilon}(X|Z),$$

that proves

$$\mathbf{SD}(\bar{K}\bar{Z}FS, U\bar{Z}FS) \leqslant \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\underline{H}_s^{\epsilon}(X|Z)}},$$

where $\bar{K} = h_S(\bar{X})$. Since $\mathbf{SD}(XZ,\bar{X}\bar{Z}) \leqslant \epsilon/2$, then $\mathbf{SD}(KZ,\bar{K}\bar{Z}) \leqslant \epsilon/2$ and $\mathbf{SD}(Z,\bar{Z}) \leqslant \epsilon/2$ (See Corollary 2.1.1, and Lemma 2.2). By triangle inequality we have

$$\mathbf{SD}(KZFS,UZFS) \leqslant \mathbf{SD}(KZFS,\bar{K}\bar{Z}FS) + \mathbf{SD}(\bar{K}S\bar{Z}F,US\bar{Z}F) + \mathbf{SD}(U\bar{Z}FS,UZFS)$$

$$= \mathbf{SD}(KZ,\bar{K}\bar{Z}) + \mathbf{SD}(\bar{K}S\bar{Z}F,US\bar{Z}F) + \mathbf{SD}(\bar{Z},Z)$$

Therefore,

$$\mathbf{SD}(KZFS,UZFS) \leqslant \epsilon + \frac{1}{2}\sqrt{|\mathcal{K}||\mathcal{F}|2^{-\underline{H}_s^{\epsilon}(X|Z)}}.$$

Equivalently, the Leftover Hash Lemma, as stated above, implies that h_S is $\sigma-$secure for P_{XZF}, if for an arbitrary $0 < \mu \leqslant \sigma$,

$$\log|\mathcal{K}| \leqslant \underline{H}_s^{\sigma-\mu}(X|Z) - \log|\mathcal{F}| + \log 4\mu^2. \qquad \blacksquare$$

Chapter 4

Secret Key Agreement in Wiretapped Tree-PIN

Abstract. This chapter considers the problem of multiterminal secret key agreement (SKA) in wiretapped source model where terminals have access to samples of correlated random variables from a publicly known joint probability distribution. The adversary has access to a side information variable, that is correlated with terminals' variables. We focus on a special type of terminal variables in this model, known as Tree-PIN, where the relation between variables of the terminals can be represented by a tree. The study of Tree-PIN source model is of practical importance as it can be realized in wireless network environments. We derive the wiretap secret key capacity of Tree-PIN, and give lower and upper bounds on the maximum achievable secret key length in finite-length regime. We then prove an upper bound and a lower bound for the wiretap secret key capacity of a wiretapped PIN and give two conditions for which these bounds are tight. We also extend our main result to two other related models and prove their corresponding capacities. At the end we argue how our analysis suggests that public interaction is required for achieving the multiterminal WSK capacity.

4.1   Introduction

In a multiterminal secret key agreement (SKA) problem, a designated group of users (terminals) collaborate to obtain a shared secret key (SK) such that users outside the group do not have any information about the key. We study the problem of SKA in source model [21], where there is a set of m terminals $\mathcal{M} = \{1, \ldots, m\}$ and the goal is to establish a shared secret key among a subset $\mathcal{A} \subseteq \mathcal{M}$ of terminals. Terminals have access to samples of correlated random variables where random variable X_j is observed by the jth terminal, and $X_{\mathcal{M}} = (X_1, \ldots, X_m)$ denotes the set variables of all terminals. To obtain a shared key, terminals use a public channel to exchange messages that are visible by the eavesdropper, Eve. All terminals, including the helper terminals in $\mathcal{A}^c = \mathcal{M} \backslash \mathcal{A}$, cooperate to establish a shared secret key. Eve, will see and record public messages, denoted by $\mathbf{F}$, and has access to the side information Z that is correlated with $X_{\mathcal{M}}$.

For a key agreement protocol that establishes a key of length ℓ, the *key rate* is defined for the case that the terminals' random variables consist of a vector of n independent and identically distributed (IID) samples of the source distribution $P_{ZX_{\mathcal{M}}}$, and is given by ℓ/n. The *key capacity* of a protocol for a given source distribution is the highest achievable key rate associated with that distribution, and for this general case of variable Z, is referred to as *wiretap secret key (WSK) capacity*. For the special case where $Z = $ constant and there is no wiretapper, the model is called *non-wiretapped* and the key capacity is called *secret key (SK) capacity*. An important special case is when the adversary "wiretaps" and their side information is obtained from a set $\mathcal{D} \subseteq \mathcal{A}^c$ of compromised helper terminals. It is assumed that the compromised terminals of $\mathcal{D}$ make their RV's public, $X_{\mathcal{D}} = (X_j | \ j \in \mathcal{D}) = Z$, and remain cooperative throughout the SKA protocol. The key capacity of such a source model is called *private key (PK) capacity*. A summary of these adversarial models and their corresponding key capacities are given in Table 4.1. Single-letter expressions for SK and PK capacities of multiterminal source model are known [21]. Single-letter characterization of WSK capacity however, remains an open question in general, even for the case of two-party

Table 4.1: Different Types of Key Capacities Based on The Assumption About The Adversary.

Source Model	Eve's Side Information	Key Capacity
Wiretapped	Z not known publicly	WSK
Compromised	$Z = X_{\mathcal{D}}$, and known publicly	PK
Non-wiretapped	$Z = $ constant	SK

SKA (that is when $|\mathcal{A}| = |\mathcal{M}| = 2$) [45, 74, 90]. WSK capacity of a few special cases are known [22, 34, 104]. In this work, we prove the WSK capacity of another special subclass of multiterminal model, referred to as the wiretapped Tree-PIN model with independent leakage. In the following, we first give a brief overview of relevant related works, and then outline our contributions.

4.1.1 Related Works

Capacity results. The SKA problem for two terminals was first considered, independently, in [20] and [19]. The SK capacity was proved to be $I(X_1; X_2)$ [19, 20]. It was also proved that $I(X_1; X_2|Z)$ is an achievable key rate if the terminals know Eve's side information Z. Therefore the conditional mutual information $I(X_1; X_2|Z)$ is an upper bound for the WSK capacity, and it was shown [20] that it is tight if the Markov Relation $X_1 - X_2 - Z$ (or $X_2 - X_1 - Z$) holds. Csiszár and Narayan extended the two-party source model of [19, 20] to the multiterminal model and proved single-letter expressions for SK and PK capacities of multiterminal source models [21]. Similar to the two-party scenario, it was showed that multiterminal PK capacity provides an upper bound on the WSK capacity. The PK (and SK) capacity achieving protocol of [21] has two steps: in the first step, terminals communicate over the public channel to obtain omniscience, that is terminals in $\mathcal{D}^c$ learn $X_{\mathcal{M}}$, and in the second step, terminals in $\mathcal{A}$ extract their copy of the key from the common shared randomness $X_{\mathcal{M}}^n$. While WSK capacity remains unknown in general, the characterization of WSK and also alternative formulations of SK and PK capacities for special cases of multi-

terminal models have been studied, extending the general results of [21]. We briefly review two of these special case models that are related to our work.

The Markov Tree model is a special case of the general multiterminal source model that was introduced and studied in [21, 22]. In a non-wiretapped Markov Tree, the correlation between source variables is given by an undirected tree $G = (\mathcal{M}, \mathcal{E})$ in which each terminal is represented by a node in G, and for any path from terminal i_1 to i_f, denoted by $\text{Path}(i_1 \rightarrow i_f) = (e_{i_1 i_2}, e_{i_2 i_3}, \cdots, e_{i_{f-1} i_f})$, the Markov chain $X_{i_1} - X_{i_2} - X_{i_3} - \cdots - X_{i_{f-1}} - X_{i_f}$ holds. The source model is called wiretapped Markov Tree, if the source variables form a Markov Tree, and the variable associated with each terminal is independently and partially leaked to Eve – i.e., with respect to each X_j there exists a Z_j component available to Eve, where Z_j is a noisy version of X_j. In a wiretapped Markov Tree, corresponding to a path from terminal i_1 to i_f as above, the Markov chain $Z_{i_1} - X_{i_1} - X_{i_2} - X_{i_3} - \cdots - X_{i_{f-1}} - X_{i_f} - Z_{i_f}$ holds. The SK and PK capacities of the Markov Tree source model where derived in [21, Example 7]. The WSK capacity of wiretapped Markov Tree however remains an open problem even for the case of two-party SKA (i.e., when $m = 2$ and $Z_1 - X_1 - X_2 - Z_2$). For the case that the variable associated with only *one* of the leaf terminals is leaked (i.e., $Z_i = \text{constant}, \forall i \neq j$ where $j \in \mathcal{M}$ is a leaf node of G), the WSK capacity of the wiretapped Markov Tree is proved in [22, Theorem 5.1].

A second special case of the multiterminal model is the Pairwise Independent Network (PIN) model [55], inspired by a wireless setting where each pair of terminals can obtain correlated variables from the channel connecting the two. Source variables in PIN are defined by an undirected graph $G = (\mathcal{M}, \mathcal{E})$ with node (vertex) set $\mathcal{M}$ and edge set $\mathcal{E}$, where for an edge $e_{ij} = e_{ji} \in \mathcal{E}$ between i and j ($i \neq j \in \mathcal{M}$), there exists a variable V_{ij} accessible to terminal i, and a second variable V_{ji} (correlated with V_{ij}) accessible to terminal j. The set of all "reciprocal correlated pairs" of variables (i.e., $\{(V_{ij}, V_{ji}) \mid e_{ij} \in \mathcal{E}\}$) are assumed mutually independent[1]. An upper bound on the SK capacity of PIN is given in [55], and a capacity

---

[1]This means that $P_{X_{\mathcal{M}}} = \prod_{e_{ij}} P_{V_{ij} V_{ji}}$.

achieving SKA protocol when $\mathcal{A} = \mathcal{M}$, or when $|\mathcal{A}| = 2$, was proposed in [60]. The PIN model has been well studied [105–108], and has inspired other multiterminal models [71, 72, 109, 110]. An important subclass of the PIN model is defined when the defining graph G is an undirected tree. This model is called Tree-PIN [34]. In this work, we focus on wiretapped Tree-PIN model. We observe that a non-wiretapped Tree-PIN is a non-wiretapped Markov Tree, but the converse is not true. Similarly, we will show that, every wiretapped Tree-PIN with independent leakage is a wiretapped Markov Tree, but the converse does not necessarily hold.

Finite-length performance. The finite-length analysis of coding schemes has found much attention in recent years [31, 49, 50, 53, 59, 65, 86, 111]. Such analysis is important theoretically, and also in practice. While SKA key capacities capture the best asymptotic efficiency of a source model, in practice one needs to obtain bounds on the achievable key length when a finite number (n) of source samples is available. For wiretapped multiterminal source model, a single-shot $(n = 1)$ upper bound on the key length is given in [59]. Finite-length upper and lower bounds for two-party SKA, when $X_1 - X_2 - Z$ holds, have been obtained in [31]. For multiterminal key agreement when Eve has no side information, a finite-length lower bound (of the form $nC_{SK} - \mathcal{O}(\sqrt{n \log n})$) is given in [111].

Communication and computation costs. The key rate measures efficiency of SKA protocol in using the initial correlated randomness, it is also important in practice to measure communication and computation costs.

The *computational efficiency* of an SKA protocol is in terms of the computational complexity of terminals' operations. An SKA protocol is considered computationally efficient if its computational complexity is quasi-linear in n, and is of the form $\mathcal{O}(n \log n)$. The known computationally efficient capacity achieving SKA protocols are given in [33, 76, 78–80]. In most cases the protocols have not been analysed for finite-length performance.

Communication efficiency of an SKA protocol is measured using (i) the public commu-

nication, that for asymptotic case can be measured in terms of asymptotic rate r_{PC}, and for finite-length case, in terms of the total number of bits, of the public communication, and (ii) the total number of rounds N_{PC} of public discussion. We define these measures in Section 4.2. Informally, the asymptotic rate of public communication measures the number of bits of public communication that is used per each observation bit. In a round of public discussion the messages of the terminals only depend on the private samples of the corresponding terminals, and the public messages of the previous rounds. The SKA protocols in [20, 21, 34] are noninteractive: they have one round of public communication, $N_{PC} = 1$. Interactive SKA's have two or more rounds of public communication; e.g., the SKA protocol of [106] has $N_{PC} = 2$ and the two-party SKA protocol of [31] has $N_{PC} \in \mathcal{O}(n)$. For source models, the minimum asymptotic rate of public communication, and the minimum number of public discussion rounds that are required for achieving the key capacity, are important parameters of the system. For SK and PK capacity, the result of [21] implies that the minimum asymptotic rate of public communication for omniscience, is an upper bound for the minimum asymptotic rate of public communication that is required for achieving the corresponding capacity. The minimum asymptotic rate of public communication for SKA for various source models were studied in [112–115].

4.1.2   Our Contributions

In this work, we introduce and study *wiretapped PIN model* and *wiretapped Tree-PIN model*. The wiretapped PIN model with independent leakage is defined as a PIN with an underlying undirected graph $G = (\mathcal{M}, \mathcal{E})$ where legitimate terminals are represented by vertices (nodes) of the graph. An undirected edge between the nodes i and j is represented by $e_{ij} \in \mathcal{E}$. Corresponding to each edge $e_{ij} \in \mathcal{E}$, there exists a variable V_{ij} accessible to terminal i, and a second variable V_{ji} accessible to terminal j. Also, with respect to each edge $e_{ij} \in \mathcal{E}$ Eve has access to a component variable Z_{ij}, and the set of all triplets of variables $\{(V_{ij}, V_{ji}, Z_{ij}) | \ e_{ij} \in \mathcal{E}\}$ are assumed mutually independent, and for each $e_{ij} \in \mathcal{E}$ either $V_{ij} - V_{ji} - Z_{ij}$ or $V_{ji} - V_{ij} - Z_{ij}$

hold[2]. Since G is undirected, we have $Z_{ij} = Z_{ji}$, and denote the adversary's side information by $Z = (Z_{ij}|i < j)$. A wiretapped Tree-PIN is a special case of wiretapped PIN for which the corresponding undirected graph G is a tree. A simple example of such wiretapped Tree-PIN is depicted in Figure 4.1.

Main results. We derive the WSK capacity of wiretapped Tree-PIN with independent leakage as described above, and present an SKA protocol that achieves this capacity. Our SKA protocol has two rounds of public communication ($N_{PC} = 2$) and as shown in Remark 4.3, has a lower asymptotic public communication rate than other SKA protocols (including the protocol in [21]) that have the two steps of achieving omniscience followed by privacy amplification. We note that the adversary in our model is more powerful than the adversary in the wiretapped Markov Tree model of [22, Theorem 5.1], as in the capacity result of [22, Theorem 5.1] Eve only wiretaps one terminal's variable, while in our model of wiretapped Tree-PIN Eve wiretaps all terminals' variables by wiretapping all pairs of correlated variables (V_{ij}, V_{ji}). For the case of two-party SKA, our capacity result also reduces to the result in [20] when $X_1 - X_2 - Z$ (or when $Z - X_1 - X_2$) holds. A simplified version of the wiretapped Tree-PIN model where it is assumed that $V_{ij} = V_{ji}$, was studied and its capacity was derived in our previous work presented in [34].

In Section 4.4, we give a finite-length upper bound and three finite-length lower bounds for the maximum achievable secret key length of a wiretapped Tree-PIN, where each lower bound is due to a different concrete construction of our SKA protocol. We will discuss and compare the three construction approaches in terms of their corresponding lower bounds, their computational complexity, and their communication costs. Our SKA protocol is capacity achieving; however, its achieved key length for n source samples (finite-length analysis) does not match the finite-length upper bound, and the construction of a capacity achieving protocol that achieves the finite-length upper bound of wiretapped Tree-PIN remains open.

---

[2]Only one wiretapped component Z_{ij} is accessible to Eve for each connection $e_{ij} \in \mathcal{E}-$ e.g., $Z_{ij} - V_{ij} - V_{ji} - Z_{ji}$ is not allowed.

$X_1 = (V_{12})$
$X_2 = (V_{21}, V_{23})$
$X_3 = (V_{32}, V_{34})$
$X_4 = (V_{43})$
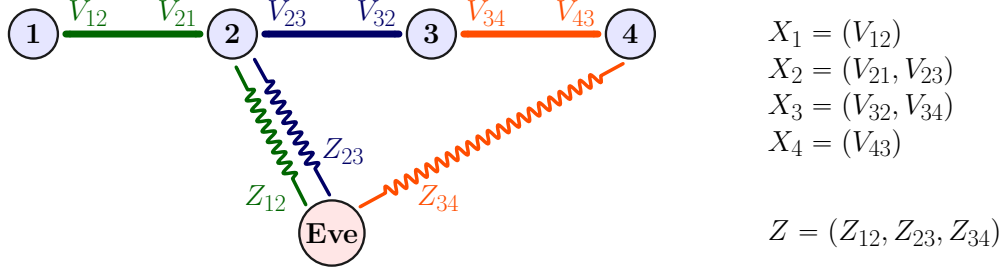
$Z = (Z_{12}, Z_{23}, Z_{34})$

Figure 4.1: An example of wiretapped Tree-PIN with independent leakages defined over $\mathcal{M} = \{1, 2, 3, 4\}$ and $\mathcal{E} = \{e_{12}, e_{23}, e_{34}\}$. The solid lines (edges) show the independent connections between terminals, and the curly lines (with the same color) show the corresponding independent wiretapping RV's of Eve. The RV associated with each terminal $i \in \mathcal{M}$ is of the form $X_i = (V_{ij} | e_{ij} \in \mathcal{E})$. In this example, the following Markov relations hold $V_{12} - V_{21} - Z_{12}$, $V_{32} - V_{23} - Z_{23}$, $V_{34} - V_{43} - Z_{34}$. Eve's RV is a collection of independent wiretapped components, i.e., $Z = (Z_{12}, Z_{23}, Z_{34})$.

Related models. Tree-PIN model has attracted attention over the past years as it can be extended and used to study a number of other related practically important models. In Section 4.5, we extend our main capacity result for wiretapped Tree-PIN to the following more general scenarios. For *wiretapped PIN models* where G can have loops, we show that, a SKA protocol based on Steiner Tree Packing can achieve the WSK capacity when $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$. This is similar to the results obtained in [34, 60], for SKA in non-wiretapped PIN. Next, we note that an important open problem in SKA is finding the WSK capacity of the two-party model when Markov Relation $Z_1 - X_1 - X_2 - Z_2$ holds, where $Z = (Z_1, Z_2)$ is Eve's wiretapped side information [22]. We extend our Tree-PIN to the case where corresponding to each $e_{ij} \in \mathcal{E}$, we have $V_{ij}^a - V_{ji}^a - Z_{ij}^a$ and $Z_{ij}^b - V_{ij}^b - V_{ji}^b$, which implies $Z_{ij}^b - X_i - X_j - Z_{ij}^a$. For $|\mathcal{M}| = 2$, this extended model is an special case of the open problem where Markov relation $Z_1 - X_1 - X_2 - Z_2$ holds. We prove the WSK capacity of this extended model which is (naturally) higher than the WSK capacity of a simple Tree-PIN – as terminals have access to more correlated sources. Lastly, we also prove the key capacity of a PIN model in which not only source variables are wiretapped but also one of the terminals is compromised and is not cooperating. In this case we show that the WSK capacity reduces to the WSK capacity

of the associated model where the compromised terminal and terminals' variables associated with the the compromised terminal are removed (ignored.)

Need for interaction. Csiszár and Narayan proved that SK and PK capacities can be achieved noninteractively [21]. For some special cases also WSK capacity can be achieved noninteractively [20, 34]. Our proposed capacity achieving SKA protocol for wiretapped Tree-PIN is *interactive*. In Section 4.6, we discuss the number of public communication rounds that is required for achieving the WSK capacity. We analyze known models and constructions [20, 31, 93] and study a number of examples that suggest that in general achieving the WSK capacity requires interaction. Proving this result however remains an interesting open question for future research.

4.1.3 Organization

The rest of this chapter is organized as follows. We review security basic notions and definitions in Section 4.2, and present our main result in Section 4.3. Section 4.4 gives finite-length analysis of wiretapped Tree-PIN, and Section 4.5 is on extensions of our main result including for the wiretap secret key capacity of PIN. Section 4.6 discusses the problem of whether interaction is necessary to attain the WSK capacity, and Section 4.7 concludes the chapter.

4.2 Multiterminal Source Model for SKA

In the general multiterminal source model [21], we have a set of m terminals denoted by $\mathcal{M} = [m] = \{1, \ldots, m\}$, and each terminal $j \in [m]$ has access to a random variable X_j. We denote the collection of m correlated random variables $X_1, \ldots, X_m$ by $X_{\mathcal{M}} = (X_1, \ldots, X_m)$. Terminals collaborate by public discussion over a public channel that is reliable and authenticated. A message that is sent by a terminal j is a function of the terminal's observations of X_j, and the previous public messages. Public discussion happens over a finite number

of rounds, denoted by, N_{PC}. We denote by $\mathbf{F}$ the set of all messages sent over the public channel.

Eve has access to the side information Z which is correlated with $X_{\mathcal{M}}$, and has full read access to public messages $\mathbf{F}$. Eve is a passive adversary, which means they will not change, or block public messages communicated messages. The joint distribution $P_{X_{\mathcal{M}}Z}$ is publicly known. We denote the multiterminal source model by $P_{X_{\mathcal{M}}Z}$ or the discrete multiple memoryless source (DMMS) notation $(X_{\mathcal{M}}, Z)$.

Let $\mathcal{A} \subseteq \mathcal{M}$ be the set of terminals who want to establish a shared secret key K. The key need not be concealed from the helper terminals in $\mathcal{A}^c$. The secret key K is secure against Eve if it satisfies the reliability and secrecy conditions.

Definition 4.1. Consider a source model $(X_{\mathcal{M}}, Z)$ with adversary's side information, Z, and $\mathcal{A} \subseteq \mathcal{M}$ denoting the set of terminals that will share a key $K \in \mathcal{K}$. The key is an (ϵ, σ)-Secret Key (in short (ϵ, σ)-SK) for $\mathcal{A}$, if there exists a protocol with public communication $\mathbf{F}$, and output RVs $\{K_j\}_{j \in \mathcal{A}}$ such that

$$\text{(reliability)} \quad \Pr\{K_j = K\} \geqslant 1 - \epsilon, \quad \forall j \in \mathcal{A}, \tag{4.1}$$

$$\text{(secrecy)} \quad \mathbf{SD}\left((K, \mathbf{F}, Z), (U, \mathbf{F}, Z)\right) \leqslant \sigma, \tag{4.2}$$

where $\mathbf{SD}$ denotes the statistical distance and U is the uniform probability distribution over alphabet $\mathcal{K}$. The length of a key K is given by $\log|\mathcal{K}|$.

Definition 4.2. For a source model $(X_{\mathcal{M}}, Z)$ with adversary's side information, Z, and $\mathcal{A} \subseteq \mathcal{M}$ denoting the set of terminals that want to share a secret key, let $S_{\epsilon,\sigma}(X_{\mathcal{A}}|Z)$ denote the maximum length $\log|\mathcal{K}|$ of all the (ϵ, σ)-SKs that can be established for $\mathcal{A} \subseteq \mathcal{M}$.

SKA for IID variables. Consider a source model $(X_{\mathcal{M}}, Z)$ described by $P_{X_{\mathcal{M}}Z}$, where all terminals cooperate for to establish a shared secret key for terminals in $\mathcal{A}$. To increase the key length, terminal $j \in \mathcal{M}$ use a vector, X_j^n, of n independent and identically distributed

$(n-\mathrm{IID})$ samples of X_j. Let Π be an SKA protocol family that, for any n, establishes a secret key $K^{(n)}$ for $\mathcal{A} \subseteq \mathcal{M}$. The public communication of Π, denoted by $\mathbf{F} = \mathbf{F}(\Pi)$, can be interactive and be comprised of $N_{PC}(\Pi) \geqslant 1$ rounds where in each round $t \in [N_{PC}(\Pi)]$ each terminal j sends up to one public message F_{tj}. A message is a function of X_j^n and all public messages of the previous rounds that is denoted by F^{t-1}, and so $F_{tj} = F_{tj}(X_j^n, F^{t-1})$. We denote all messages of round t by $F_t = (F_{t1}, \ldots, F_{tm})$. The public messages of terminals in each round do not depend on other messages of that round, and can be sent in any order. The maximum number of the rounds of public communication, $N_{PC}(\Pi)$, may in general be a function of n. The SKA protocol Π with public communication $\mathbf{F}$ is called *noninteractive* if $N_{PC}(\Pi) = 1$, meaning that in one round each terminal sends up to a single public message, and $\mathbf{F} = (F_1, \ldots, F_m)$, where $F_j = F_j(X_j^n)$.

The *asymptotic public communication rate* of Π is defined by

$$r_{PC}(\Pi) = \limsup_{n \to \infty} \frac{1}{n} \log(\mathrm{supp}(\mathbf{F}(\Pi))),$$

where $\mathbf{F}(\Pi)$ is the public communication of Π. Public communication cost of Π can be quantified by $r_{PC}(\Pi)$ and $N_{PC}(\Pi)$.

Suppose SKA protocol Π establishes an $(\epsilon_n, \sigma_n)-$SK $K^{(n)}$ for a subset $\mathcal{A} \subseteq \mathcal{M}$, and let $\ell_\Pi(n) = \log |\mathcal{K}^{(n)}|$ denote the length of $K^{(n)}$. The key rate of Π for $n-$IID observations is given by $1/n\ell_\Pi(n)$, and $r_K(\Pi) = \liminf_{n \to \infty} 1/n\ell_\Pi(n)$ is called the *asymptotic key rate* of Π. The asymptotic key rate $r_K(\Pi)$ is *achievable* if $\lim_{n \to \infty} \epsilon_n = \lim_{n \to \infty} \sigma_n = 0$. The *key capacity of a source model* is the maximum of all achievable asymptotic key rates of SKA protocols for the model. See Definition 4.3. For an integer $n \in \mathbb{N}$, and $\epsilon, \sigma \in [0, 1)$, define $S_{\epsilon, \sigma}(X_{\mathcal{A}}^n | Z^n)$ to be the maximum length of all (ϵ, σ)-SK protocols for establishing a secret key for $\mathcal{A} \subseteq \mathcal{M}$.

Definition 4.3 (Key Capacity – Definition 17.16 of [90]). Consider multiterminal SKA for a subset $\mathcal{A} \subseteq \mathcal{M}$ in a the source model $(X_{\mathcal{M}}, Z)$ for the joint distribution $P_{X_{\mathcal{M}}Z}$, where Z denotes Eve's side information about $X_{\mathcal{M}}$. A real number $R \geqslant 0$ is an achievable

SK rate if there exists an SKA protocol that for a given n establishes an $(\epsilon_n, \sigma_n)-$SK $K \in \mathcal{K}$ where $\lim_{n\to\infty} \epsilon_n = 0$, $\lim_{n\to\infty} \sigma_n = 0$, and $\liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{K}| = R$. The maximum of all achievable SK rates is called the key capacity of the model.

SK, PK, and WSK Capacities. When $Z = $ constant (i.e., independent of $X_\mathcal{M}$), the capacity is called SK capacity and is denoted by $C_{SK}^\mathcal{A}(P_{X_\mathcal{M}})$. When $Z = X_\mathcal{D} = (X_j \mid j \in \mathcal{D})$ with $\mathcal{D}$ being the set of (known) compromised terminals, the capacity is called PK capacity and is denoted by $C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{X_\mathcal{M}})$. In this case it is assumed that Z is known publicly. In the general case when the side information Z is correlated with $X_\mathcal{M}$ and is not known by the terminals, the key capacity is called WSK capacity and is denoted by $C_{WSK}^\mathcal{A}(P_{X_\mathcal{M}Z})$. An SKA protocol Π is capacity achieving for a source model if $r_K(\Pi)$ is equal to the key capacity of the source.

For a source model $(X_\mathcal{M}, Z)$ with the joint probability distribution $P_{X_\mathcal{M}Z}$, let $R_{SK}(X_\mathcal{M})$, $R_{PK}(X_\mathcal{M}|X_\mathcal{D})$, and $R_{WSK}(X_\mathcal{M}|Z)$ denote the minimum public communication rate to achieve the SK, PK, and WSK capacities, respectively. These quantities give the minimum public communication cost of the SKA, and are often referred to as *communication complexity* of $(X_\mathcal{M}, Z)$ [113–115]. Characterizations of $R_{SK}(X_\mathcal{M})$ for two-party SKA, and for a special case of PIN models, are given in [112] and [115], receptively. An SKA protocol Π that achieves the WSK capacity of a source model $(X_\mathcal{M}, Z)$ implies $R_{WSK}(X_\mathcal{M}|Z) \leqslant r_{PC}(\Pi)$. A similar statement holds for the case of SK and PK.

The single-letter characterization of SK and PK capacities of the general multiterminal source model was derived in [21]. Next Theorem states this result.

Theorem 4.1 (PK Capacity [21]). *In a given source model $X_\mathcal{M}$ for sharing a secret key among terminals in $\mathcal{A} \subsetneq \mathcal{M}$, with compromised terminals $\mathcal{D} \subseteq \mathcal{A}^c$, the PK capacity is*

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{X_\mathcal{M}}) = H(X_\mathcal{M}|X_\mathcal{D}) - R_{CO}(X_\mathcal{A}|X_\mathcal{D}), \tag{4.3}$$

where $R_{CO}(X_\mathcal{A}|X_\mathcal{D}) = \min\limits_{R_{\mathcal{D}^c} \in \mathcal{R}_{CO}} \mathrm{sum}(R_{\mathcal{D}^c})$ *and*

$$\mathcal{R}_{CO} = \{R_{\mathcal{D}^c}|\mathrm{sum}(R_\mathcal{B}) \geqslant H(X_\mathcal{B}|X_{\mathcal{B}^c}), \ \forall \mathcal{B} \subset \mathcal{D}^c, \mathcal{A} \nsubseteq \mathcal{B}\}\,.$$

Remark 4.1. Equation (4.3) also leads to the SK capacity when $\mathcal{D} = \varnothing$. The achievability result is based on a protocol in which first, the compromised terminals (that are cooperative) publicly reveal their observed random variables (as it is the assumption for the PK capacity,) and then the rest of the terminals in $\mathcal{D}^c$ communicate over the public channel to obtain omniscience (i.e., the state that terminals in $\mathcal{D}^c$ learn each other's initial observations). Finally, terminals in $\mathcal{A}$ extract the key from the common shared randomness $X_\mathcal{M}^n$. It was noted that this SKA protocol is noninteractive; meaning that, $N_{PC} = 1$, and $\mathbf{F} = (F_1, \ldots, F_m)$, where $F_j = X_j^n$ for all $j \in \mathcal{D}$ and $F_j = F_j(X_j^n)$ for all $j \in \mathcal{D}^c$. See the achievablity part of the proof of Theorem 2, in Section IV of [21]. The asymptotic public communication rate of this SKA protocol is given by $r_{PC} = R_{CO}(X_\mathcal{A}|X_\mathcal{D})$, which implies that $R_{PK}(X_\mathcal{M}|X_\mathcal{D}) \leqslant R_{CO}(X_\mathcal{M}|X_\mathcal{D})$ (and $R_{SK}(X_\mathcal{M}) \leqslant R_{CO}(X_\mathcal{M})$).

Remark 4.2. We note that SK or PK capacity of multiterminal models depend on the correlation among the variables and in some cases may be zero, in which case it is impossible to establish a group key with information-theoretic security. See Theorem 5 of [21].

Unfortunately, the WSK capacity of the general source model as defined previously, remains an open problem even for the special case of two terminals ($|\mathcal{M}| = 2$) [45]. For the case of two-party SKA, the source model WSK capacity is upper bounded by $I(X_1; X_2|Z)$, which is proved to be a tight bound under the additional assumption that the Markov Chain $X_1 - X_2 - Z$ holds [19, 20]. As was mentioned before, the multiterminal WSK capacity is only known for a few limited special cases [22, 34]. However, PK capacity (see Theorem 4.1) gives a general upper bound to the WSK capacity. We show in the next section that this upper bound is tight for the case wiretapped Tree-PIN.

Lemma 4.2 (Lemma 5.1 of [22]). *For a given wiretapped source model* $(X_\mathcal{M}, Z)$, *let* $C^\mathcal{A}_{WSK}(P_{X_\mathcal{M}Z})$ *denote the WSK capacity of the wiretapped model. Let* $C^{\mathcal{A}|\{m+1\}}_{PK}(P_{X_\mathcal{M}Z})$ *be the PK capacity of an auxiliary model with* $m+1$ *terminals such that* $X_j = X_j$ *for all* $j \leqslant m$, *and* $X_{m+1} = Z$, *where terminal* $m+1$ *is compromised (i.e.,* $\mathcal{D} = \{m+1\}$). *For any given wiretapped model such auxiliary model can be defined. By definition of the PK capacity we have* $C^\mathcal{A}_{WSK}(P_{X_\mathcal{M}Z}) \leqslant C^{\mathcal{A}|\{m+1\}}_{PK}(P_{X_\mathcal{M}Z})$.

4.3 WSK Capacity of Tree-PIN

Here, we first define the wiretapped PIN (Pairwise Independent Network) and wiretapped Tree-PIN models. The non-wiretapped PIN model was first defined in [55] and its SK capacity was later studied in [60]. Let $G = (\mathcal{M}, \mathcal{E})$ be an undirected graph. We denote the edge that connects the nodes i and j by e_{ij}, and *assume* $e_{ij} = e_{ji}$. In a graph $G = (\mathcal{M}, \mathcal{E})$, we denote the neighbours of a node $j \in \mathcal{M}$ by $\Gamma(j) = \{i \mid i \in \mathcal{M}, e_{ij} \in \mathcal{E}\}$.

Definition 4.4 (Wiretapped PIN & Wiretapped Tree-PIN). A set of m terminals form a PIN if there exists a tree $G = (\mathcal{M}, \mathcal{E})$ with $\mathcal{M} = [m]$ such that the RV of any terminal $j \in \mathcal{M}$ can be represented by $X_j = (V_{ji} \mid i \in \Gamma(j))$, where all pairs of RVs in $\{(V_{ij}, V_{ji}) \mid i < j \text{ and } e_{ij} \in \mathcal{E}\}$ are mutually independent. Note that $V_{ij} \neq V_{ji}$. A PIN model is called wiretapped if Eve has access to side information Z which is correlated with all terminals' variables. That is, the correlation between Z and all V_{ij} variables can be in any general form. A wiretapped PIN is called with *independent leakage* if Eve's variable is of the form $Z = (Z_{ij} \mid i < j)$, such that the set of all triplets of variables $\{(V_{ij}, V_{ji}, Z_{ij}) \mid e_{ij} \in \mathcal{E}\}$ are mutually independent and for each $e_{ij} \in \mathcal{E}$ either $V_{ij} - V_{ji} - Z_{ij}$ or $V_{ji} - V_{ij} - Z_{ij}$ hold. A Tree-PIN is a PIN model for which G is an undirected tree. A (Tree-)PIN model is called non-wiretapped if $Z = $ constant.

In our model of wiretapped (Tree-)PIN with independent leakage, Eve has wiretapped side information correlated with every component variable of every terminal, and thus our

wiretap model not only strongly resembles the case of general wiretapped PIN model it is also a special case of the wiretapped Markov Tree model for which the WSK capacity is still unknown [22]. The main results of this chapter are giving the WSK capacity of wiretapped Tree-PIN with independent leakage for any $\mathcal{A}$ (Theorem 4.3), and wiretapped PIN with independent leakage for $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$ (Corollary 4.12.1). These results are more general than previous results on wiretapped multiterminal models. We will compare our results with the aforementioned past results in Section 4.5. In this section, we focus on wiretapped Tree-PIN. The WSK capacity of wiretapped Tree-PIN is given by the following theorem[3].

Theorem 4.3. *WSK capacity of a given wiretapped Tree-PIN $(X_{\mathcal{M}}, Z)$ with independent leakage, defined as in Definition 4.4, for any subset $\mathcal{A} \subseteq \mathcal{M}$ is given by*

$$C^{\mathcal{A}}_{WSK}(P_{X_{\mathcal{M}}Z}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}), \tag{4.4}$$

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ is the subgraph of G with the smallest number of edges connecting all nodes of $\mathcal{A}$.

We emphasis that the WSK capacity of a more general wiretapped PIN model in which $Z = (Z_{ij}|\ i < j)$ and for any i and j the Markov relation $V_{ij} - V_{ji} - Z_{ij}$ does not necessarily hold remains an open problem, even for the case of two-party SKA, $m = 2$.

Proof of Theorem 4.3: The proof is in two parts: (i) the converse, and (ii) the achievability. In the converse part of the proof we prove an upper bound on WSK capacity, that is given by Lemma 4.4.

Lemma 4.4 (The converse). *For a Tree-PIN $(X_{\mathcal{M}}, Z)$ defined as in Definition 4.4, we*

---

[3]The proof for a special case of Theorem 4.3 when $V_{ij} = V_{ji}$ was presented in ISIT 2019 [34]. An extension of this special model to the case of finite linear sources [105] with a linear wiretapper was studied in [104].

have

$$C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z}) \leqslant C_{PK}^{\mathcal{A}|\{m+1\}}(P_{X_{\mathcal{M}}Z}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}),$$

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ is the subtree of G with the least number of edges that connects all nodes of $\mathcal{A}$ and dummy terminal $m+1$ represents the adversary.

In the achievability (direct) part we prove that the above upper bound is indeed achievable. That is given by Lemma 4.5.

Lemma 4.5 (The achievability). *For a wiretapped Tree-PIN $(X_{\mathcal{M}}, Z)$ defined by $G = (\mathcal{M}, \mathcal{E})$, and $P_{ZX_{\mathcal{M}}}$, and for any subset $\mathcal{A} \subseteq \mathcal{M}$, the largest asymptotically achievable key rate of SKA protocol 6 is given by*

$$r_K(\boldsymbol{\Pi_{TP}}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}).$$

The proof of Theorem 4.3 is immediately complete by Lemmas 4.4 and 4.5. ∎

4.3.1 Proof Sketch of the Converse and Achievability

In the following, we give an outline of the proof of the converse, and explain how protocol 6 of Lemma 4.5 achieves the key capacity. The full proofs of Lemmas 4.4 and 4.5 are given in Appendix 4.8.1 and Appendix 4.8.2, respectively.

The Converse. For simplicity, assume $\mathcal{A} = \mathcal{M}$. Also, recall that by Lemma 4.2, we have $C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \leqslant C_{PK}^{\mathcal{M}|\{m+1\}}(P_{X_{\mathcal{M}}Z})$, and due to Theorem 4.1 we know that $C_{PK}^{\mathcal{A}|\{m+1\}}(P_{X_{\mathcal{M}}Z}) = H(X_{\mathcal{M}}|Z) - R_{CO}(X_{\mathcal{M}}|Z)$. Here, $R_{CO}(X_{\mathcal{M}}|Z)$ denotes the solution to the real-valued Linear Programming (LP) problem represented in Figure 4.2.

We prove that

$$R_{CO}(X_{\mathcal{M}}|Z) = H(X_{\mathcal{M}}|Z) - \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}). \tag{4.5}$$

$$
\boxed{
\begin{aligned}
&\text{Minimize:} \quad \sum_{j \in \mathcal{M}} R_j \\
&\text{Subject to:} \sum_{j \in \mathcal{B}} R_j \geqslant H(X_{\mathcal{B}}|X_{\mathcal{B}^c}, Z), \quad \forall \mathcal{B} \subsetneq \mathcal{M}, \\
&\qquad\qquad\quad R_j \in \mathbb{R}^+, \quad \forall j \in \mathcal{M}.
\end{aligned}
}
$$

Figure 4.2: The LP problem of finding $R_{CO}(X_{\mathcal{M}}|Z)$.

First, consider an arbitrary edge $e_{i'j'} \in \mathcal{E}$. By cutting this edge, the set of terminals will be partitioned into two parts $\mathcal{B}$ and $\mathcal{B}^c$ ($\mathcal{B} \cap \mathcal{B}^c = \varnothing$ and $\mathcal{B} \cup \mathcal{B}^c = \mathcal{M}$). Let R_j be the rate of public communication of terminal j. Rewriting the inequalities of LP of Figure 4.2 for these two sets of terminals, and considering the facts that $\{(V_{ij}, V_{ji}, Z_{ij})\}$'s are mutually independent, we get $H(X_{\mathcal{M}}|Z) = \sum_{i,j} H(V_{ij}, V_{ji}|Z_{ij})$ and thus, for any $e_{i'j'} \in \mathcal{E}$ we have

$$
\sum_{j \in \mathcal{B}} R_j \geqslant \sum_{\substack{i \in \mathcal{B} \\ j \in \mathcal{B}}} H(V_{ij}, V_{ji}|Z_{i'j'}) + H(V_{i'j'}|V_{j'i'}, Z_{i'j'}),
$$

$$
\sum_{j \in \mathcal{B}^c} R_j \geqslant \sum_{\substack{i \in \mathcal{B}^c \\ j \in \mathcal{B}^c}} H(V_{ij}, V_{ji}|Z_{ji}) + H(V_{j'i'}|V_{i'j'}, Z_{i'j'}).
$$

By adding these two inequalities, we arrive at

$$
\sum_{j \in \mathcal{M}} R_j \geqslant H(X_{\mathcal{M}}|Z) - \big(H(V_{i'j'}, V_{j'i'}|Z_{i'j'}) - H(V_{i'j'}|V_{j'i'}, Z_{i'j'}) - H(V_{j'i'}|V_{i'j'}, Z_{ji})\big),
$$

$$
= H(X_{\mathcal{M}}|Z) - I(V_{i'j'}; V_{j'i'}|Z_{i'j'}).
$$

This holds for any arbitrary $e_{i'j'} \in \mathcal{E}$, and thus, we have proved that $R_{CO}(X_{\mathcal{M}}|Z) \geqslant H(X_{\mathcal{M}}|Z) - \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$. See Appendix 4.8.1 for the full proof when $\mathcal{A} \neq \mathcal{M}$. This lower bound on R_{CO} implies that $C_{PK}^{\mathcal{A}|\{m+1\}}(P_{X_{\mathcal{M}}Z}) \leqslant \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$, which is essentially sufficient to prove the converse. However, we further prove that this bound is tight and the equality in (4.5) holds. To do so, we show that there exist a heuristic rate assignment for R_1 to R_m such that $\sum_{j \in \mathcal{M}} R_j$ is always equal to the right hand side of (4.5). The proof is in

Protocol 6: SKA for Tree-PIN ($\mathbf{\Pi_{TP}}$)

    Known: Undirected tree $G = (\mathcal{M}, \mathcal{E})$ with $\mathcal{M} = [m]$, and joint distribution $P_{ZX_{\mathcal{M}}}$

    Assumption: Node 2 is the only neighbor of node 1, i.e., $\Gamma(1) = \{2\}$

    Input: Descriptions of $m-1$ two-party SKA protocols $\{\pi_{ij}|\ i < j \text{ and } e_{ij} \in \mathcal{E}\}$

    Input: $n-$IID samples $(X_1^n, X_2^n, \ldots, X_m^n)$

    Final Key Length: ℓ

    Output: Terminals' copies of the final key $(K_1, \ldots, K_m)$, each with length ℓ

    `// Establishing Pairwise Secret Keys`

1 **for** $i \in \mathcal{M}$

2   **for** $j > i$

3   **if** $j \in \Gamma(i)$ **then**   `// Nodes (terminals)` i `and` j `are adjacent`

4   Terminals i and j **do** reconcile on V_{ij}^n using public communication Q_{ij}

5   Terminals i and j **do** extract pairwise keys $S_{ij} = S_{ji} = \pi_{ij}(V_{ij}^n, V_{ji}^n)$

6   Terminals i and j **do** save the first ℓ bits of S_{ij} in $S'_{ij} \leftarrow S_{ij}|_\ell$

    `// XOR Key Distribution`

7 **for** $j \geqslant 2$

8   **if** $|\Gamma(j)| > 1$ **then**   `// Node (terminal)` j `has more than one neighbor`

9   Terminal j **do** find node $j^* \in \Gamma(j)$ s.t. $d(1, j^*) < d(1, i)\ \forall i \in \Gamma(j) \backslash \{j^*\}$, **and**

10   **foreach** $i \in \Gamma(j) \backslash \{j^*\}$, *terminal j* **do** broadcasts $F_{ji} = S'_{jj*} \oplus S'_{ji}$

    `// Local Final Key Calculation`

11 Terminals 1 and 2 set their keys to $K_1 = K_2 = S'_{12}$.

12 **for** $j \geqslant 3$

13   Terminal j **do** find node $j^* \in \Gamma(j)$ s.t. $d(2, j^*) < d(2, i)\ \forall i \in \Gamma(j) \backslash \{j^*\}$, **then**

14   **do** find $\text{Path}(j \to 2)$, the path from node j to node 2, **then**

15   **do** compute $K_j = S'_{jj*} \displaystyle\bigoplus_{\substack{i_a, i_b\ \in \mathcal{M} \\ \text{s.t. } e_{i_a i_b} \in \text{Path}(j \to 2)}} F_{i_a i_b}$

Appendix 4.8.1. This exact formulation of $R_{CO}(X_{\mathcal{M}}|Z)$ will be used later in Remark 4.3 for

arguing the public communication efficiency of SKA protocol 6.

The achievability. We show that the upper bound given in Lemma 4.4 is achievable.

More precisely, we prove that for every n, Protocol 6 generates an $(\epsilon_n, \sigma_n)-$SK K with

length ℓ, such that $\lim_{n\to\infty} \epsilon_n = \lim_{n\to\infty} \sigma_n = 0$, and

$$r_K(\mathbf{\Pi_{TP}}) = \lim_{n\to\infty} \frac{\ell}{n} = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} I(V_{ij}; V_{ji}|Z_{ij}).$$

The protocol works by using the public communication channel in two rounds. First, each pair of connected terminals i and j execute two-party SKA protocols π_{ij} (in parallel) to establish pairwise keys S'_{ij} of length ℓ, where for each e_{ij} the pairwise key length $\ell \approx nI(V_{ij}, V_{ji}|Z_{ij}) - o(n)$ is achievable due to [20, Theorem 1] – see also Theorem 2.11-b. In the second round, terminals use the public channel to reconcile on one of the pairwise keys, namely S'_{12}. In this step, non-leaf nodes (terminals) send enough messages that enables all terminals to calculated $K = S'_{12}$ while keeping the leakage of information to Eve to a minimum amount.

A complete description of this SKA protocol is given in Protocol 6. In Protocol 6, without loss of generality, we assume that terminal 2 is the only terminal connected to terminal 1; i.e., $\Gamma(1) = \{2\}$. In line 15 of Protocol 6 $\text{Path}(i_1 \to i_f) = (e_{i_1 i_2}, e_{i_2 i_3}, \cdots, e_{i_{f-1} i_f})$ denotes the path from terminal i_1 to i_f. Since G is an undirected tree, between each terminal $i \in \mathcal{M}$ and $j \in \mathcal{M}$ there is always a unique path. We show in the proof of Lemma 4.5 that if pairwise keys are $(\epsilon, \sigma)-$SKs established by executing two-party SKA protocols π_{ij}, then the final key of Protocol 6 is an $(|\mathcal{E}|\epsilon, 2|\mathcal{E}|\sigma)-$SK. The full proof of achievability is in Appendix 4.8.2.

Example 4.1. In the following, we revisit the example of Figure 4.1, and illustrate how protocol 6 works. This wiretapped Tree-PIN with $\mathcal{M} = \{1, 2, 3, 4\}$ is a simple path from terminal 1 to terminal 4.

Protocol 6 works as follows. First, each pair of connected terminals establish pairwise secret keys S_{ij} by employing two-party SKA protocols π_{ij}. Then, let ℓ be the length of the smallest pairwise key. All parties then keep only the first ℓ bits of their pairwise keys. Let S'_{ij} denote the first ℓ bits of S_{ij}. Note that in this example terminal 2 has two pairwise keys $\{S'_{12}, S'_{23}\}$ and terminal 3 also has two pairwise keys $\{S'_{23}, S'_{34}\}$. In the next phase of the

protocol, terminal 2 broadcasts $F_{23} = S'_{12} \oplus S'_{23}$ and terminal 3 broadcasts $F_{34} = S'_{23} \oplus S'_{34}$.

In the last phase, each terminal j computes the key K_j according to the following

$$K_1 = S'_{12},$$

$$K_2 = S'_{12},$$

$$K_3 = S'_{23} \oplus F_{23},$$

$$K_4 = S'_{34} \oplus F_{34} \oplus F_{23}.$$

One can easily see that above equations imply that we have $K_1 = K_2 = K_3 = K_4 = S'_{12}$.

4.3.2   Public Communication Cost of Protocol 6

The Protocol 6 is the only known protocol that achieves the WSK capacity of Tree-PIN; however, when Z is known, it can be compared with other protocols that achieve the PK capacity. This protocol is interactive with two rounds of public communication but does not require omniscience. We show that the public communicate cost of Protocol 6, that is the asymptotic rate of its public communication, is no larger than other protocols that require omniscience for achieving the PK capacity.

Remark 4.3. Let $R_{SK}(X_{\mathcal{M}})$ denote the minimum public communication rate required for achieving $C_{SK}^{\mathcal{M}}(P_{X_{\mathcal{M}}})$. That is $R_{SK}(X_{\mathcal{M}}) = \min\{r_{PC}(\Pi)|\ \Pi$ achieves $C_{SK}^{\mathcal{M}}(P_{X_{\mathcal{M}}})\}$. It was proved in [115] that for PIN model with $V_{ij} = V_{ji}$, we have $R_{SK}(X_{\mathcal{M}}) = (m-2)C_{SK}^{\mathcal{M}}(P_{X_{\mathcal{M}}})$. Similarly, define $R_{WSK}(X_{\mathcal{M}}|Z) = \min\{r_{PC}(\Pi)|\ \Pi$ achieves $C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z})\}$. We show that for any wiretapped Tree-PIN, when $V_{ij} \neq V_{ji}$, we have

$$R_{WSK}(X_{\mathcal{M}}|Z) \leqslant \Big(\sum_{i,j} H(V_{ij}|V_{ji})\Big) + (m-2)C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \leqslant R_{CO}(X_{\mathcal{M}}|Z),$$

where $R_{CO}(X_{\mathcal{M}}|Z)$ is defined in Theorem 4.1. It is not known whether the left bound is tight. When Z is known, both Protocol 6 and protocol of [21] achieve the PK capacity of

Tree-PIN. Protocol 6 does not require achieving omniscience while protocol of [21] does. The above inequality shows that Protocol 6 uses less public communication than the protocol of [21] (Also see [21, Example 7]).

Proof of Remark 4.3: First we prove the first bound by noting the fact that $R_{WSK}(X_{\mathcal{M}}|Z) \leqslant r_{PC}(\mathbf{\Pi_{TP}})$ as Protocol 6 ($\mathbf{\Pi_{TP}}$) achieves the WSK capacity. We now calculate $r_{PC}(\mathbf{\Pi_{TP}})$. Protocol 6 has two rounds of public communication. In the first round terminals agree on their pairwise keys. For each e_{ij} either $V_{ij} - V_{ji} - Z_{ij}$ or $V_{ji} - V_{ij} - Z_{ij}$ holds. With an abuse of notation, assume that $V_{ij} - V_{ji} - Z_{ij}$ for all e_{ij}. Then public communication rate of the first round for each e_{ij} is given by $H(V_{ij}|V_{ji})$ [20]. Since in the first round, pairwise keys are generated in parallel and independently, the total amount of public communication rate of this round is given by $\sum_{i,j} H(V_{ij}|V_{ji})$. In the second round, any terminal $j \in \mathcal{M}$ finds its unique[4] neighbour j^* that is closest to the node 1 and broadcasts $|\Gamma(j)| - 1$ encoded messages $\{F_{ji}| \ \forall i \in \Gamma(j)\backslash\{j^*\}\}$, where each message has the same length ℓ as the final key K. Thus, the public communication rate of the protocol 6 is

$$
\begin{aligned}
r_{PC}(\mathbf{\Pi_{TP}}) &= \sum_{i,j} H(V_{ij}|V_{ji}) + \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{m} \ell \times (|\Gamma(j)| - 1) \\
&= \sum_{i,j} H(V_{ij}|V_{ji}) + \lim_{n\to\infty} (\ell/n) \times \left(\sum_{j=1}^{m} |\Gamma(j)| - m \right) \\
&= \sum_{i,j} H(V_{ij}|V_{ji}) + \lim_{n\to\infty} (\ell/n) (2|\mathcal{E}| - m) \\
&= \sum_{i,j} H(V_{ij}|V_{ji}) + \lim_{n\to\infty} (m - 2)\ell/n,
\end{aligned}
$$

where we used the facts that for a graph $G = (\mathcal{M}, \mathcal{E})$, we have $\sum_{j\in\mathcal{M}} |\Gamma(j)| = 2|\mathcal{E}|$, and for an undirected tree with m vertexes we have $|\mathcal{E}| = m - 1$. By $\ell \approx n \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}) - o(n)$, and the fact that Protocol 6 achieves the WSK capacity of a Tree-PIN, namely $\min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$, proves the first (left) inequality for any given Tree-PIN $G = (\mathcal{M}, \mathcal{E})$.

---

[4]Exists because of tree structure of the variables.

Next, we prove the second (right) inequality by showing that $r_{PC}(\mathbf{\Pi_{TP}}) \leqslant R_{CO}(X_{\mathcal{M}}|Z)$.

$$
\begin{aligned}
r_{PC}(\mathbf{\Pi_{TP}}) &= \sum_{i,j} H(V_{ij}|V_{ji}) + (m-2)C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \\
&\leqslant \sum_{i,j} H(V_{ij}|V_{ji}) + (m-1)C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) - C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \\
&\leqslant \sum_{i,j} H(V_{ij}|V_{ji}) + \sum_{i,j} H(V_{ij}|Z_{ij}) - H(V_{ij}|V_{ji}) - C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \\
&= \sum_{i,j} H(V_{ij}|Z_{ij}) - C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \\
&\leqslant H(X_{\mathcal{M}}|Z) - C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) \\
&= R_{CO}(X_{\mathcal{M}}|Z)
\end{aligned}
$$

where the last equality is due to (4.22).   ∎

4.4   Finite-length Bounds for Wiretapped Tree-PIN

Finite-length analysis of information theoretic tasks such as SKA is important in practice, as in real-life deployment of SKA protocols the number of samples, n, accessible to each terminal is finite. In this case, better estimations and bounds on the maximum achievable key length (i.e., $S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z^n)$) are desired (see Definition 4.2). In this section, we give a finite-length upper bound, and three finite-length lower bounds for the maximum achievable key length in a wiretapped Tree-PIN.

4.4.1   The Finite-length Upper Bound

Theorem 4.6. *For any given wiretapped Tree-PIN* $(X_{\mathcal{M}}, Z)$, *described by* $P_{ZX_{\mathcal{M}}}$, *and for every* $n \in \mathbb{N}$, *every* $\epsilon, \sigma > 0$, *with* $\epsilon + \sigma < 1$, *and any subset* $\mathcal{A} \subseteq \mathcal{M}$, *we have that* $S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z^n)$ *is upper bounded by*

$$
\min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \left\{ nR_{ij} - \sqrt{n\Delta_{ij}}Q^{-1}(\epsilon+\sigma) \right\} + \frac{3}{2}\log n + \mathcal{O}(1), \tag{4.6}
$$

where $R_{ij} = I(V_{ij}; V_{ji}|Z_{ij})$.

For the proof of Theorem 4.6 we use the Hypothesis testing upper bound of Tyagi and Watanabe [116] which is a general single-shot bound for any wiretapped multiterminal source model. Hayashi et al. used the upper bound of [116] to prove a finite-length upper bound for the case of two-party SKA. To our knowledge, Theorem 4.6 is the first multiterminal finite-length upper bound based on the Hypothesis testing upper bound.

To prove Theorem 4.6, we first recall the notion hypothesis testing and a couple of lemmas.

The binary hypothesis testing problem is defined as follows. For a random variable X, there are two possible distributions P_X and Q_X. Using a test algorithm T we shall decide between P_X or Q_X. Let the null hypothesis be $H_0 = P_X$. If we reject the null hypothesis P_X when the actual distribution is P_X then type I error is occurred, and if we accept the null hypothesis when the actual distribution is Q_X then type II error is occurred. Let $\beta_\eta(P_X, Q_X)$ denote the infimum of type II error probability given that type I error probability is less than η. That is,

$$\beta_\eta(P_X, Q_X) = \inf_{T: E_1(T) \leqslant \eta} E_2(T),$$

where $E_1(T) = \sum_{x \in \mathcal{X}} P_X(x) \Pr\{\text{Rej } H_0|x\}$, and $E_2(T) = \sum_{x \in \mathcal{X}} Q_X(x) \Pr\{\text{Acc } H_0|x\}$, are respectively the type I and type II errors of a given hypothesis testing algorithm T.

Lemma 4.7 (Hypothesis testing upper bound [116]). *Given an arbitrary multiterminal source model $(X_{\mathcal{M}}, Z)$, and any given partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_l\}$ of $\mathcal{M}$, for every $\epsilon, \sigma > 0$, with $\epsilon + \sigma < 1$, and every $0 < \eta < 1 - \epsilon - \sigma$, we have*

$$S_{\epsilon, \sigma}(X_{\mathcal{M}}|Z) \leqslant \frac{1}{|\mathcal{P}| - 1} \left[-\log \beta_{\epsilon + \sigma + \eta} \left(P_{X_{\mathcal{M}}Z}, Q^{\mathcal{P}}_{X_{\mathcal{M}}Z} \right) + |\mathcal{P}| \log \frac{1}{\eta} \right],$$

where $Q^{\mathcal{P}}_{X_{\mathcal{M}}Z}$ is any probability distribution for which $Q^{\mathcal{P}}_{X_{\mathcal{M}}|Z} = \Pi^l_{j=1} Q_{X_{\mathcal{P}_j}|Z}$ holds.

Lemma 4.8 (Also see Lemma 4.1.2 of [24]). *Consider a hypothesis testing problem where P_X and Q_X are respectively the null and alternative hypotheses. For any $\lambda > 0$, we*

have

$$-\log \beta_\epsilon(P_X, Q_X) \leqslant \lambda - \log \left(P_X \left(\left\{ x: \ \log \frac{P_X(x)}{Q_X(x)} \leqslant \lambda \right\} \right) - \epsilon \right).$$

Proof: Let

$$\mathcal{C} = \left\{ x: \ \log \frac{P_X(x)}{Q_X(x)} \geqslant \lambda \right\}.$$

Suppose that the hypothesis testing algorithm T is such that accepts the null hypothesis P_X if the observed value x belongs to $\mathcal{C}$. Also, let ϵ denote the type I error of test T. That is,

$$\epsilon = E_1(\mathrm{T}) = P_X \left(\left\{ x: \ \log \frac{P_X(x)}{Q_X(x)} < \lambda \right\} \right) = \sum_{x \in \mathcal{X}} P_X(x) \mathbb{1} \left(x \notin \mathcal{C} \right).$$

Due to the Neyman-Pearson lemma, T gives the least type II error of all tests with type I error of at most ϵ. To simplify the proof, let

$$\mathcal{S} = \left\{ x: \ \log \frac{P_X(x)}{Q_X(x)} \leqslant \lambda \right\}.$$

Using the Neyman-Pearson lemma we have,

$$
\begin{aligned}
P_X \left(\left\{ x: \ \log \frac{P_X(x)}{Q_X(x)} \leqslant \lambda \right\} \right) &= \sum_{x \in \mathcal{X}} P_X(x) \mathbb{1} \left(x \in \mathcal{S} \cap \mathcal{C}^c \right) + \sum_{x \in \mathcal{X}} P_X(x) \mathbb{1} \left(x \in \mathcal{S} \cap \mathcal{C} \right) \\
&\overset{(a)}{\leqslant} \sum_{x \in \mathcal{X}} P_X(x) \mathbb{1} \left(x \notin \mathcal{C} \right) + \sum_{x \in \mathcal{X}} 2^\lambda Q_X(x) \mathbb{1} \left(x \in \mathcal{S} \cap \mathcal{C} \right) \\
&\leqslant \sum_{x \in \mathcal{X}} P_X(x) \mathbb{1} \left(x \notin \mathcal{C} \right) + \sum_{x \in \mathcal{X}} 2^\lambda Q_X(x) \mathbb{1} \left(x \in \mathcal{C} \right) \\
&\overset{(b)}{=} \epsilon + 2^\lambda \beta_\epsilon(P_X, Q_X),
\end{aligned}
$$

where in (a) we use that $P_X(x) \leqslant 2^\lambda Q_X(x) \ \forall x \in \mathcal{S}$, and in (b) we use Neyman-Pearson lemma. The proof is complete by taking logarithm from both sides of the inequality. $\blacksquare$

Theorem (Berry-Esseen, see Theorem 2.4). *Let W^n be an $n-$IID variable, and $-\infty <*

$\alpha < \infty$, then

$$\left| \Pr \left\{ \sum_{j=1}^{n} W_j \leqslant n\mu - \alpha\sqrt{\Delta n} \right\} - Q(\alpha) \right| \leqslant \frac{3\rho}{\Delta^{3/2}\sqrt{n}},$$

where $\mu = \mathbb{E}\{W\}, \Delta = \text{Var}\{W\}, \rho = \mathbb{E}\{|W - \mu|^3\}$, and $Q(\cdot)$ is the tail probability of the standard Gaussian distribution.

We now prove the upper bound of Theorem 4.6.

Proof: Denote the set of all terminals in G_A by $\mathcal{B} = \mathcal{M}_A \subseteq \mathcal{M}$. For SKA in the Tree-PIN G_A, lemma 4.7 implies that for an arbitrary partition $\mathcal{P}$ of $\mathcal{B}$, we have

$$S_{\epsilon,\sigma}(X_A^n|Z) = S_{\epsilon,\sigma}(X_{\mathcal{B}}^n|Z) \leqslant \frac{1}{|\mathcal{P}| - 1}\left[-\log\beta_{\epsilon+\sigma+\eta}\left(P_{X_{\mathcal{B}}^n Z^n}, Q_{X_{\mathcal{B}}^n Z^n}^{\mathcal{P}} \right) + |\mathcal{P}|\log\frac{1}{\eta} \right]. \quad (4.7)$$

Fix an edge $e_{i'j'} \in \mathcal{E}_A$ of G_A that connects nodes (terminals) i' and j'. Cutting this edge induces a partition $\mathcal{P}_{i'j'} = \{\mathcal{P}_1, \mathcal{P}_2\}$, such that $i' \in \mathcal{P}_1$ and $j' \in \mathcal{P}_2$. By applying (4.7) and lemma 4.8, with $\mathcal{P} = \mathcal{P}_{i'j'}$, $P_{X_{\mathcal{B}}^n Z^n} = \prod_{e_{ij}} P_{V_{ij}^n V_{ji}^n Z_{ij}^n}$, $Q_{X_{\mathcal{B}}^n Z^n}^{\mathcal{P}} = P_{V_{i'j'}^n|Z_{i'j'}^n} P_{V_{j'i'}^n Z_{i'j'}^n}\prod_{e_{ij}\neq e_{i'j'}} P_{V_{ij}^n V_{ji}^n Z_{ij}^n}$, and $\eta = \frac{1}{\sqrt{n}}$, we get

$$S_{\epsilon,\sigma}(X_A^n|Z) \leqslant \lambda - \log\left(\Pr\left\{ \log\frac{P_{X_{\mathcal{B}}^n Z^n}}{Q_{X_{\mathcal{B}}^n Z^n}^{\mathcal{P}_{i'j'}}} \leqslant \lambda \right\} - \epsilon - \sigma - \frac{1}{\sqrt{n}} \right) + \log n. \quad (4.8)$$

Let

$$\theta_n = \frac{2}{\sqrt{n}} + \frac{3\rho_{i'j'}}{\Delta_{i'j'}^{3/2}\sqrt{n}},$$

where

$$\Delta_{ij} = \text{Var}\left\{ \log\frac{P_{V_{ij}V_{ji}|Z}(V_{ij}, V_{ji}|Z)}{P_{V_{ij}|Z}(V_{ij}|Z)P_{V_{ji}|Z}(V_{ji}|Z)} \right\},$$

and

$$\rho_{ij} = \mathbb{E}\left\{ \left| \log\frac{P_{V_{ij}V_{ji}|Z}(V_{ij}, V_{ji}|Z)}{P_{V_{ij}|Z}(V_{ij}|Z)P_{V_{ji}|Z}(V_{ji}|Z)} - I(V_{i'j'}; V_{j'i'}|Z) \right|^3 \right\}.$$

By choosing

$$\lambda = nI(V_{i'j'}; V_{j'i'}|Z) - \sqrt{n\Delta_{i'j'}}Q^{-1}(\epsilon + \sigma + \theta_n),$$

125

and by the Berry-Esseen theorem we get

$$\Pr\left\{\log\frac{P_{X_{\mathcal{B}}^n Z^n}}{Q_{X_{\mathcal{B}}^n Z^n}^{\mathcal{P}_{i'j'}}} \leqslant \lambda\right\} = \Pr\left\{\log\frac{P_{V_{i'j'}^n, V_{j'i'}^n, Z_{i'j'}^n}}{P_{V_{i'j'}^n|Z_{i'j'}^n} P_{V_{j'i'}^n Z_{i'j'}^n}} \leqslant \lambda\right\}$$

$$= \Pr\left\{\log\frac{P_{V_{i'j'}^n, V_{j'i'}^n|Z^n}}{P_{V_{i'j'}^n|Z^n} P_{V_{j'i'}^n|Z^n}} \leqslant \lambda\right\}$$

$$\geqslant \epsilon + \sigma + \frac{2}{\sqrt{n}}.$$

Note that $\mathbb{E}\left\{\log\frac{P_{V_{ij} V_{ji}|Z}(V_{ij}, V_{ji}|Z)}{P_{V_{ij}|Z}(V_{ij}|Z) P_{V_{ji}|Z}(V_{ji}|Z)}\right\} = I(V_{ij}; V_{ji}|Z)$. Applying the above inequality in (4.8) gives

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z) \leqslant nI(V_{i'j'}; V_{j'i'}|Z) - \sqrt{n\Delta_{i'j'}}Q^{-1}(\epsilon + \sigma + \theta_n) - \log\left(\frac{1}{\sqrt{n}}\right) + \log n.$$

By using Taylor approximation of $Q(\cdot)$ to remove θ_n we get

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z) \leqslant nI(V_{i'j'}; V_{j'i'}|Z) - \sqrt{n\Delta_{i'j'}}Q^{-1}(\epsilon + \sigma) + \frac{3}{2}\log n + \mathcal{O}(1),$$

that holds for any edge $e_{i'j'}$ of $G_{\mathcal{A}}$. The proof is complete by minimizing over all $e_{i'j'}$'s. ∎

4.4.2 Finite-length Lower Bounds

The achievability (lower) bounds are based on variations of the SKA protocol that achieves the WSK capacity of wiretapped Tree-PIN given in Theorem 4.3. This protocol has two main steps. In the first step, each pair of connected terminals i and j (i.e., $e_{ij} \in \mathcal{E}$) preform a two-party SKA protocol to obtain a pairwise secret key. For this task, terminals can use, for example, the two-party SKA protocols 4 or 5 – see also [19, 20, 31–33]. In the second step, terminals use their pairwise keys and public communication to agree on the final shared secret key. See the details of this SKA protocol in Appendix 4.8.2.

For the case of two-party SKA, Hayashi et al. [31] proved that for a given source model

126

(V_{ij}, V_{ji}, Z_{ij}), if $V_{ij} - V_{ji} - Z_{ij}$, and for every $n \in \mathbb{N}$ and $\epsilon, \sigma > 0$, with $\epsilon + \sigma < 1$, we have

$$S_{\epsilon,\sigma}(V_{ij}^n, V_{ji}^n | Z_{ij}^n) = nR_{ij} - \sqrt{n\Delta_{ij}}Q^{-1}(\epsilon + \sigma) \pm \mathcal{O}(\log n),$$

where $S_{\epsilon,\sigma}(\cdot)$ denotes the maximum achievable key length,

$$\Delta_{ij} = \mathrm{Var}\left\{\log \frac{P_{V_{ij}V_{ji}|Z_{ij}}(V_{ij}, V_{ji}|Z_{ij})}{P_{V_{ij}|Z_{ij}}(V_{ij}|Z_{ij})P_{V_{ji}|Z_{ij}}(V_{ji}|Z_{ij})}\right\},$$

$R_{ij} = I(V_{ij}; V_{ji}|Z_{ij})$ is the two-party WSK capacity of (V_{ij}, V_{ji}, Z_{ij}), and $Q(\cdot)$ is the tail probability of the standard Gaussian distribution. This second-order approximation of the key length is achievable by the interactive protocol of [31]. Sharifian et al. [32] gave also two finite-length approximations corresponding to a one-way two-party SKA protocol. See Chapter 3, Section 3.4, Protocol 4. One-way SKA protocols are more efficient in terms of the public communication than the interactive construction of [31], while in finite-length regime, the SKA protocol of [31] is closer to the two-party capacity (R_{ij}) than the SKA protocol of [32]. However, by a numerical example in Section 4.2 we illustrate that the lower bound that is based on [32] can be very close to the lower bound which is based on [31].

By using the SKA protocols of [31] and [32] in the first step of our SKA protocol for obtaining pairwise keys, we prove the following lower bounds for wiretapped Tree-PIN.

Proposition 4.9 (Lower bounds). *For any given wiretapped Tree-PIN, described by $P_{ZX_{\mathcal{M}}}$, and for every $n \in \mathbb{N}$, every $\epsilon, \sigma > 0$, with $\epsilon + \sigma < 1$, and any subset $\mathcal{A} \subseteq \mathcal{M}$, we have*

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n | Z^n) \geqslant F_1(X_{\mathcal{A}}^n | Z^n) - \frac{11}{2}\log n + \mathcal{O}(1) \tag{4.9}$$

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n | Z^n) \geqslant F_2(X_{\mathcal{A}}^n | Z^n) - \log n + \mathcal{O}(1) \tag{4.10}$$

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n | Z^n) \geqslant F_3(X_{\mathcal{A}}^n | Z^n) - \log n + \mathcal{O}(1) \tag{4.11}$$

where

$$F_1(X_{\mathcal{A}}^n|Z^n) = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \left\{ nR_{ij} - \sqrt{n\Delta_{ij}}Q^{-1}(\frac{2\epsilon+\sigma}{2|\mathcal{E}_{\mathcal{A}}|}) \right\},$$

$$F_2(X_{\mathcal{A}}^n|Z^n) = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \left\{ nR_{ij} - Q^{-1}(\frac{\epsilon}{|\mathcal{E}_{\mathcal{A}}|})\sqrt{n\Delta'_{ij}} - Q^{-1}(\frac{\sigma}{2|\mathcal{E}_{\mathcal{A}}|})\sqrt{n\Delta''_{ij}} \right\},$$

$$F_3(X_{\mathcal{A}}^n|Z^n) = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \{nR_{ij}\} - \sqrt{2n}\log(|\mathcal{X}|+3)(\sqrt{\log\frac{|\mathcal{E}_{\mathcal{A}}|}{\epsilon}} + \sqrt{\log\frac{2|\mathcal{E}_{\mathcal{A}}|}{\sigma}}),$$

with $R_{ij} = I(V_{ij}; V_{ji}|Z_{ij})$, $\Delta'_{ij} = \text{Var}\left\{-\log P_{V_{ij}|V_{ji}}\right\}$, $\Delta''_{ij} = \text{Var}\left\{-\log P_{V_{ij}|Z_{ji}}\right\}$, *and* $|\mathcal{E}_{\mathcal{A}}|$ *is the number of edges in the sub-tree* $G_{\mathcal{A}}$.

For Theorem 4.9 we note that according to the proof of Theorem 4.3, obtaining pairwise (ϵ, σ)−SKs leads to a finial $(|\mathcal{E}_{\mathcal{A}}|\epsilon, 2|\mathcal{E}_{\mathcal{A}}|\sigma)$−SK. Thus, for all of the above achievability (lower) bounds, parties first establish pairwise $(\frac{\epsilon}{|\mathcal{E}_{\mathcal{A}}|}, \frac{\sigma}{2|\mathcal{E}_{\mathcal{A}}|})$ secret keys, and then use Protocol 6 to agree on the final key. None of the bounds require omniscience. Lower bound of (4.9) is based on Protocol 6 which uses the two-party protocol of [31] for generating pairwise keys, and lower bounds in (4.10) and (4.11) are based on Protocol 6 when the one-way two-party protocol of [32] (Protocol 4) is used for pairwise key generation. Lower bounds in (4.9) and (4.10) assume that samples are IID and lower bound of (4.11) only assumes that samples are independent (and not necessarily IID.) The full proof of Theorem 4.9 is given in Appendix 4.8.3.

Note that the second-order terms (in $\mathcal{O}(\sqrt{n})$) of the upper and lower bounds do not match. Finding tighter bounds with matching second-order terms is an interesting open problem.

Example 4.2. The following numerical example compares the finite-length bounds given in (4.6) and (4.9)-(4.11). Consider a source model with $m = 3$ terminals, $\mathcal{M} = \{1, 2, 3\}$, and $\mathcal{A} = \mathcal{M}$. Let $X_1 = (V_{12}, V_{13})$ such that V_{1j}'s are binary uniform variables. Also for $p, q \in (0, 1)$ and for $j \in \{2, 3\}$, let $X_j = V_{j1} = \text{BSC}_p(V_{1j})$ and $Z_{1j} = \text{BSC}_q(V_{j1})$. Here, $\text{BSC}_p(\cdot)$ denotes
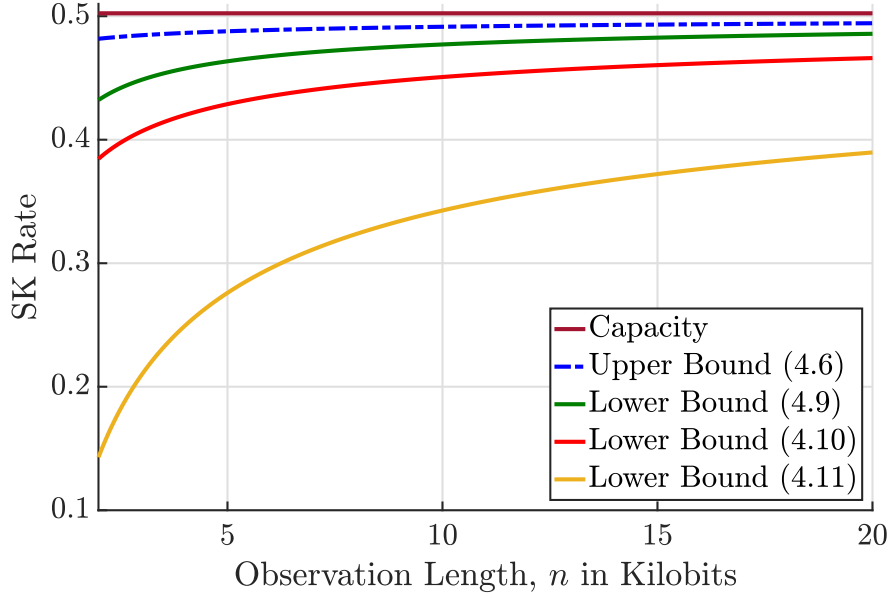
Figure 4.3: Comparing finite-length bounds of the example in Section 4.2. Here, $m = 3$, $\epsilon = \sigma = 0.05$, and the WSK capacity is 0.502. Lower bound of (4.9) is the tightest lower bound and is by Protocol 6 if the two-party interactive SKA of [31] is used for pairwise key generation. Lower bounds in (4.10) and (4.11) are based on Protocol 6 if the two-party one-way SKA protocol of [32] (Protocol 4) is used for pairwise key generation.

a binary symmetric channel with crossover probability of p. For this example, the WSK capacity is $C_{WSK} = h_2(p * q) - h_2(p)$, where $p * q = p(1 - q) + (1 - p)q$, and h_2 is the binary entropy given by $h_2(p) = -p \log p - (1 - p) \log(1 - p)$. Consider, $p = 0.0093$, $q = 0.13$, and $\epsilon = \sigma = 0.05$. Then, $C_{WSK} = 0.502$, and the finite-length approximations of (4.6) and (4.9)-(4.11) calculated for this example are depicted in Figure 4.3 for $n \in [2000, 20000]$. The bounds are converted to rate (both sides are divided by n) to show the gap to the WSK capacity. Note that (4.9) is the tightest lower bound. Though, we also observe that (4.10) is very close to (4.9).

4.4.3   A Lower Bound for a Special Case

In this section, we consider the wiretapped Tree-PIN with $V_{ij} = V_{ji}$ that is studied in [34]. For this case, it was proved that the WSK capacity is $C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z}) = \min\limits_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} H(V_{ij}|Z_{ij})$ [34]. We use the lower bound in [86, Theorem 1], and give the following finite-length lower bound for $S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z^n)$.

Proposition 4.10. *For wiretapped Tree-PIN $(X_{\mathcal{M}}, Z)$ described by $P_{ZX_{\mathcal{M}}}$, with $V_{ij} = V_{ji}$ and for every $n \in \mathbb{N}$, every $\epsilon, \sigma > 0$, with $\epsilon + \sigma < 1$, and any subset $\mathcal{A} \subseteq \mathcal{M}$, we have*

$$S_{\epsilon,\sigma}(X_{\mathcal{A}}^n|Z^n) \geqslant F_4(X_{\mathcal{A}}^n|Z^n) - \frac{1}{2}\log n + \mathcal{O}(1), \qquad (4.12)$$

where

$$F_4(X_{\mathcal{A}}^n|Z^n) = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \left\{ nR_{ij} - \sqrt{n\Delta_{ij}''}Q^{-1}(\frac{\sigma}{2|\mathcal{E}_{\mathcal{A}}|}) \right\},$$

$R_{ij} = H(V_{ij}|Z_{ij})$, $\Delta_{ij}'' = \mathrm{Var}\left\{ -\log P_{V_{ij}|Z_{ji}} \right\}$, and $|\mathcal{E}_{\mathcal{A}}|$ is the number of edges of $G_{\mathcal{A}}$.

Note that (4.12) does not depend on ϵ as the reconciliation phase is not required for obtaining pairwise keys, and for $\epsilon = 0$, the lower bounds in (4.9) and (4.12) are equal up to their second-order term. The proof is in Appendix 4.8.3.

4.5   Extended Models

In this section, we extend our capacity result of wiretapped Tree-PIN. While doing so, we compare our results with some important related previous works. We give an upper and a lower bound for the WSK capacity of wiretapped PIN, which is a generalization of the bounds given in [60] for (non-wiretapped) PIN. More importantly, these bounds lead to capacity results for the case wiretapped PIN when $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$. We then, review the notion of *wiretapped Markov Trees* which was introduced in [21]. The WSK capacity of wiretapped Markov Trees is an open problem. We show that a wiretapped PIN is a wiretapped Markov

Tree but the converse is not true. Thus, Theorem 4.3 resolves the capacity problem for a large class of wiretapped Markov Trees – i.e., wiretapped PIN. Moreover, we show that Theorem 4.3 can be extended furthermore and gives WSK capacity for an even larger class of wiretapped Markov Trees. Finally, we consider the case when in a wiretapped PIN there is a non-cooperative compromised terminal. For this case we show that WSK capacity is equal to the PK capacity of the same wiretapped PIN in which the compromised terminal is cooperative. In fact, this result generalizes Proposition 4.1 of [60].

4.5.1   WSK Capacity of Wiretapped PIN

For the case of wiretapped PIN (as defined in 4.4), we give a lower bound and an upper on the WSK capacity. These bounds are tight for the special cases of $\mathcal{A} = \mathcal{M}$ and $|\mathcal{A}| = 2$. Finding the WSK capacity of a wiretapped PIN as defined in Definition 4.4 for any given $\mathcal{A}$ remains an open problem.

Proposition 4.11. *For any given wiretapped PIN* $(X_{\mathcal{M}}, Z)$*, described by* $G = (\mathcal{M}, \mathcal{E})$ *and* $P_{ZX_{\mathcal{M}}}$*, and for any* $\mathcal{A} \subseteq \mathcal{M}$*, let* $R_{ij} = I(V_{ij}; V_{ji}|Z_{ij})$*, then we have*

$$
C^{\mathcal{A}}_{WSK}(P_{X_{\mathcal{M}}Z}) \leqslant \min_{\mathcal{P}} \left(\frac{1}{|\mathcal{P}| - 1} \right) \left[\sum_{\substack{i<j \ s.t. \\ (i,j) \ crosses \ \mathcal{P}}} R_{ij} \right],
$$

where the minimization is over all partitions of $\mathcal{M}$ *such that for every part of the partition there exists a node in that part that is also in* $\mathcal{A}$*. In a partition* $\mathcal{P}$ *a pair of nodes* (i, j) *crosses* $\mathcal{P}$*, if* i *and* j *are in different parts of* $\mathcal{P}$*.*

Proof: The proof goes along the same lines as the proof in [21, Example 4]. According to Lemma 4.2 we know $C^{\mathcal{A}}_{WSK}(P_{X_{\mathcal{M}}Z}) \leqslant C^{\mathcal{A}|\{m+1\}}_{PK}(P_{X_{\mathcal{M}}Z})$, and for any $\mathcal{B} \subset \mathcal{M}$ we have

$$
\sum_{j \in \mathcal{B}} R_j \geqslant \sum_{\substack{i<j \\ s.t. \ e_{ij} \in \mathcal{E}_{\mathcal{B}}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i<j \\ s.t. \ i \in \mathcal{B}, j \notin \mathcal{B}}} H(V_{ij}|V_{ji}, Z_{ij}). \tag{4.13}
$$

Consider a partition $\mathcal{P} = \{B_1, \ldots, B_{|\mathcal{P}|}\}$ of $\mathcal{M}$. Then, corresponding to each part of $\mathcal{P}$ we have

$$\sum_{j \in \mathcal{B}_k^c} R_j \geqslant H(X_{\mathcal{M}}|Z) - \sum_{\substack{i<j \\ \text{s.t. } e_{ij} \notin \mathcal{E}_{\mathcal{B}_k}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i<j \\ \text{s.t. } (i,j) \text{ crosses } \{\mathcal{B}_k, \mathcal{B}_k^c\}}} H(V_{ij}|V_{ji}, Z_{ij}).$$

By adding all $|\mathcal{P}|$ inequalities, and remembering the fact that $H(V_{ij}, V_{ji}|Z_{ij}) = H(V_{ij}|V_{ji}, Z_{ij}) + H(V_{ji}|V_{ij}, Z_{ij}) + I(V_{ij}; V_{ji}|Z_{ij})$, we get

$$(|\mathcal{P}|-1) \sum_{j \in \mathcal{M}} R_j \geqslant |\mathcal{P}| H(X_{\mathcal{M}}|Z) - \sum_{k=1}^{|\mathcal{P}|} \sum_{\substack{i<j \\ \text{s.t. } e_{ij} \notin \mathcal{E}_{\mathcal{B}_k}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i,j \\ \text{s.t. } (i,j) \text{ crosses } \mathcal{P}}} H(V_{ij}|V_{ji}, Z_{ij}).$$

$$= (|\mathcal{P}|-1) H(X_{\mathcal{M}}|Z) - \sum_{\substack{i<j \text{ s.t.} \\ (i,j) \text{ crosses } \mathcal{P}}} I(V_{ij}; V_{ji}|Z_{ij}),$$

which implies,

$$R_{CO}(X_{\mathcal{A}}|Z) \geqslant H(X_{\mathcal{M}}|Z) - \frac{1}{|\mathcal{P}|-1} \sum_{\substack{i<j \text{ s.t.} \\ (i,j) \text{ crosses } \mathcal{P}}} I(V_{ij}; V_{ji}|Z_{ij}),$$

and thus due to Theorem 4.1

$$C_{PK}^{\mathcal{A}|\{m+1\}}(P_{X_{\mathcal{M}}Z}) \leqslant \frac{1}{|\mathcal{P}|-1} \sum_{\substack{i<j \text{ s.t.} \\ (i,j) \text{ crosses } \mathcal{P}}} I(V_{ij}; V_{ji}|Z_{ij}).$$

Which is also an upper on the WSK capacity $C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z})$. ∎

We show that the Steiner tree packing methods of [60] for key agreement, leads to the following lower bound on the WSK capacity of PIN. A Steiner tree of G for terminals of $\mathcal{A}$ is a subtree of G that spans (connects) all terminals in $\mathcal{A}$. A family of edge-disjoint Steiner trees is called a Steiner tree packing [117]. We show that for each family with ℓ Steiner trees, a secret key of length ℓ can be generated. Let $\mu(G, \mathcal{A})$ denote the maximum cardinality of such family. Therefore, for a general wiretapped PIN we have the following.

132

Proposition 4.12. *The WSK capacity of a wiretapped PIN $(X_\mathcal{M}, Z)$ defined by $G = (\mathcal{M}, \mathcal{E})$ and $P_{ZX_\mathcal{M}}$ for any $\mathcal{A} \subseteq \mathcal{M}$ is lower-bounded by*

$$C_{WSK}^{\mathcal{A}}(P_{X_\mathcal{M} Z}) \geqslant \sup_{n \in \mathcal{N}} \frac{1}{n} \mu(G^n, \mathcal{A}),$$

where $\mathcal{N}$ is the set of n's such that $nI(V_{ij}; V_{ji}|Z_{ij})$ for any (i, j) is integer-valued and for each n, we define a multigraph $G^n = (\mathcal{M}, \mathcal{E}^n)$ such that for any $e_{ij} \in \mathcal{E}$ of G there exists $nI(V_{ij}; V_{ji}|Z_{ij})$ edges between nodes i and j in $\mathcal{E}^n$.

Proof: For a given $n \in \mathcal{N}$, each pair of connected nodes (i, j) establish a pairwise key S_{ij} of length approximately equal to $nI(V_{ij}; V_{ji}|Z_{ij})$. There exists a Steiner tree packing with cardinality $\mu(G^n, \mathcal{A})$; thus, for any Steiner tree of this Steiner packing, the terminals in $\mathcal{A}$ can establish one bit of shared secret key due to Theorem 4.3. Thus, the asymptotic SK rate is $\sup_{n \in \mathcal{N}} \frac{1}{n} \mu(G^n, \mathcal{A})$. Let pairwise keys S_{ij} be all (ϵ_n, σ_n)-SK's such that $\epsilon_n, \sigma_n \in \mathcal{O}(2^{-n})$. We prove that the final key K is an ϵ'_n, σ'_n-SK such that $\lim_{n \to \infty} \epsilon'_n = \lim_{n \to \infty} \sigma'_n = 0$. The reliability of the final key follows similar to the proof of Theorem 4.3, and $\epsilon'_n = |\mathcal{E}|\epsilon_n$. The security of the final key is as follows. By Corollary 2.1.1 each bit of pairwise keys is also σ_n secure. By Lemma 4.5 each bit of the final key is $2(m-1)\sigma_n$ secure, and by Corollary 2.1.3 the final key is $\sigma'_n = 2(m-1)(\log|\mathcal{K}|)\sigma_n$ secure[5]. Since we chose $\epsilon_n, \sigma_n \in \mathcal{O}(2^{-n})$, we have $\lim_{n \to \infty} \epsilon'_n = \lim_{n \to \infty} \sigma'_n = 0$. ∎

Corollary 4.12.1. *For the special case of $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$, the problem of calculating $\mu(G^n, \mathcal{A})$ is efficiently solvable [117]; rendering the above lower bound of Preposition 4.12 achieving the upper bound of Preposition 4.11 if $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$.*

Proof: It has been proven [117, See Menger's theorem in Section 3.3] that When $|\mathcal{A}| = 2$ then the problem of maximal Steiner Tree Packing in multigraph $G^n = (\mathcal{M}, \mathcal{E}^n)$ will reduce to the problem of finding maximum number of edge-disjoint paths connecting

---

[5]We note that one can use our techniques presented in the security part of the proof of Lemma 4.5 to show a tighter secrecy bound, that is $\sigma'_n = 3|\mathcal{E}|\sigma_n$, without requiring σ_n to decay exponentially in n. However, the presented proof here is more straightforward and suffices for the capacity results in Corollary 4.12.1.

the two terminals in $\mathcal{A}$. Thus, for any multigraph $G^n = (\mathcal{M}, \mathcal{E}^n)$ and any arbitrary subset $\mathcal{A} \subseteq \mathcal{M}$ with $|\mathcal{A}| = 2$ we have

$$\mu(G^n, \mathcal{A}) = \min_{\substack{\mathcal{B} \subsetneq \mathcal{M} \\ \text{s.t. } \mathcal{A} \nsubseteq \mathcal{B}}} |\{e_{ij} \in \mathcal{E}^n | (i, j) \text{ crosses } \mathcal{P} = \{\mathcal{B}, \mathcal{B}^c\}\}| .$$

Therefore, we will have the following lower bound.

$$
\begin{aligned}
C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z}) &\overset{(a)}{\geqslant} \sup_{n \in \mathcal{N}} \frac{1}{n} \mu(G^n, \mathcal{A}) \\
&\overset{(b)}{=} \min_{\substack{\mathcal{B} \subsetneq \mathcal{M} \\ \text{s.t. } \mathcal{A} \nsubseteq \mathcal{B}}} \left[\sum_{\substack{i < j \text{ s.t.} \\ (i,j) \text{ crosses } \mathcal{P} = \{\mathcal{B}, \mathcal{B}^c\}}} I(V_{ij}; V_{ji} | Z_{ij}) \right] \\
&\overset{(c)}{=} C_{PK}^{\mathcal{A} | \{m+1\}}(P_{X_{\mathcal{M}}Z}),
\end{aligned}
$$

where (a) is due to Corollary 4.12, (b) is due to Menger's Theorem, and (c) is due to Lemma 4.11. This proves the tightness of the bound in Corollary 4.12 for $|\mathcal{A}| = 2$.

For the special case of $\mathcal{A} = \mathcal{M}$, in the problem of maximal Steiner Tree Packing in multigraph $G^n = (\mathcal{M}, \mathcal{E}^n)$ the exact value of $\mu(G^n, \mathcal{M})$ is known due to the Tutte/Nash-Williams Theorem [117, Section 3.5], which is

$$\mu(G^n, \mathcal{M}) = \min_{\mathcal{P}} \lfloor \frac{|\{e_{ij} \in \mathcal{E}^n | (i, j) \text{ crosses } \mathcal{P}\}|}{|\mathcal{P}| - 1} \rfloor.$$

Therefore, we have

$$
\begin{aligned}
C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) &\overset{(a)}{\geqslant} \sup_{n \in \mathcal{N}} \frac{1}{n} \mu(G^n, \mathcal{M}) \\
&\overset{(b)}{=} \min_{\mathcal{P}} \left(\frac{1}{|\mathcal{P}| - 1} \right) \left[\sum_{\substack{i < j \text{ s.t.} \\ (i,j) \text{ crosses } \mathcal{P}}} I(V_{ij}; V_{ji} | Z_{ij}) \right], \\
&\overset{(c)}{=} C_{PK}^{\mathcal{M} | \{m+1\}}(P_{X_{\mathcal{M}}Z}),
\end{aligned}
$$

134
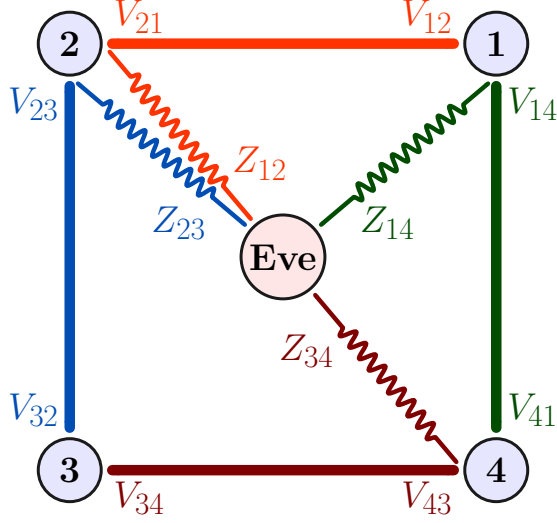
Figure 4.4: The wiretapped PIN of Example 4.3. Here we have 4 terminals, $\mathcal{M} = \mathcal{A} = \{1, 2, 3, 4\}$ and the connectivity graph is given by $G = (\mathcal{M}, \mathcal{E})$, where $\mathcal{E} = \{e_{12}, e_{23}, e_{34}, e_{41}\}$. Terminals variables are $X_1 = (V_{12}, V_{14})$, $X_2 = (V_{21}, V_{23})$, $X_3 = (V_{32}, V_{34})$, and $X_4 = (V_{41}, V_{43})$. Eve's side information is $Z = (Z_{12}, Z_{23}, Z_{34}, Z_{41})$, where the following Markov relations hold: $V_{12} - V_{21} - Z_{12}$, $V_{32} - V_{23} - Z_{23}$, $V_{34} - V_{43} - Z_{34}$, and $V_{41} - V_{14} - Z_{14}$.

where (a) is due to Corollary 4.12, (b) is due to Tutte/Nash-Williams Theorem, and (c) is due to Lemma 4.11. This proves the tightness of the bound in Corollary 4.12 for $\mathcal{A} = \mathcal{M}$.

∎

Example 4.3. To illustrate the result of Corollary 4.12.1, we give the following simple example. Let $m = 4$, and $\mathcal{A} = \mathcal{M} = \{1, 2, 3, 4\}$ and assume that $G = (\mathcal{M}, \mathcal{E})$ is a square as depicted in Figure 4.4. We also assume that for any $e_{ij} \in \mathcal{E}$, $V_{ij} - V_{ji} - Z_{ij}$, such that $R_{ij} = I(V_{ij}; V_{ji}|Z_{ij}) = 1/2$. According to Corollary 4.12.1 and Preposition 4.11, for this example we have

$$C_{WSK}^{\mathcal{M}}(P_{X_{\mathcal{M}}Z}) = \frac{1}{3} \sum_{\substack{(i,j) \text{ crosses } \mathcal{P}}} R_{ij} = \frac{2}{3},$$

where the minimizing partition is $\mathcal{P} = \{\{1\}, \{2\}, \{3\}, \{4\}\}$. To see how the Steiner tree packing method attains this WSK capacity, we first note that if according to each edge $e_{ij} \in \mathcal{E}$, terminals have obtained pairwise SKs of length 3 bits, then a group secret key of length 4 bits can be generated. The reason is that, $G^3 = (\mathcal{M}, \mathcal{E}^3)$ of the square can be

decomposed by 4 edge-disjoint trees, and corresponding to each tree one bit of group SK can be generated. This Steiner tree packing is demonstrated in Figure 4.5.

Recall that for any large enough n, each pair of connected terminals can obtain pairwise keys of length $\ell_{ij}(n) \approx n \times R_{ij} = n/2$ for all $e_{ij} \in \mathcal{E}$. Thus, for any n we find a and b such that $\ell_{ij}(n) = 3 \times b + a$, and thus the final group key will have length of $\ell(n) = 4 \times b + a$. As $n \to \infty$, we will have $\ell(n)/n \to 2/3$, that is the WSK capacity given by Preposition 4.11.

4.5.2 Comparison with Wiretapped Markov Trees and Generalizing Wiretapped Tree-PIN

As examples of the general source model, Csiszár and Narayan introduced the notion of Markov chain on a tree and its wiretapped analogue. We first define the notion of Markov chain on a tree (or Markov Tree in short) as defined in [21].

Definition 4.5 (Markov Tree). Let $\mathcal{M} = [m]$ be a set of m terminals, and let $G = (\mathcal{M}, \mathcal{E})$ be an undirected tree. Note that for any $e_{ij} \in \mathcal{E}$ we can partition $\mathcal{M}$ into two sets $\mathcal{B}_i$ and $\mathcal{B}_j$ such that $\mathcal{M} = \mathcal{B}_i \cup \mathcal{B}_j$, $i \in \mathcal{B}_i$, and $j \in \mathcal{B}_j$. A source model $P_{X_{\mathcal{M}}}$ forms a Markov chain on G if for any $e_{ij} \in \mathcal{E}$ we have $\Pr\left\{X_i | X_{\mathcal{B}_j}\right\} = \Pr\left\{X_i | X_j\right\}$. A special case of such source models is the case when we have $X_1 - X_2 - X_3 - \cdots - X_m$.

For any Markov Tree described by $P_{X_{\mathcal{M}}}$, it is proved that

$$C_{SK}^{\mathcal{A}}(P_{X_{\mathcal{M}}}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(X_i; X_j), \tag{4.14}$$
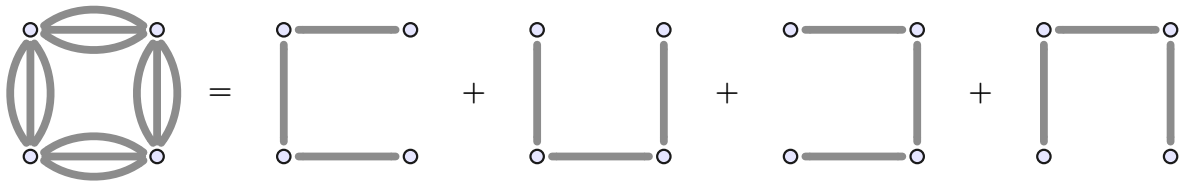
Figure 4.5: Steiner packing of G^3 into 4 edge-disjoint trees.

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ is the smallest subtree connecting all nodes of $\mathcal{A}$. See Example 7, Equation (36) of [21]. The same equation also holds for any given non-wiretapped Tree-PIN – that is implied by Theorem 4.3 when $Z =$ constant. In fact we observe that any Tree-PIN is also a Markov Tree but the converse is not true.

Next, we define the notion of wiretapped Markov chain on a Tree (or wiretapped Markov Tree for short), which was defined first in [21].

Definition 4.6 (Wiretapped Markov Tree). Consider a model $P_{ZX_{\mathcal{M}}}$ where $\mathcal{M} = [m]$ is the set of m terminals and Z is Eve's side information. If Z is of the form $Z = \{Z_1, Z_2, \ldots, Z_m\}$ then we can define an auxiliary model as follows. Let $\mathcal{M}' = \{m+1, \ldots, 2m\}$ be the set of m dummy terminals. Let terminals in $\mathcal{M}$ have access to RVs X_j for all $j \in \mathcal{M}$, and let dummy terminals in $\mathcal{M}'$ have access to RVs Z_{j-m} for all $j \in \mathcal{M}'$. Thus the probability distribution of the auxiliary model defined over $\overline{\mathcal{M}} = \mathcal{M} \cup \mathcal{M}' = \{1, 2, \ldots, 2m\}$, is $P_{X_{\mathcal{M}} Z_{\mathcal{M}}} = P_{X_{\mathcal{M}} Z}$. Any wiretapped SKA model with distribution $P_{X_{\mathcal{M}} Z}$ is called a wiretapped Markov chain on a Tree if, Eve's side information Z is of the form $Z = \{Z_1, Z_2, \ldots, Z_m\}$ such that $\Pr\{Z_{\mathcal{M}} | X_{\mathcal{M}}\} = \Pi_{j \in \mathcal{M}} \Pr\{Z_j | X_j\}$, and if its corresponding auxiliary model defined over $\overline{\mathcal{M}} = \{1, 2, \ldots, 2m\}$ forms a Markov chain on a tree (according to definition 4.5). See an example of such model in the figure 4.6 below.
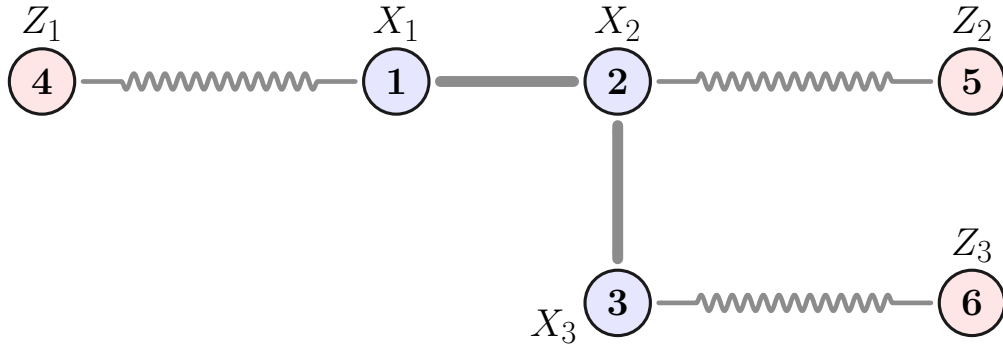
Figure 4.6: A simple wiretapped Markov Chain on a Tree with three terminals. Here, the terminals variables X_1, X_2, X_3 and Eve's side information components Z_1, Z_2, Z_3 form a Markov Tree. The WSK capacity of this model is still unknown.

Unfortunately, for the wiretapped Markov Tree model defined in definition 4.6, where

all terminals are wiretapped, the WSK capacity is not known[6], even for the special case when $m = 2$ (see Figure 4.7 below). The WSK capacity is proved [22] for wiretapped Markov Trees where only one terminal (say terminal 1) is wiretapped, that is $Z = Z_1$ and $Z_j = $ constant $\forall j \neq 1$.

Figure 4.7: A simple wiretapped Markov Chain on a Tree with two terminals. Here, the terminals variables and Eve's side information satisfy the Markov relation of $Z_1 - X_1 - X_2 - Z_2$. The WSK capacity of this model is still unknown.

We observe that every wiretapped Tree-PIN is a wiretapped Markov Tree but the converse is not true. Even though the WSK capacity is not known for all wiretapped Markov Trees, Theorem 4.3, proves the WSK capacity for a large subset of wiretapped Markov Trees. For the special case of $m = 2$ our wiretapped Tree-PIN model and our main result reduces to the well-known case of $X_1 - X_2 - Z$ [19, 20], where $X_1 = V_{12}$, $X_2 = V_{21}$, and $Z = Z_{12}$.

We can extend our model of wiretapped Tree-PIN and obtain a generalized version of Theorem 4.3. In this case for each pair of connected terminals i and j we assume two sets of correlated variables $(V_{ij}^a, V_{ji}^a, Z_{ij}^a)$ and $(V_{ij}^b, V_{ji}^b, Z_{ji}^b)$.

Definition 4.7 (General Wiretapped Tree-PIN). A set of m terminals form a "General Wiretapped Tree-PIN" if there exists a tree $G = (\mathcal{M}, \mathcal{E})$ with $\mathcal{M} = [m]$ such that the RV of any terminal $j \in \mathcal{M}$ can be represented by $X_j = (V_{ji}^\theta | \ i \in \Gamma(j), \theta \in \{a, b\})$, where Eve's side information is of the form $Z = (Z_{ij}^\theta, i \in \mathcal{M}, j \in \mathcal{M}, \theta \in \{a, b\})$ and all pairs of RVs in $\{(V_{ij}^\theta, V_{ji}^\theta, Z_{ij}^\theta) | \ \theta \in \{a, b\}, i < j \text{ and } e_{ij} \in \mathcal{E}\}$ are mutually independent, such that $V_{ij}^\theta - V_{ji}^\theta - Z_{ij}^\theta$ for all $i, j \in \mathcal{M}$ and any $\theta \in \{a, b\}$.

Note that any general wiretapped Tree-PIN is a wiretapped Markov Tree, but the converse is not true.

---

[6]In [21] the authors mistakenly claim to prove the WSK capacity of all wiretapped Markov Trees. See the remark after Theorem 5.1 in [22].
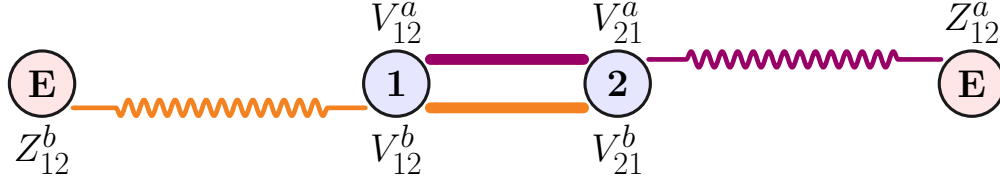
Figure 4.8: A general wiretapped Tree-PIN with two terminals. Here both terminals labeled "E" represent the adversary Eve. Since the Markov relations $V_{12}^a - V_{21}^a - Z_{12}^a$ and $V_{21}^b - V_{12}^b - Z_{12}^b$ hold, we have $Z_{12}^a - V_1 - V_2 - Z_{12}^b$ which resembles the Markov relation in the Markov Tree example of Figure 4.7.

Example 4.4. For the two-party SKA, the general wiretapped Tree-PIN model of definition 4.7 reduces to a case where both terminals are wiretapped. See figure below.

For this case we prove that

$$C_{WSK}(P_{X_1,X_2,Z}) = I(V_{12}^a; V_{21}^a | Z_{12}^a) + I(V_{12}^b; V_{21}^b | Z_{12}^b). \tag{4.15}$$

Proof of Equation 4.15: The achievablity follows directly from Lemma 4.5 applied two times, once for $\theta = a$ and once for $\theta = b$. The converse follows from lemma 4.2 and Theorem 4.1. That is

$$C_{WSK}(P_{X_1,X_2,Z}) \leqslant C_{PK}(P_{X_1,X_2,Z})$$

$$= H(V_{12}^a, V_{21}^a | Z_{12}^a) + H(V_{12}^b, V_{21}^b | Z_{12}^b) - R_{CO}(X_1, X_2 | Z)$$

$$= H(V_{12}^a, V_{21}^a | Z_{12}^a) - H(V_{12}^a | V_{21}^a, Z_{12}^a) - H(V_{21}^a | V_{12}^a, Z_{12}^a)$$

$$+ H(V_{12}^b, V_{21}^b | Z_{12}^b) - H(V_{12}^b | V_{21}^b, Z_{12}^b) - H(V_{21}^b | V_{12}^b, Z_{12}^b)$$

$$= I(V_{12}^a; V_{21}^a | Z_{12}^a) + I(V_{12}^b; V_{21}^b | Z_{12}^b). \qquad \blacksquare$$

Note that the two-party SKA model of Figure 4.7 is more general than the model in Figure 4.8. The WSK capacity of the model of Figure 4.7 is still unresolved, while for the case of general wiretapped Tree-PIN models, including the model of Figure 4.7 can be proved. Moreover, it is easy to see that the following holds – the proof follows the same argument of

the proof of Theorem 4.3 with considering the independence of (a) and (b) variables.

Proposition 4.13 (WSK capacity of general wiretapped Tree-PIN). *The WSK capacity of a given general wiretapped Tree-PIN* $(X_{\mathcal{M}}, Z)$, *defined as in Definition 4.7, for any subset* $\mathcal{A} \subseteq \mathcal{M}$ *is*

$$C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}^a; V_{ji}^a | Z_{ij}^a) + I(V_{ij}^b; V_{ji}^b | Z_{ij}^b), \tag{4.16}$$

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ *is the smallest subtree connecting all nodes of* $\mathcal{A}$.

Note that Preposition 4.13 generalizes Theorem 4.3 as it implies the case of Theorem 4.3 when $\theta \in \{a\}$.

4.5.3 The Case of a Non-cooperative Compromised Terminal

Consider a wiretapped PIN defined by $G = (\mathcal{M}, \mathcal{E})$. Recall that terminals RVs are defined by $X_j = (V_{ji} | \ i \in \Gamma(j))$. Furtherer assume that one terminal (denoted by $\mathcal{D} = \{d\}$) is compromised and is not cooperating with the SKA. Thus, Eve's side information is given by $Z = (Z_{jk} | \ e_{jk} \in \mathcal{E})$ and $X_{\mathcal{D}}$. The following theorem gives the secrecy capacity of this model, which we denote by $C_W(G)$ for simplicity.

Proposition 4.14. *For a given wiretapped PIN defined by* $G = (\mathcal{M}, \mathcal{E})$ *with a non-cooperative compromised terminal denoted by* $\mathcal{D} = \{d\}$, *define the following associated model. Let* $\tilde{G} = (\tilde{\mathcal{M}}, \tilde{\mathcal{E}})$, *where* $\tilde{\mathcal{M}} = \mathcal{M} \backslash \mathcal{D}$ *and* $\tilde{\mathcal{E}} = \mathcal{E} \backslash \{e_{dj} | \ j \in \Gamma(d)\}$. *Eve's side information of the associated model is also defined by* $\tilde{Z} = (Z_{jk} | \ e_{jk} \in \tilde{\mathcal{E}})$. *Then* $C_W(G) = C_W(\tilde{G})$ *where* $C_W(\tilde{G})$ *is the WSK capacity of the associated wiretapped PIN model.*

Proof: The proof follows along the same line as for the proof of Proposition 4.1 of [60]. We show that

$$C_W(\tilde{G}) \overset{(a)}{\leqslant} C_W(G) \overset{(b)}{\leqslant} C_P(G) \overset{(c)}{\leqslant} C_W(\tilde{G}),$$

where $C_P(G)$ denoted the secrecy capacity of model G when compromised terminal is cooperative. To prove (a) we argue that a secrete key for model $\tilde{G}$ also constitutes a valid secret key for model G. Let $Z = (\tilde{Z}, Z_d, X_d)$ where $Z_d = (Z_{dj}|\ j \in \Gamma(d))$. Let K be secrete key established for model $\tilde{G}$ by public communication $\mathbf{F}$ By the independence of (Z_d, X_d) from $(K, \mathbf{F}, \tilde{Z})$ and due to corollary 2.1.2, we have

$$\mathbf{SD}((K, \mathbf{F}, Z), (U, \mathbf{F}, Z)) = \mathbf{SD}((K, \mathbf{F}, \tilde{Z}), (U, \mathbf{F}, \tilde{Z})),$$

which completes the proof of (a). Relation (b) is due to Lemma 4.2 and to prove (c) we show that a secret key based on the protocol that achieves $C_P(G)$ can be used to generate key for model $\tilde{G}$. In model $\tilde{G}$ one terminal, e.g., terminal 1, can use local randomization and simulate X_d^n (since the source distribution is assumed to be known) and reveal it via public communication. Then all terminals can independently simulate their correlated RVs with respect to the compromised terminal d. Therefore, a model is simulated (or emulated) by terminals such that terminal d is compromised and its RV is revealed. Thus, the protocol that achieves $C_P(G)$ can be executed for SKA. Hence, $C_P(G)$ constitutes a lower bound for $C_W(\tilde{G})$. $\blacksquare$

The above results can be regarded as a generalization for Proposition 4.1 of [60] in which $(Z_{jk}|\ e_{jk} \in \mathcal{E})$ = constant.

4.6 Need for Interaction in Source Model SKA

Let N_{PC} denote the number of public communication rounds of an SKA protocol. For noninteractive SKA protocols we have $N_{PC} = 1$, and for interactive ones $N_{PC} > 1$. For two-party SKA in source model, considering the key capacity achieving protocols that use at least use one public message, the following three types of interactions have been studied [20, 31, 93]. (We note that, as shown in [118, 119], for two-party non-wiretapped source model, achieving the maximum rate of common randomness extraction requires public communication, and

two-party SK capacity $C_{SK} = I(X_1; X_2)$ in general is not achievable without using at least a single public message.)

First, is "*one-way*" in which only one party (terminal 1, or Alice) sends a public message to the other party (terminal 2, or Bob). Second, is when each party sends a single public message that is independent of other parties' message. Both these are noninteractive. The third type is "*interactive*" SKA where $N_{PC} > 1$ and in each round, each terminal (party) sends a single message that is a function of the terminal's private samples and previous public messages, and is independent of the other message in the same round. The next round begins when all sent public messages are received by all terminals. See Figure 4.9. The general key capacity of an adversarial model SK, PK, or WSK upper bounds the noninteractive key capacity of the model, and in general we have $C_{XK}^{\rightarrow} \leqslant C_{XK}^{NI} \leqslant C_{XK}$, where $C_{XK}^{\rightarrow}$ denotes the one-way key capacity, C_{XK}^{NI} denotes noninteractive key capacity, C_{XK} denotes the key capacity when interaction is allowed, and $XK \in \{SK, PK, WSK\}$. In the following, we review previous results obtained regarding the required interaction to achieve the key capacity.

Two-party SKA. Ahlswede and Csiszár showed that both two-party SK and PK capacities can be achieved with one-way SKA [20, Preposition 1 and Theorem 3]. That is,

$$\text{(when } m = 2\text{)} \quad \begin{matrix} C_{SK}^{\rightarrow} = C_{SK}^{NI} = C_{SK} \\ C_{PK}^{\rightarrow} = C_{PK}^{NI} = C_{PK} \end{matrix}. \tag{4.17}$$

A single-letter characterization of two-party one-way WSK capacity was derived in [20], where the corresponding one-way capacity achieving SKA protocol is showed to also achieve the general WSK capacity if the Markov condition $X_1 - X_2 - Z$ holds [20, Theorem 1 and its Corollary]. That is,

$$\text{(when } m = 2 \text{ and } X_1 - X_2 - Z\text{)} \quad C_{WSK}^{\rightarrow} = C_{WSK}^{NI} = C_{WSK}. \tag{4.18}$$
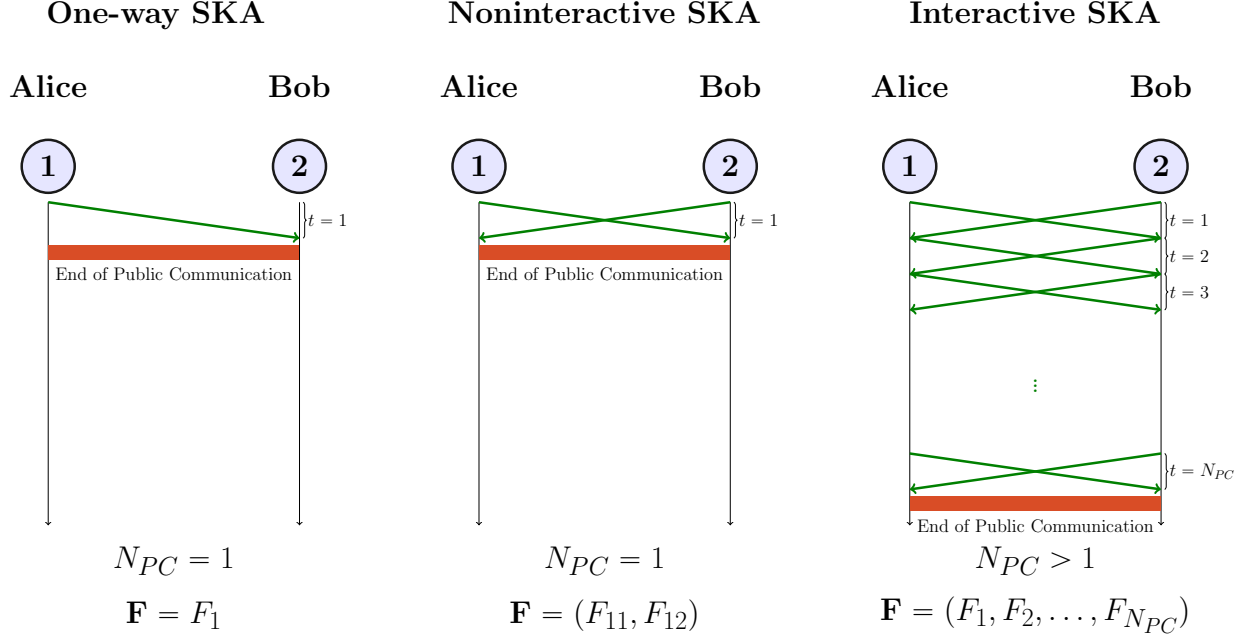
Figure 4.9: Three levels (modes) of interaction for two-party SKA. Note that one-way SKA is an special case of the general noninteractive SKA.

An example is given in [93, Section V, Proof of Theorem 7] for which the one-way WSK capacity is strictly less than the WSK capacity which can be achieved by a noninteractive SKA where both Alice and Bob each send one public message to each other. See also Example 4.4 which is similar to the example given in [93]. This result, proves that in general there is a non-zero gap between the one-way and general WSK capacities, i.e.,

$$\text{(when } m = 2) \quad C_{WSK} - C_{WSK}^{\rightarrow} > 0. \tag{4.19}$$

See the source model of Fig.1 and the last part of the proof for Theorem 7 in [93] for the proof. In other words, one-way SKA is not sufficient to achieve the two-party WSK capacity.

Multiterminal SKA. Extending the statements of (4.17), Csiszár and Narayan showed that for multiterminal SKA, the SK and PK capacities can be achieved noninteractively [21, Theorems 1 and 2]. The best known general lower bound for multiterminal WSK capacity is
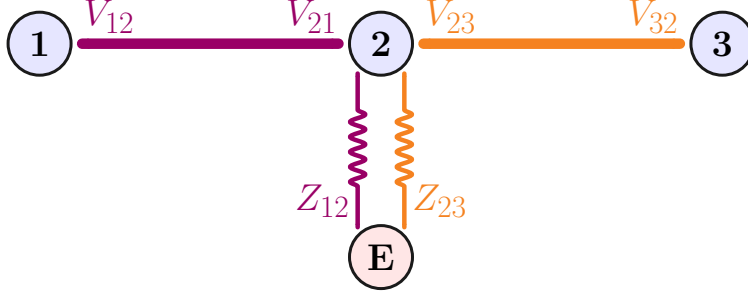
Figure 4.10: The Tree-PIN model of Example 4.5. Here $X_1 = V_{12}$, $X_2 = (V_{21}, V_{23})$, $X_3 = V_{32}$, and Eve's wiretapped side information is $Z = (Z_{12}, Z_{23})$.

the interactive lower bound of [93]. For special cases of Tree-PIN model, WSK capacity can be achieved noninteractively [34, 104]. In this chapter, we gave an interactive SKA protocol (with $N_{PC} = 2$) that achieves the WSK capacity of Tree-PIN sources with independent leakages. However, it remains unknown if interaction is required for achieving the WSK capacity in general.

To investigate if there is a non-zero gap between the general multiterminal WSK capacity and the noninteractive WSK capacity, it is sufficient to know expressions for both capacities at least for a special class of multiterminal source models. For Tree-PIN, we proved an expression for WSK capacity, but the noninteractive WSK capacity of Tree-PIN is not known. In the following, we use a specific example of a Tree-PIN source model (see Figure 4.10) to show that there is a non-zero gap between the WSK capacity and the highest key rate of known noninteractive SKA methods. We prove a lower bound on the noninteractive WSK capacity of this example source model which is strictly less than the WSK capacity. However, we leave the problem of tightening (or closing) this gap for future work.

Example 4.5. Consider the wiretapped Tree-PIN source model of Figure 4.10. In this setting, $\mathcal{M} = \{1, 2, 3\}$, $X_1 = V_{12}$, $X_2 = (V_{21}, V_{23})$, $X_3 = V_{32}$, and Eve's wiretapped side information is $Z = (Z_{12}, Z_{23})$, and the Markov relations $V_{12} - V_{21} - Z_{12}$ and $V_{32} - V_{23} - Z_{23}$ hold. Further, assume $I(V_{21}; Z_{12}), I(V_{23}; Z_{23}) > 0$. When $\mathcal{A} = \mathcal{M}$, the WSK capacity of this

model is given by Theorem 4.3 as

$$C_{WSK} = \min\{I(V_{12}; V_{21}|Z_{12}), I(V_{23}; V_{32}|Z_{23})\}.$$

We prove the following lower bound on the noninteractive WSK capacity of this model

$$C_{WSK}^{NI} \geqslant r_L^{NI} := H(X_2|Z) - \max\{H(X_2|X_1), H(X_2|X_3)\}, \qquad (4.20)$$

which is less that the general WSK capacity, i.e.,

$$C_{WSK} - r_L^{NI} > 0. \qquad (4.21)$$

Proof of Inequalities (4.20) *and* (4.21)*:* We first calculate the noninteractive lower bound r_L^{NI} of (4.20), by considering Protocol 7 ($\boldsymbol{\Pi_{E5}^a}$). The key rate of this protocol immediately follows from the source coding Theorem 2.6 and the generalized privacy amplification Lemma 2.13.

The noninteractive Protocol 7 ($\boldsymbol{\Pi_{E5}^a}$), is in the style of one-way SKA and the SKA protocol of [93] in which some terminals participate in public discussion and some don't (are silent.) Protocol 7 works as follows. Terminal 2, sends a public message such that terminal 1 and terminal 3 can recover X_2^n. Using the common randomness X_2^n all terminals extract their copies of the final key by using universal hashing.

The asymptotic key rate of this protocol can be calculated using Lemma 2.13 as

$$\begin{aligned} r_K(\boldsymbol{\Pi_{E5}^a}) &\overset{(a)}{=} H(X_2|Z) - \min_{F_2} \lim_{n \to \infty} \frac{1}{n} \log \operatorname{supp}(F_2) \\ &\overset{(b)}{=} H(X_2|Z) - \max\{H(X_2|X_1), H(X_2|X_3)\}, \end{aligned}$$

where (a) follows from the fact that the common randomness which is used for group key extraction is RV X_2 and (b) is due to source coding Theorem 2.6.

Protocol 7: First Noninteractive SKA for Tree-PIN of Example 5 ($\mathbf{\Pi_{E5}^a}$)

> **Public Knowledge:** $P_{ZX_{\mathcal{M}}}$ and a family $\mathcal{H}$ of universal hash functions
> $$h_s : \mathcal{X}_2^n \to \mathcal{K} \text{ where } s \in \mathcal{S}.$$
> **Input:** Observations (n–IID samples) X_1^n, X_2^n, X_3^n
> **Output:** Copies of the final key K_1, K_2, K_3
>
> // Information Reconciliation
> **1** Terminal 2 sends public message F_2
> **2** All terminals recover X_2^n
>
> // Privacy Amplification
> **3** All terminals agree on a random seed $s \in \mathcal{S}$ using the public channel
> **4** All terminals extract their keys from X_2^n by $K_j = h_s(X_2^n) \ \forall j \in \{1, 2, 3\}$

Thus, the noninteractive lower bound is then given by

$$r_L^{NI} = H(X_2|Z) - \max\{H(X_2|X_1), H(X_2|X_3)\}.$$

Next, we prove inequality (4.21). Assume that $C_{WSK} = I(V_{12}; V_{21}|Z_{12})$. Then,

$$r_L^{NI} = H(X_2|Z) - \max\{H(X_2|X_1), H(X_2|X_3)\}$$

$$\leqslant H(X_2|Z) - H(X_2|X_1)$$

$$= H(V_{21}|Z_{12}) + H(V_{23}|Z_{23}) - H(V_{21}|V_{12}) - H(V_{23})$$

$$= I(V_{12}; V_{21}|Z_{12}) - I(V_{23}|Z_{23})$$

$$< C_{WSK},$$

where the last inequity holds since $I(V_{23}; Z_{23}) > 0$. Using the same line of argument we can show that $r_L^{NI} < C_{WSK}$ if the WSK capacity is $C_{WSK} = I(V_{23}; V_{32}|Z_{23})$. ∎

Remark 4.4. Finally, we point out that, to our knowledge, Protocol 7 ($\mathbf{\Pi_{E5}^a}$) gives the highest known noninteractive key rate for this example. In fact in the following, we show that the alternative noninteractive approach of SKA by omniscience[7] also leads to the same

---

[7]See also Section 2.3.2.

Protocol 8: Second Noninteractive SKA for Tree-PIN of Example 5 ($\Pi_{\mathbf{E5}}^{\mathbf{b}}$)

Public Knowledge: $P_{ZX_{\mathcal{M}}}$ and a family $\mathcal{H}$ of universal hash functions
$h_s : \mathcal{X}_{\mathcal{M}}^n \to \mathcal{K}$ where $s \in \mathcal{S}$.
Input: Observations ($n-$IID samples) X_1^n, X_2^n, X_3^n
Output: Copies of the final key K_1, K_2, K_3

    // Information Reconciliation
1 Terminal 1 sends public message F_1
2 Terminal 2 sends public message F_2
3 Terminal 3 sends public message F_3
4 Terminals 1 and 3 recover X_2^n
5 Terminals 1 and 2, use F_3 and X_2^n to recover X_3
6 Terminals 3 and 2, use F_1 and X_2^n to recover X_1

    // Privacy Amplification
7 All terminals agree on a random seed $s \in \mathcal{S}$ using the public channel
8 All terminals extract their keys from $X_{\mathcal{M}}^n$ by $K_j = h_s(X_{\mathcal{M}}^n) \ \forall j \in \{1, 2, 3\}$

---

lower bound.

Consider the noninteractive Protocol 8 ($\Pi_{\mathbf{E5}}^{\mathbf{b}}$), which is in the style of SKA by omniscience, similar to the SKA protocol of [21]. Protocol 8 works as follows. Terminal 2, sends a public message such that terminal 1 and terminal 3 can recover X_2^n. Then, terminal 1 (and 3), send public messages F_1 (and F_3), such that other terminals can recover X_1^n (and X_3^n). Using the common randomness $X_{\mathcal{M}}^n$ all terminals extract their copies of the final key by using universal hashing. Let $\mathbf{F} = (F_1, F_2, F_3)$ denote the overall public communication of this protocol.

The asymptotic key rate of this protocol also can be calculated using Lemma 2.13 as

$$
\begin{aligned}
r_K(\Pi_{\mathbf{E5}}^{\mathbf{b}}) &\overset{(a)}{=} H(X_{\mathcal{M}}|Z) - \min_{\mathbf{F}} \lim_{n \to \infty} \frac{1}{n} \log \operatorname{supp}(\mathbf{F}) \\
&\overset{(b)}{=} H(X_{\mathcal{M}}|Z) - \max\{H(X_2|X_1), H(X_2|X_3)\} - H(X_1|X_2) - H(X_3|X_2),
\end{aligned}
$$

where (a) follows from the fact that the common randomness which is used for group key extraction is RV $X_{\mathcal{M}} = (X_1, X_2, X_3)$ and (b) is due to source coding Theorem 2.6.

Noting that $H(X_1|X_2 Z) = H(X_1|X_2)$ and $H(X_3|X_2 X_1 Z) = H(X_3|X_2)$, implies that both SKA protocols have the same asymptotic key rate, $r_K(\mathbf{\Pi^a_{E5}}) = r_K(\mathbf{\Pi^b_{E5}})$.

In summary, the above example, suggests that known noninteractive SKA approaches cannot achieve the general WSK capacity.

4.7 Conclusion

We considered the wiretapped PIN and wiretapped Tree-PIN models. For wiretapped Tree-PIN we proved the WSK capacity and proposed an efficient capacity achieving SKA protocol. The protocol has two rounds and uses any capacity achieving two-party SKA as a subroutine so terminals can obtain pairwise keys. By extending the two-party capacity achieving protocols of [31] and [32] to the case of Tree-PIN, we derived new finite-length lower bounds on the maximum achievable key length. We also proved a finite-length upper bound for the general wiretapped Tree-PIN, and another lower bound for the special case of Tree-PIN studied in [34]. Finally, for wiretapped PIN, we proved a lower and an upper bound for WSK capacity. The bounds are tight when $\mathcal{A} = \mathcal{M}$ or $|\mathcal{A}| = 2$. We extended the Tree-PIN model to two other general cases and proved corresponding WSK capacities. Finally, we investigated the problem of noninteractive key agreement in an example of wiretapped Tree-PIN model, and our analysis suggests that the noninteractive approach for SKA is not sufficient for achieving the general WSK capacity.

4.8 Appendix

4.8.1 Proof of Upper Bound Lemma 4.4

In this section, we prove Lemma 4.4. We prove that for a Tree-PIN specified by the graph $G = (\mathcal{M}, \mathcal{E})$ and probability distribution $P_{ZX_{\mathcal{M}}}$, we have

$$C^{\mathcal{A}}_{WSK}(P_{X_{\mathcal{M}}Z}) \leqslant \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji} | Z_{ij}),$$

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ is the subtree of G with the least number of edges that connects all nodes of $\mathcal{A}$.

Proof of Lemma 4.4: Recall that (due to Lemma 4.2, see also [21, Theorem 4])

$$C^{\mathcal{A}}_{WSK}(P_{X_{\mathcal{M}}Z}) \leqslant C^{\mathcal{A}|\{m+1\}}_{PK}(P_{X_{\mathcal{M}}Z}),$$

where C_{PK} denotes the PK capacity of the associated PIN model given by $\mathcal{M}' = [m+1]$ and $G' = (\mathcal{M}', \mathcal{E}')$ with a dummy node $m+1$ representing the adversary (i.e., $X_{m+1} = Z$). From Theorem 4.1 we know

$$C^{\mathcal{A}|\{m+1\}}_{PK}(P_{X_{\mathcal{M}}Z}) = H(X_{\mathcal{M}}|Z) - R_{CO}(X_{\mathcal{A}}|Z),$$

where $R_{CO}(X_{\mathcal{A}}|Z)$ denotes the solution to the Linear Programming (LP) problem of Figure 4.11, defined over real numbers [21].

Minimize: $\sum_{j \in \mathcal{M}} R_j$

Subject to: $\sum_{j \in \mathcal{B}} R_j \geqslant H(X_{\mathcal{B}}|X_{\mathcal{B}^c}, Z), \quad \forall \mathcal{B} \subsetneq \mathcal{M}, \ \mathcal{A} \nsubseteq \mathcal{B}$ (a)

$R_j \in \mathbb{R}^+, \quad \forall j \in \mathcal{M}.$ (b)

Figure 4.11: The LP problem of finding $R_{CO}(X_{\mathcal{A}}|Z)$.

We prove that

$$R_{CO}(X_{\mathcal{A}}|Z) = H(X_{\mathcal{M}}|Z) - \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}). \tag{4.22}$$

The proof is by first, proving the following lower bound (4.23) and then presenting a rate assignment that achieves the equality, hence proving Equation (4.22).

$$R_{CO}(X_{\mathcal{A}}|Z) \geqslant H(X_{\mathcal{M}}|Z) - \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}). \tag{4.23}$$

Proof of Inequality (4.23): The terminals in $\mathcal{M}$ form a Tree-PIN $G = (\mathcal{M}, \mathcal{E})$. By cutting (removing) an arbitrary edge $e_{i'j'} \in \mathcal{E}$ that connects nodes i' and j', we will have two trees $G_{\langle\mathcal{B}\rangle} = (\mathcal{B}, \mathcal{E}_{\mathcal{B}})$ and $G_{\langle\mathcal{B}^c\rangle} = (\mathcal{B}^c, \mathcal{E}_{\mathcal{B}^c})$, such that $\mathcal{P} = \{\mathcal{B}, \mathcal{B}^c\}$ is a partition of $\mathcal{M}$, and nodes i' and j' each belong to one part of the partition – and $\mathcal{E}_{\mathcal{B}^c} \cup \mathcal{E}_{\mathcal{B}} = \mathcal{E}\backslash\{e_{i'j'}\}$.

Consider the constraints of the LP problem in Figure 4.11 written two times for subsets $\mathcal{B}$ and $\mathcal{B}^c$ individually, and note that $\mathcal{A} \nsubseteq \mathcal{B}$ and $\mathcal{A} \nsubseteq \mathcal{B}^c$. We will have,

$$\sum_{j\in\mathcal{B}} R_j \geqslant H(X_{\mathcal{B}}|X_{\mathcal{B}^c}, Z), \tag{4.24}$$

$$\sum_{j\in\mathcal{B}^c} R_j \geqslant H(X_{\mathcal{B}^c}|X_{\mathcal{B}}, Z). \tag{4.25}$$

From Slepian-Wolf source coding theorem we know that inequality (4.24), implies that if a decoder has access to side information $X_{\mathcal{B}^c}$ and Z, then by receiving the public messages broadcasted by terminals in $\mathcal{B}$, the decoder can reliably recover $X_{\mathcal{B}}$. Also, recall that $X_{\mathcal{B}} = \bigcup_{i\in\mathcal{B}} V_{ij}$. Due to the mutual independence of $\{(V_{ij}, V_{ji}, Z_{ij})\}$'s, we get $H(X_{\mathcal{M}}|Z) = \sum_{i,j} H(V_{ij}, V_{ji}|Z_{ij})$,

and thus we can translate inequalities (4.24) and (4.25) to

$$\sum_{j\in\mathcal{B}} R_j \geqslant \sum_{\substack{i<j \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{B}}} H(V_{ij},V_{ji}|Z_{ij}) + H(V_{i'j'}|V_{j'i'},Z_{ij}),$$

$$\sum_{j\in\mathcal{B}^c} R_j \geqslant \sum_{\substack{i<j \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{B}^c}}} H(V_{ij},V_{ji}|Z_{ij}) + H(V_{j'i'}|V_{i'j'},Z_{ij}).$$

By adding these two inequalities, we arrive at

$$\sum_{j\in\mathcal{M}} R_j \geqslant \sum_{\substack{i<j \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{B}}} H(V_{ij},V_{ji}|Z_{ij}) + \sum_{\substack{i<j \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{B}^c}}} H(V_{ij},V_{ji}|Z_{ij})$$

$$+ H(V_{i'j'}|V_{j'i'},Z_{ij}) + H(V_{j'i'}|V_{i'j'},Z_{ij})$$

$$= \sum_{\substack{i<j \\ \text{s.t. } e_{ij}\in\mathcal{E}}} H(V_{ij},V_{ji}|Z_{ij}) - I(V_{i'j'};V_{j'i'}|Z_{ij})$$

$$= H(X_\mathcal{M}|Z) - I(V_{i'j'};V_{j'i'}|Z_{ij}),$$

where $e_{i'j'}$ denotes the edge that connects the two trees $G_{\langle\mathcal{B}\rangle}$ and $G_{\langle\mathcal{B}^c\rangle}$. We also used the

facts that $\mathcal{E}_{\mathcal{B}^c} \cup \mathcal{E}_\mathcal{B} = \mathcal{E}\backslash\{e_{i'j'}\}$ and that the sets $\{X_j|\ \forall j \in \mathcal{M}\}$ and $\{V_{jk}|\ j < k, e_{jk} \in \mathcal{E}\}$

are indeed equivalent. The above inequality holds for any pair i' and j' of terminals with

$e_{i'j'} \in \mathcal{E}$ and their induced partition $\{\mathcal{B},\mathcal{B}^c\}$, where $\mathcal{A} \nsubseteq \mathcal{B}$ and $\mathcal{A} \nsubseteq \mathcal{B}^c$. Thus,

$$R_{CO}(X_\mathcal{A}|Z) \geqslant \max_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} \{H(X_\mathcal{M}|Z) - I(V_{ij};V_{ji}|Z_{ij})\},$$

$$= H(X_\mathcal{M}|Z) - \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} I(V_{ij};V_{ji}|Z_{ij}),$$

which proves the Inequality (4.23).   ∎

To complete the proof of Equation (4.22), we prove that there exists a rate assignment

protocol that achieves the bound in (4.23).

$$
\boxed{
\begin{array}{ll}
\text{Let } (i^*, j^*) \text{ s.t.} & I(V_{i*j*}; V_{j*i*}|Z_{i*j*}) \\
& \qquad = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{A}}} I(V_{ij}; V_{ji}|Z_{ij}), \text{ and} \\[2em]
\text{for any } j \in \mathcal{M} \text{ let} & R_j = \sum_{i \in \Gamma(j)} \widetilde{R}_i^{(j)}. \\[1.5em]
\hline \\
\text{To minimize } \sum_{j \in \mathcal{M}} R_j & \\[1em]
\text{assign} & \widetilde{R}_{i*}^{(j*)} = H(V_{j*i*}|V_{j*i*}, Z_{i*j*}), \\
& \widetilde{R}_{j*}^{(i*)} = H(V_{i*j*}|V_{i*j*}, Z_{i*j*}), \text{ and} \\
\forall\, e_{ij} \neq e_{i*j*}, & \text{with } d(i,i^*) < d(j,i^*), \\
\text{assign} & \widetilde{R}_i^{(j)} = H(V_{ji}|V_{ij}, Z_{ij}), \text{ and} \\
& \widetilde{R}_j^{(i)} = H(V_{ij}|Z_{ij}).
\end{array}
}
$$
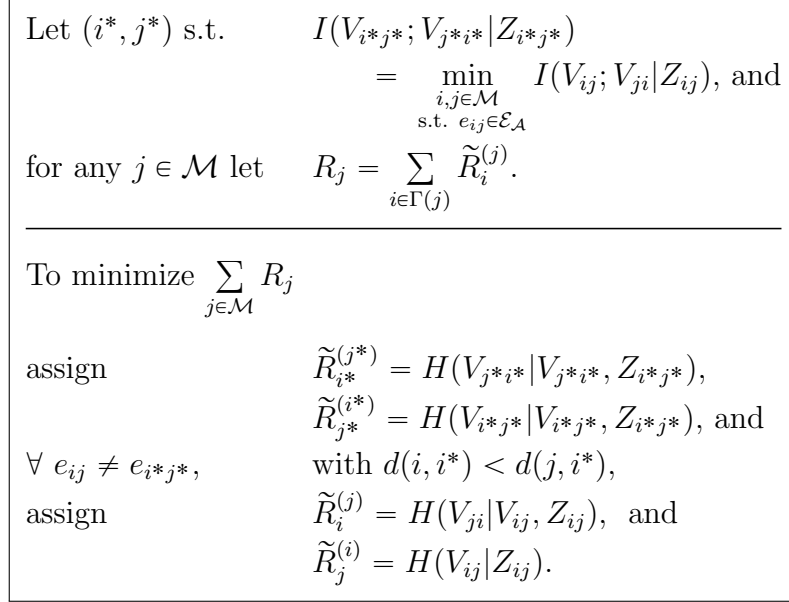
Figure 4.12: The rate assignment that achieves $R_{CO}(X_\mathcal{A}|Z)$.

Proof of Equation (4.22): First, let (i^*, j^*) be defined as follows,

$$
I(V_{i*j*}; V_{j*i*}|Z_{i*j*}) = \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{A}}} I(V_{ij}; V_{ji}|Z_{ij}).
$$

Then for each terminal $j \in \mathcal{M}$ we let the communication rate R_j be chosen according to the rate assignment in Figure 4.12, where $R_j = \sum_{i \in \Gamma(j)} \widetilde{R}_i^{(j)}$ and the rate assignment protocol assigns values to all $\widetilde{R}_i^{(j)}$ components.

This rate assignment satisfies the following equations,

$$
\widetilde{R}_j^{(i)} + \widetilde{R}_i^{(j)} = H(V_{ij}, V_{ji}|Z_{ij}), \quad \forall i, j \in \mathcal{M} \text{ s.t. } e_{ij} \in \mathcal{E} \backslash \{e_{i*j*}\}, \tag{4.26}
$$

$$
\widetilde{R}_{j*}^{(i*)} = H(V_{i*j*}|V_{j*i*}, Z_{i*j*}), \tag{4.27}
$$

$$
\widetilde{R}_{i*}^{(j*)} = H(V_{j*i*}|V_{i*j*}, Z_{i*j*}), \tag{4.28}
$$

which leads to the following sum rate:

$$
\begin{aligned}
\sum_{j \in \mathcal{M}} R_j = \sum_{j \in \mathcal{M}} \sum_{i \in \Gamma(j)} \widetilde{R}_i^{(j)} &= \sum_{\substack{i<j \\ \text{s.t. } e_{ij} \in \mathcal{E}}} \widetilde{R}_i^{(j)} + \widetilde{R}_j^{(i)} \\
&= \sum_{i<j} H(V_{ij}, V_{ji} | Z_{ij}) - I(V_{i*j*}; V_{j*i*} | Z_{i*j*}) \\
&= H(X_{\mathcal{M}} | Z) - I(V_{i*j*}; V_{j*i*} | Z_{i*j*}) \\
&= H(X_{\mathcal{M}} | Z) - \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji} | Z_{ij}).
\end{aligned}
\tag{4.29}
$$

Thus, the rate assignment indeed achieves the lower-bound of Inequality (4.23). We, however, need to show that this rate assignment satisfies the constraints of the LP problem described in Figure 4.11.

First, note that condition (b) in the LP in Figure 4.11 is satisfied as all assigned rates are non-negative. The constraints (a) in the LP can be rewritten for an arbitrary subset of terminals (nodes) $\mathcal{B} \subsetneq \mathcal{M}, \mathcal{A} \nsubseteq \mathcal{B}$ as

$$
\sum_{j \in \mathcal{B}} R_j \geqslant \sum_{i \in \mathcal{B}, j \in \mathcal{B}} H(V_{ij}, V_{ji} | Z_{ij}) + \sum_{i \in \mathcal{B}, j \notin \mathcal{B}} H(V_{ij} | V_{ji}, Z_{ij}).
\tag{4.30}
$$

We show in the following that the rate assignment of Figure 4.12, satisfies the inequality (4.30) for any arbitrary subset $\mathcal{B} \subsetneq \mathcal{M}, \mathcal{A} \nsubseteq \mathcal{B}$. For a given subset $\mathcal{B}$ let $\mathcal{E}_{\mathcal{B}}$ be the set of all edges contained in $\mathcal{B}$ (i.e., $\mathcal{E}_{\mathcal{B}} = \{e_{ij} | \ e_{ij} \in \mathcal{E}, \text{ and } i \in \mathcal{B}, \text{ and } j \in \mathcal{B}\}$). Then, depending on a given subset $\mathcal{B}$ there are two different cases: $I)$ $e_{i*j*} \notin \mathcal{E}_{\mathcal{B}}$, and $II)$ $e_{i*j*} \in \mathcal{E}_{\mathcal{B}}$. The proof is given for all the cases.

Case I) $e_{i*j*} \notin \mathcal{E}_\mathcal{B}$ — The left hand side of the inequality (4.30), can be written as,

$$\sum_{j \in \mathcal{B}} R_j = \sum_{j \in \mathcal{B}} \sum_{i \in \Gamma(j)} \widetilde{R}_i^{(j)}$$

$$= \sum_{j \in \mathcal{B}} \left(\sum_{\substack{i \in \Gamma(j) \\ i \in \mathcal{B}}} \widetilde{R}_i^{(j)} + \sum_{\substack{i \in \Gamma(j) \\ i \notin \mathcal{B}}} \widetilde{R}_i^{(j)} \right)$$

$$\overset{(a)}{\geqslant} \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B}}} \widetilde{R}_i^{(j)} + \widetilde{R}_j^{(i)} + \sum_{\substack{i < j \\ \text{s.t. } i \in \mathcal{B}, j \notin \mathcal{B}}} H(V_{ij}|V_{ji}, Z_{ij})$$

$$\overset{(b)}{=} \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i < j \\ \text{s.t. } i \in \mathcal{B}, j \notin \mathcal{B}}} H(V_{ij}|V_{ji}, Z_{ij}),$$

where, in the (a) we used the fact that $H(V_{ij}|Z_{ij}) \geqslant H(V_{ij}|V_{ji}, Z_{ij})$, and in (b) we used Equation (4.26).

Case II) $e_{i*j*} \in \mathcal{E}_\mathcal{B}$ — The left hand side of the inequality (4.30), can be written as,

$$\sum_{j \in \mathcal{B}} R_j = \sum_{j \in \mathcal{B}} \sum_{i \in \Gamma(j)} \widetilde{R}_i^{(j)}$$

$$= \sum_{j \in \mathcal{B}} \left(\sum_{\substack{i \in \Gamma(j) \\ i \in \mathcal{B}}} \widetilde{R}_i^{(j)} + \sum_{\substack{i \in \Gamma(j) \\ i \notin \mathcal{B}}} \widetilde{R}_i^{(j)} \right)$$

$$= \widetilde{R}_{i*}^{(j*)} + \widetilde{R}_{j*}^{(i*)} + \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B} \setminus \{e_{i*j*}\}}} \widetilde{R}_i^{(j)} + \widetilde{R}_j^{(i)} + \sum_{j \in \mathcal{B}} \sum_{\substack{i \in \Gamma(j) \\ i \notin \mathcal{B}}} \widetilde{R}_i^{(j)}$$

$$\overset{(a)}{=} \widetilde{R}_{i*}^{(j*)} + \widetilde{R}_{j*}^{(i*)} + \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B} \setminus \{e_{i*j*}\}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{j \in \mathcal{B}} \sum_{\substack{i \in \Gamma(j) \\ i \notin \mathcal{B}}} H(V_{ji}|Z_{ij})$$

$$\overset{(b)}{=} \widetilde{R}_{i*}^{(j*)} + \widetilde{R}_{j*}^{(i*)} + \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B} \setminus \{e_{i*j*}\}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i < j \\ j \in \mathcal{B}, i \notin \mathcal{B}}} H(V_{ji}|V_{ij}, Z_{ij}) + I(V_{ij}; V_{ji}|Z_{ij})$$

$$\overset{(c)}{\geqslant} \sum_{\substack{i < j \\ \text{s.t. } e_{ij} \in \mathcal{E}_\mathcal{B}}} H(V_{ij}, V_{ji}|Z_{ij}) + \sum_{\substack{i < j \\ j \in \mathcal{B}, i \notin \mathcal{B}}} H(V_{ji}|V_{ij}, Z_{ij}).$$

154

In (a), we used Equation (4.26), and the rules of the rate assignment protocol, and in (b) we used $H(V_{ji}|Z_{ij}) = H(V_{ji}|V_{ij}, Z_{ij}) + I(V_{ij}; V_{ji}|Z_{ij})$, and in (c) we observe that $\mathcal{B} \subsetneq \mathcal{M}$, which means there always exists at least one node $i \notin \mathcal{B}$ in $G_{\mathcal{A}}$ such that $i \in \Gamma(j)$ for some node $j \in \mathcal{B}$. Thus, on the right-hand-side of (c) there is always a $I(V_{ij}; V_{ji}|Z_{ij})$ and by definition $I(V_{ij}; V_{ji}|Z_{ij}) \geqslant I(V_{i*j*}; V_{j*i*}|Z_{i*j*})$. Also, note that due to (4.27) and (4.28) we have $I(V_{i*j*}; V_{j*i*}|Z_{i*j*}) + \widetilde{R}_{i*}^{(j*)} + \widetilde{R}_{j*}^{(i*)} = H(V_{i*j*}, V_{j*i*}|Z_{i*j*})$.

With the proof of Case I and Case II, the proof of Equation (4.22) is complete. ∎

Equation (4.22) immediately implies that

$$C_{WSK}^{\mathcal{A}}(P_{X_{\mathcal{M}}Z}) \leqslant \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(V_{ij}; V_{ji}|Z_{ij}). \qquad \blacksquare$$

4.8.2   Proof of Lower Bound Lemma 4.5

We prove that SKA protocol 6 achieves the key capacity of any wiretapped Tree-PIN. The proof has three parts: (i) proof of key rate, (ii) proof of reliability, and (iii) proof of secrecy.

Proof of Lemma 4.5: We prove that for any given Tree-PIN with terminals $\mathcal{M} = [m]$ and graph $G = (\mathcal{M}, \mathcal{E})$ and distribution $P_{ZX_{\mathcal{M}}}$, there exists an SKA protocol that achieves the upper-bound of Lemma 4.4 on the wiretap secret key capacity of key agreement for $\mathcal{A} = \mathcal{M}$. We assume that each terminal $j \in \mathcal{M}$ can execute $|\Gamma(j)|$ two-party (pairwise) SKA protocols $\{\pi_{ij}| i \in \Gamma(j)\}$, for extracting pairwise secure keys between terminal (node) j and its neighbors.

Without loss of generality, we assume that the Tree-PIN, is labeled such that node 1 is adjacent to node 2 and $|\Gamma(1)| = 1$. Thus, the edge e_{12} will be included in all paths from node 1 to other nodes in the tree. If the path from i_1 to node i_f, goes through the nodes $i_2, i_3, \ldots, i_{f-1}$, then we denote the path from i_1 to i_f by $\text{Path}(i_1 \to i_f) = (e_{i_1 i_2}, e_{i_2 i_3}, \cdots, e_{i_{f-1} i_f})$.

All terminals in $\mathcal{M}$ will participate in an SKA protocol, described in the pseudo-code 6. In the first phase of the protocol, each terminal j obtains a shared secret key with each member of $\Gamma(j)$. Let $S_{ij} = \pi_{ij}(V_{ij}^n, V_{ji}^n)$ denote the pairwise shared key for any adjacent

nodes i and j. Then, all terminals cut the first ℓ bits of their obtained keys, so that all pairwise keys have the same length. The shortened pairwise keys are $S'_{ij} = S_{ij}|_{\ell}$. The parameter ℓ is a protocol parameter that has to be calculated before running the protocol, according to the known joint distribution $P_{ZX_{\mathcal{M}}}$.

During the public communication phase of protocol 6, each node j finds the unique node $j^* \in \Gamma(j)$ that is closest to node 2. For any other node $k \in \Gamma(j)\backslash\{j^*\}$, node j broadcasts $F_{jk} = S'_{jj*} \oplus S'_{jk}$. Thus, the total number of broadcasts by node j is $|\Gamma(j)| - 1$. Note that each broadcast only uses local variables of node j.

In the last phase of the protocol, terminals 1 and 2 set their final shared keys to be $K_1 = K_2 = S'_{12}$, and the rest of the terminals calculate their obtained keys K_j using the public broadcasted messages (see Protocol 6, line 15).

Proof of Key Rate: It is known that [19, 20] the two-party WSK capacity of a pair of terminals i and j with access to $n-$IID copies of random variables V_{ij} and V_{ji} is $I(V_{ij}; V_{ji}|Z_{ij})$ when $V_{ij} - V_{ji} - Z_{ij}$ holds – see Theorem 2.11-b. That is, there exists a family of (ϵ_n, σ_n) SKA protocols with $\lim_{n\to\infty}(\epsilon_n) = \lim_{n\to\infty}(\sigma_n) = 0$, where $\text{length}(S_{ij}) = \lfloor n\left(I(V_{ij}; V_{ji}|Z_{ij}) - \Delta_n\right)\rfloor$ for some $\Delta_n(\epsilon_n + \sigma_n)$ such that $\lim_{n\to\infty}\Delta_n = 0$. To start protocol 6, fix an arbitrary $\delta > 0$ which is smaller that $\min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$ and choose any ℓ such that

$$\ell \leqslant n\left(\min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}) - \delta - \Delta_n\right).$$

Due to the reliability of the protocol (proved next), every node $j \in \mathcal{M}$, can obtain the same key $K = S'_{12}$ with length ℓ. Thus, the SKA protocol 6, can achieve the asymptotic SK rate of

$$r_K(\mathbf{\Pi_{TP}}) = \lim_{n\to\infty} \frac{1}{n}\text{length}(S'_{12}) = \lim_{n\to\infty} \frac{1}{n}\ell$$
$$\leqslant \lim_{n\to\infty} \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}) - \delta - \Delta_n$$
$$= \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij}) - \delta.$$

156

Since, δ can take any small value, then as $\delta \to 0$, the SK rate of 6 will be arbitrary close to $C = \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$. ∎

Next, we show that the WSK capacity achieving SKA protocol 6 is secure and reliable for any given Tree-PIN. To prove this claim, we need to show

- **Reliability**: Showing that $\Pr\{K_1 = K_2 = \cdots = K_m = K\} \to 1$ as $n \to \infty$, and

- **Secrecy**: Showing that $\mathbf{SD}\left((K, \mathbf{F}, Z), (U, \mathbf{F}, Z)\right) \to 0$ as $n \to \infty$.

Proof of Reliability: Let K_j denote the final key calculated by terminal j. We show that $K_1 = K_2 = \cdots = K_m = S'_{12} = K$, if all $m-1$ pairwise (ϵ_n, σ_n)–SKs S_{ij} are established. For any node $j \in \mathcal{M}\backslash\{1, 2\}$ there is only one path to node 2. This path is of the form $\text{Path}(j \to 2) = (e_{jj^*}, e_{j^*i_1}, e_{i_1i_2}, e_{i_2i_3}, \cdots, e_{i_f2})$, where node j^* is the unique neighbor of j which is closest to node 2 and i_k's $(i = 1\ldots f)$ are the labels for all the nodes (except for j, j^* and 2) that are in the path of j to 2.

In protocol 6, line 10, node k broadcasts $F_{kj} = S'_{i_1k} \oplus S'_{kj}$. Thus, node j who has access to the key S'_{kj} can perfectly recover S'_{i_1k} by computing $S'_{kj} \oplus F_{kj}$. Also, node i_1 (which is connected to i_2 and k) has broadcasted $F_{i_1k} = S'_{i_2i_1} \oplus S'_{i_1k}$. Node j who has now have recovered S'_{i_1k}, can recover $S'_{i_2i_1}$ as well, by computing $S'_{i_1k} \oplus F_{i_1k}$. This chain of recovering local keys will continue until S'_{12} is recovered by computing $K_j = S'_{kj} \oplus F_{kj} \oplus F_{i_1k} \oplus F_{i_2i_1} \oplus F_{i_3i_2} \oplus \cdots \oplus F_{2i_f}$, which proves that $K_j = S'_{12}$ for any $j \in \mathcal{M}$.

This requires all $m-1$ pairwise (ϵ_n, σ_n)–SKs S_{ij} to be established. The error probability of each pairwise key is bounded by ϵ_n, thus the error probability of establishing the global key is $(m-1)\epsilon_n = |\mathcal{E}|\epsilon_n$. Therefore, $\Pr\{K_1 = K_2 = \cdots = K_m = K\} \leqslant 1 - \epsilon'_n$, with $\epsilon'_n = |\mathcal{E}|\epsilon_n$ where ϵ_n such that $\lim_{n\to\infty} \epsilon_n = 0$. ∎

Proof of Secrecy: We need to prove the secrecy of the global shared key K. Without loss of generality, assume that all adjacent terminal pairs i and j with $e_{ij} \in \mathcal{E}$ have established a binary pairwise (ϵ_n, σ_n)–SK S_{ij} with length $\ell = \lfloor n(C - \delta) \rfloor$, where $C = \min_{i,j} I(V_{ij}; V_{ji}|Z_{ij})$. Note that for any $e_{ij} \in \mathcal{E}$ we have $\mathbf{SD}((S_{ij}, Q_{ij}, Z), (U, Q_{ij}, Z_{ij})) \leqslant \sigma_n$,

where U is the uniform distribution over $\{0,1\}^{\ell}$ and Q_{ij} denotes the public communication used to generate S_{ij}. To recall the definition of statistical distance please see Definition 2.7.

Let Q denote the collection of all public communications required to establish all $|\mathcal{E}| = m-1$ pairwise keys S_{ij}, and let F denote the collection of all public communications broadcasted by all terminals during the SKA protocol 6 and $\mathbf{F} = (F, Q)$ be the overall public communication. For any given Tree-PIN $P_{ZX_{\mathcal{M}}}$ with $G = (\mathcal{M}, \mathcal{E})$ we prove that

$$\mathbf{SD}((K, \mathbf{F}, Z), (U, \mathbf{F}, Z)) = \mathbf{SD}((K, F, Q, Z), (U, F, Q, Z))$$
$$\leqslant \mathbf{SD}((K, F, Q, Z), (U, U^{|\mathcal{E}|-1}, Q, Z)) +$$
$$\mathbf{SD}((U, U^{|\mathcal{E}|-1}, Q, Z), (U, F, Q, Z))$$
$$\leqslant |\mathcal{E}|\sigma_n + |\mathcal{E}|\sigma_n = 2|\mathcal{E}|\sigma_n,$$

where U^d is the uniform distribution over $\mathcal{K}^d = \{0,1\}^{d\ell}$.

First we show that *"the combination (K, F) uniquely gives all pairwise keys $\{S_{ij}\}_{i<j}$"*. Recall that any pairwise key belongs to the alphabet $\mathcal{K} = \{0,1\}^{\ell}$. Let $\mathbf{s} = \{s_{ij}\}_{i<j} \in \mathcal{K}^{|\mathcal{E}|}$ be an instance of all pairwise keys. Note that $F = F(\mathbf{S})$ is a set of $m-2$ linear functions of the random vector $\mathbf{S}$. According to Protocol 6 each terminal $j \in \mathcal{M}$ broadcasts $|\Gamma(j)| - 1$ messages. Also recall that for any tree $|\mathcal{E}| = m-1$, so, the total number of public messages is $\sum_{j \in \mathcal{M}} |\Gamma(j)| - 1 = 2|\mathcal{E}| - m = m - 2$. Thus, the $m-2$ elements of F are not sufficient for uniquely finding all $m-1$ pairwise keys in $\mathbf{S}$. However, the combination of F and the final key K resulted by the SKA protocol 6 is sufficient for unique recalculation of all pairwise keys. Remember that $K = S_{12}$ and with all the public messages of terminal 2 one can recover all pairwise keys accessible to terminal 2 since they are all of the form $F_{2j} = S_{12} \oplus S_{2j}$ for all $j \in \Gamma(2) \backslash \{1\}$. Now with access to these pairwise keys one can recover all pairwise keys accessible to any terminal $j \in \Gamma(2) \backslash \{1\}$. This chain of calculation will continue until all pairwise keys are recovered.

Since (K, F) uniquely gives $\{S_{ij}\}_{i<j}$, then

$$\mathbf{SD}((K, F, Q, Z), (U, U^{|\mathcal{E}|-1}, Q, Z)) \leqslant \mathbf{SD}((\{S_{ij}\}, Q, Z), (U^{|\mathcal{E}|}, Q, Z)) \leqslant |\mathcal{E}|\sigma_n.$$

Also, we have $\mathbf{SD}((U, U^{|\mathcal{E}|-1}, Q, Z), (U, F, Q, Z)) \leqslant |\mathcal{E}|\sigma_n$, because,

$$
\begin{aligned}
\mathbf{SD}&((U, U^{|\mathcal{E}|-1}, Q, Z), (U, F, Q, Z)) \\
&= \mathbf{SD}((F, Q, Z), (U^{|\mathcal{E}|-1}, Q, Z)) \\
&\overset{(a)}{=} \sum_{(f,q,z)\in\mathcal{T}^*} P_{QZ}(q, z)P_{F|QZ}(f|q, z) - P_{QZ}(q, z)P_{U^{|\mathcal{E}|-1}}(f) \\
&= \sum_{(q,z)\in\mathcal{T}^*} P_{QZ}(q, z) \sum_{f\in\mathcal{T}^*} P_{F|QZ}(f|q, z) - \frac{1}{|\mathcal{K}|^{|\mathcal{E}|-1}} \\
&\overset{(b)}{\leqslant} \sum_{(q,z)\in\mathcal{T}^*} P_{QZ}(q, z) \sum_{f\in\mathcal{T}^*} P_{F|QZ}(f|q, z) - \frac{1}{|\mathcal{K}|^{|\mathcal{E}|}} \\
&\overset{(c)}{=} \sum_{(q,z)\in\mathcal{T}^*} P_{QZ}(q, z) \sum_{\mathbf{s}\in\mathcal{S}^*(\mathcal{T}^*)} \prod_{i<j} P_{S_{ij}|Q_{ij}Z}(s_{ij}|q_{ij}, z) - \frac{1}{|\mathcal{K}|^{|\mathcal{E}|}} \\
&\overset{(d)}{\leqslant} \max_{\mathcal{T}\subseteq\mathcal{Q}\times\mathcal{Z}\times\mathcal{K}^{|\mathcal{E}|}} \sum_{(q,z)\in\mathcal{T}} P_{QZ}(q, z) \sum_{\mathbf{s}\in\mathcal{T}} \prod_{i<j} P_{S_{ij}|Q_{ij}Z}(s_{ij}|q_{ij}, z) - \prod_{i<j} P_U(s_{ij}) \\
&\overset{(e)}{=} \mathbf{SD}((\{S_{ij}\}, Q, Z), (U^{|\mathcal{E}|}, Q, Z)) \\
&\overset{(f)}{\leqslant} \sum_{i<j} \mathbf{SD}((S_{ij}, Q_{ij}, Z_{ij}), (U, Q_{ij}, Z_{ij})) \\
&\leqslant |\mathcal{E}|\sigma_n,
\end{aligned}
$$

where in (a) $\mathcal{T}^* = \{(f, q, z)|P_{QZ}(q, z)P_{F|QZ}(f|q, z) \geqslant P_{QZ}(q, z)P_{U^{|\mathcal{E}|-1}}(f)\}$ which is due to definition statistical distance (see Definition 2.7,) and in equality (c) $\mathcal{S}^*(\mathcal{T}^*)$ is defined as $\mathcal{S}^*(\mathcal{T}^*) = \{\mathbf{s} \mid \mathbf{s} \in \mathcal{K}^{|\mathcal{E}|}$ and $F(\mathbf{s}) = f, \ \forall f \in \mathcal{T}^*\}$. Inequality (b) is due to the fact that for any $(f, q, z) \in \mathcal{T}^*$ we have $P_{F|QZ}(f|q, z) \geqslant P_{U^{|\mathcal{E}|-1}}(f)$. Relations (d) and (e) are due to the definition of the statistical distance. Inequality (f) follows from Corollary 2.1.3.

Hence, the final key K obtained from the SKA protocol 6 is an $(|\mathcal{E}|\epsilon_n, 2|\mathcal{E}|\sigma_n)-$SK where $\lim_{n\to\infty}(\epsilon_n) = \lim_{n\to\infty}(\sigma_n) = 0$ and the security proof is complete. ∎

With the reliability, security, and key rate proofs, the proof of Lemma 4.5 is complete. ∎

4.8.3   Proof of Theorem 4.9 and Proposition 4.10

Proof: We first recall the SKA protocol that attains the WSK capacity of Tree-PIN. Terminals in $G_{\mathcal{A}}$ –the smallest sub-tree that connects terminals in $\mathcal{A}$ – will generate pairwise keys. Note that in this step, terminals will generate pairwise $(\epsilon', \sigma')-$SKs, where $\epsilon' = \frac{\epsilon}{|\mathcal{E}_{\mathcal{A}}|}$ and $\sigma' = \frac{\sigma}{2|\mathcal{E}_{\mathcal{A}}|}$. Next, all terminals will announce the length of their pairwise keys, and then all pairwise keys will be cut to the minimum length so every pairwise key has the same length. After this, middle nodes (terminals) will broadcast appropriate XOR public messages according to the SKA protocol described earlier. According to the proof of Lemma 4.5, the final extracted key is an $(\epsilon, \sigma)-$SK.

If the pairwise keys are generated by the interactive protocol of Hayashi et. al [31], Theorem 15, terminals i and j can obtain a pairwise key of length

$$\ell_{ij} = nI(V_{ij}; V_{ji}|Z_{ij}) - \sqrt{n\Delta_{ij}}Q^{-1}(\epsilon' + \sigma') - \frac{11}{2}\log n + \mathcal{O}(1).$$

If the pairwise keys are generated by the OW-SKA Protocol 4, then terminals i and j can obtain a pairwise key of length

$$\ell_{ij} = nI(V_{ij}; V_{ji}|Z_{ij}) - Q^{-1}(\epsilon')\sqrt{n\Delta'_{ij}} - Q^{-1}(\sigma')\sqrt{n\Delta''_{ij}} - \log n + \mathcal{O}(1),$$

or

$$\ell_{ij} = nI(V_{ij}; V_{ji}|Z_{ij}) - \sqrt{2n}\log(|\mathcal{X}| + 3)(\sqrt{\log\frac{1}{\epsilon'}} + \sqrt{\log\frac{1}{\sigma'}}) - \log n + \mathcal{O}(1).$$

To understand the difference between these two achievability approximations and their applications, see Chapter 3.

For the special case when $V_{ij} = V_{ji}$, there is no need for information reconciliation, and thus we can use the key extraction bound of [86]. Thus, for this case, terminals i and j can

obtain a pairwise key of length

$$\ell_{ij} = nH(V_{ij}|Z_{ij}) - \sqrt{n\Delta''_{ij}}Q^{-1}(\sigma') - \frac{1}{2}\log n + \mathcal{O}(1).$$

By utilizing either SKA approaches, the length of the final key agreed by all terminals in $G_{\mathcal{A}}$ is

$$\ell = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_{\mathcal{A}}}} \ell_{ij},$$

and hence the proof is complete. It's easy to see that with either of these approaches, the SKA protocol 6 attains the capacity of Theorem 4.3. ∎

Chapter 5

A Channel Model of Transceivers for Multiterminal Secret Key Agreement

Abstract. In this chapter, we propose a new multiterminal channel model for information-theoretic secret key agreement (SKA) that realistically models wireless communication settings and generalizes previous models. Multiterminal channel models for SKA are defined by an underlying noisy discrete memoryless channel (DMC) that connects a set of terminals. Terminals use the noisy DMC and communication over a reliable public channel to agree on a shared secret key. Previous channel models assume that each terminal either controls one input to the channel, or receives one output variable of the channel. In our channel model, which we call *the transceiver model*, the underlying channel may be wiretapped and each terminal controls an input variable and observes an output variable of the noisy DMC. First, we give upper and lower bounds for the highest achievable key rate, known as *key capacity*. We then prove the non-adaptive key capacity of general non-wiretapped transceiver model for the case that the input variables of the noisy channel are, IID, and generated independently and non-adaptively without using the public communication as a feedback link. We compare our results with existing literature, and discuss directions for future work.

5.1 Introduction

Multiterminal secret key agreement (SKA) is an important primitive in multi-user security systems. In a multiterminal SKA protocol, a set of terminals cooperate to establish a shared secret key among a target subset of terminals. The obtained secret key can be used for secure message transmission, or other cryptographic protocols. Secret key agreement was first considered in a two-party setting [19], and independently in [20]. The results were later extended to multiterminal scenarios [21–23], which has been extensively studied thereafter (see e.g., [74, 75, 93, 120–122]).

The SKA problem has been studied in the *source model* and the *channel model*. In multiterminal source model of SKA, terminals have access to many IID (independent and identically distributed) samples of correlated random variables (RVs) [21, 93]. In multiterminal channel model of SKA, terminals are connected by a noisy discrete memoryless channel (DMC), which is used to generate correlation among the terminals [22, 23, 75, 120]. In this chapter, we introduce a new multiterminal channel model for SKA that captures real-life wireless settings, and generalizes existing channel models.

In a *multiterminal channel model*, there are m terminals, denoted by $\mathcal{M} = \{1, \ldots, m\}$, and the goal of the SKA protocol is to establish a shared secret key among a designated subset $\mathcal{A} \subseteq \mathcal{M}$ of terminals. Terminals not in the target subset $\mathcal{A}$ are called *helper terminals*; i.e., terminals in $\mathcal{A}^c = \mathcal{M}\backslash\mathcal{A}$. There exists an underlying noisy DMC connecting the terminals. This DMC might be *wiretapped* in general; that is it might leak some side information about the transmitted symbols to a passive wiretapping adversary, Eve. Terminals can also (interactively) send messages over a reliable (noiseless), authenticated (known sender), and public channel that is assumed free. Messages that are sent over this channel are accessible to all terminals and Eve. An SKA protocol has a finite number of rounds. Each round starts with symbol transmission over the noisy DMC, followed by a public discussion among terminals over the public channel. At the end of the protocol, terminals in $\mathcal{A}$ compute their copy of the secret key. They key may or may not be learned by the helper terminals. An

SKA protocol is called *reliable* if the same key is obtained by all terminals, and *secure*, if Eve has no information about the shared key. The *secret key rate* of an SKA protocol is the ratio of the key length to the number of times that the noisy DMC was used. The *key capacity* of a model is the highest achievable secret key rate [21, 23].

Eve's information can include leaked side information during symbol transmissions over the DMC, leaked information from compromised terminals, and the public messages that are sent by terminals over the public channel. The following types of secret key capacity are defined with respect to Eve's information [21, 23].

- **Secret Key (SK) capacity**: Eve has no side information about the symbol transmissions over the DMC.

- **Private Key (PK) capacity**: Eve has compromised a subset of helper terminals $\mathcal{D} \subseteq \mathcal{A}^c$ and has access to all the symbols transmitted or received by the compromised terminals. All helper terminals, including compromised terminals, cooperate in the SKA protocol. It is assumed that the compromised terminals of $\mathcal{D}$ make their observations and variables public.

- **Wiretap Secret Key (WSK) capacity**: The adversary has access to side information about the symbol transmissions over the DMC. Eve's side information is modeled as an output variable of the DMC.

The problems of finding these capacities are unresolved for many general channel models, and are only known for some special cases. WSK capacity is the most general notion of key capacity, which remains an open problem even for the case of two-party SKA (i.e., $m = 2$).

5.1.1 Our Contributions

Existing channel models assume that a terminal either controls an input or has access to an output symbol of the underlying DMC. In this chapter, we consider scenarios where terminals

can send to, and receive from, the underlying channel. Such terminals model *transceiver* wireless devices [123, Chapter 14]. Here, we introduce a new channel model that we call the "channel model of transceivers" (or the "transceiver model" for short), in which each terminal provides input to, *and* receives output from, the channel. A similar multiterminal model of transceivers has been considered and studied for multi-user communications in [124]. The variable associated with a terminal $j \in \mathcal{M}$, is of the form $V_j = (X_j, Y_j)$, where X_j's are input variables and Y_j's are output variables of the DMC. This model has the channel model of [22] and the multiaccess model of [23] as special cases.

We prove general lower bounds on the SK, PK, and WSK capacities of our proposed model using the proof ideas from [22, 23]. We also prove general upper bounds on the SK, PK, and WSK capacities of transceiver model, by relating any upper bound on the SK and PK capacities of the multiaccess model as a corresponding upper in our proposed model. The bounds, however, are not tight in general. We also use our methods to prove the SK capacity for a special cases of transceiver model, where we assume that input symbols are IID and chosen independently (public channel is not used in between the uses of the DMC.)

Our work raises many interesting questions for future work, including finding tighter bounds for the SK and PK capacities, and investigation of interactive protocols for achieving the key capacity of transceiver models.

5.1.2 Related Works

Existing channel models in literature differ in the way terminals control input or access the output of the noisy DMC, the type of side information that is available to Eve, and the way terminals use the public channel. Single-input multi-output multiterminal DMC was first considered in the channel model of [22] where $\mathcal{A} \subseteq \mathcal{M}$, and all terminals are allowed to send public messages. In the single-input multi-output channel model of [75] however, $\mathcal{A} = \mathcal{M}$, and a subset of terminals $\mathcal{U} \subseteq \mathcal{M}$ participate in public communication while the remaining terminals are silent (i.e., not sending public messages). An important generalization of the

model in [22] is the *multiaccess* channel model of [23], in which $\mathcal{A} \subseteq \mathcal{M}$, a subset of terminals provide input to the DMC and the remaining terminals (which is a disjoint subset from the first subset) are receiving channel outputs. All terminals can send public messages.

For two-party SKA in channel model it is known that [20, Proposition 1]

$$C_{SK} = \max_{P_X} I(X;Y),$$

where X denotes the input variable of terminal 1 to the DMC, and Y denotes the output variable that terminal 2 observes. The known results on the SK and PK capacities of single-input multi-output models are, the SK and PK capacities in [22], and upper and lower bounds of [75]. For the multiaccess channel model, the SK and PK capacities are not known in general. Upper bounds and lower bounds on the SK and PK capacities of the multiaccess model were given in [23], where the lower bounds are based on source emulation approach. We use this approach to derive a lower bound for our proposed model (see Section 5.3.2). The SK capacity is proved for the *symmetric multiaccess channel*[1] with single output under the constraint that input terminals are silent [120]. It was showed that this SK capacity is achievable by an interactive SKA protocol.

Single-letter characterization of WSK capacity for any given channel model remains an open problem, even for the case of $m = 2$, and it is known only for few special cases [45, 74]. It is proved [20, Theorem 2] that WSK capacity of the two-party channel model is upper bounded by

$$C_{WSK} \leqslant \max_{P_X} I(X;Y|Z),$$

where X denotes the input variable of terminal 1, and Y and Z denote the output variables observed by terminal 2 and adversary, respectively. The above bound is tight if (i) the Markov relation $X - Y - Z$ holds or if (ii) the adversary's variable Z is revealed to terminals 1 and 2 [20]. The two-party WSK capacity is also known [20, Theorem 2] to be equal to the

---

[1]A two-input single-output channel is called symmetric if the conditional probability distribution of the channel satisfies $P_{V_3|V_1V_2} = P_{V_3|V_2V_1}$. For the general definition see Section VII of [120].

wiretap secrecy capacity [57] of the underlying DMC when we restrict public communication to noninteractive one-way messages from one terminal to another. That is

$$C_{WSK}^{\rightarrow} = \max_{P_{X'X}}\{I(X';Y) - I(X';Z)\},$$

where the maximum is over all distributions $P_{X'X}$ such that $X' - X - (Y, Z)$ holds. The WSK capacity of a special class of multiterminal models (called Markov Tree) was derived in [22, Theorem 5.1]. In this model Eve's side information is about the variable of a single terminal. Authors of [125], generalized our transceiver model [36] by allowing rate limited private communication, and proved a general upper bound that implies the general upper bound for transceiver model (Theorem 3 of [36]), and the upper bound given in [23] for the multiaccess model.

5.1.3 Organization

We introduce our transceiver model in Section 5.2, and prove general upper and lower bounds on SK and PK capacities in Section 5.3. In Section 5.4, we derive the non-adaptive SK capacity of the transceiver model, and we conclude the chapter in Section 5.5.

5.2 A General Channel Model of Transceivers

5.2.1 The Model

Consider a set of m terminals denoted by $\mathcal{M} = [m] := \{1, \ldots, m\}$. The goal of an SKA protocol is for terminals in $\mathcal{M}$ to cooperate (using the public communication) so that terminals in a subset $\mathcal{A} \subseteq \mathcal{M}$ can establish a shared secret key K. Terminals in $\mathcal{A}^c = \mathcal{M}\backslash\mathcal{A}$ are called *helper terminals*. The key K is not required to be concealed from the helper terminals. All terminals have access to a public, reliable, and authenticated channel. A public message sent by a terminal j will be received by all terminals and everyone else, including the passive

adversary Eve, who will not interfere with the public communication.

There exists an underlying DMC (discrete memoryless channel) which will be used for generating the correlation among terminals. For each transmission over the channel, all terminals provide input to the noisy channel *and* receive output from it; i.e., we assume a set of "*transceivers.*" Each terminal j has two RVs, X_j which is an input variable to the DMC, and Y_j which is an output variable of the DMC, and so the RV associated with each terminal j is given by $V_j = (X_j, Y_j)$, where $\mathcal{V}_j = \mathcal{X}_j \times \mathcal{Y}_j$. Let $V_{\mathcal{M}} = (V_1, \ldots, V_m)$ denote the set of all RV's accessible to all terminals. Eve may also have access to side information Z which is an output RV of the DMC and is correlated with $V_{\mathcal{M}}$. The underlying multi-input multi-output DMC is denoted by $W = (\mathcal{X}_{\mathcal{M}}, P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}, \mathcal{Y}_{\mathcal{M}} \times \mathcal{Z})$, where

$$P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} : \mathcal{X}_1 \times \cdots \times \mathcal{X}_m \to \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_m \times \mathcal{Z} \tag{5.1}$$

is the transition matrix (conditional probability distribution) defined over the finite input alphabet $\mathcal{X}_1 \times \cdots \times \mathcal{X}_m$ and finite output alphabet $\mathcal{Y}_1 \times \cdots \times \mathcal{Y}_m \times \mathcal{Z}$.

Before starting any SKA protocol, terminals are allowed to use the public channel for initialization (e.g., agreeing on public parameters or variables). An SKA protocol consists of n rounds, where each round consists of one invocation of the noisy channel, followed by public communication by terminals in $\mathcal{M}$ over the public channel. Let $\mathbf{F}^t$ denote the random variable representing all public messages of the m terminals in round $1 \leqslant t \leqslant n$, and let $\mathbf{F} = (\mathbf{F}^1, \ldots, \mathbf{F}^n)$ denote the entire public communication during the SKA protocol. Each public message of terminal j in round $1 \leqslant t \leqslant n$ is a function of all previous samples $V_{j1}, V_{j2}, \ldots, V_{jt}$, its local randomness, public messages of the previous rounds $\mathbf{F}^1, \mathbf{F}^2, \ldots, \mathbf{F}^{t-1}$, and previous public message sent in round t. Each input symbol X_{jt} of round $t \geqslant 2$ may depend on previous public discussions $\mathbf{F}^1, \mathbf{F}^2, \ldots, \mathbf{F}^{t-1}$, and previous samples $V_{j1}, V_{j2}, \ldots, V_{j(t-1)}$. After the n rounds of the SKA protocol, the RV associated to each terminal j is given by $V_j^n = (X_j^n, Y_j^n)$. Also, let $V_{\mathcal{M}}^n$ denote the collection of all RVs accessible to all terminals after round
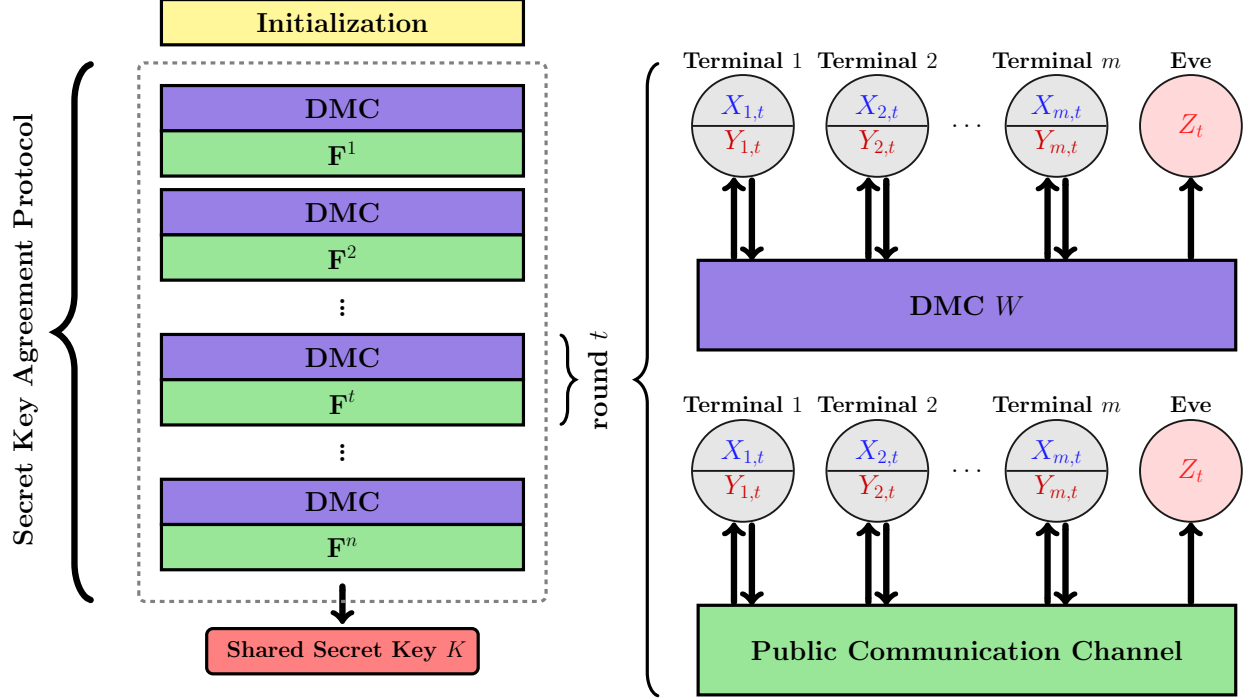
Figure 5.1: An SKA protocol runs over multiple rounds. Each round starts with one invocation of the noisy DMC, followed by a public discussion over the public channel. Eve received all public messages, $\mathbf{F}$, and side information $Z^n = (Z_1, Z_2, \ldots, Z_n)$.

n. The protocol ends when each terminal j computes their version of the key $K_j(V_j^n, \mathbf{F})$ that is a function of all the symbols they send to, or receive from, the noisy DMC (V_j^n), and all of the exchanged public messages ($\mathbf{F}$). Eve has access to all public messages, $\mathbf{F}$, and the side information Z^n, which is correlated with $V_{\mathcal{M}}^n$.

Throughout this work, when adversary Eve, has access to side information Z^n we call the model *wiretapped* and denote the DMC by $W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$. When there is no side information accessible to Eve, we call the model *non-wiretapped* and denote the DMC by $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$. In a non-wiretapped model, there is no Z variable and thus equivalently we assume $Z = $ constant.

5.2.2 Definitions

Definition 5.1. Consider a set of m terminals $\mathcal{M}$, where $\mathcal{A} \subseteq \mathcal{M}$ denotes the set of terminals that will share a key K with alphabet $\mathcal{K}$. Let Z^n denote Eve's side information about $V_{\mathcal{M}}^n$. The key K is an (ϵ, σ)-Secret Key (in short (ϵ, σ)-SK) for $\mathcal{A}$, if there exists an SKA protocol with public communication $\mathbf{F}$, and output RVs $\{K_j\}_{j \in \mathcal{A}}$ for each terminal, such that

$$\text{(reliability)} \quad \Pr\{K_j = K\} \geqslant 1 - \epsilon, \quad \forall j \in \mathcal{A}, \tag{5.2}$$

$$\text{(secrecy)} \quad \mathbf{SD}\left((K, \mathbf{F}, Z^n), (U, \mathbf{F}, Z^n)\right) \leqslant \sigma, \tag{5.3}$$

where $\mathbf{SD}$ denotes the statistical distance and U is the uniform probability distribution over alphabet $\mathcal{K}$.

Definition 5.2 (Key Capacity – see Definition 17.16 of [90]). Consider multiterminal SKA for a subset $\mathcal{A} \subseteq \mathcal{M}$. Let Z^n denote Eve's side information about $V_{\mathcal{M}}^n$. For a given channel model W, where W is the conditional distribution of the underlying DMC, a real number $R \geqslant 0$ is an achievable SK rate if there exists an SKA protocol that for every n establishes an (ϵ_n, σ_n)–SK $K \in \mathcal{K}$ where $\lim_{n \to \infty} \epsilon_n = \lim_{n \to \infty} \sigma_n = 0$, and $\liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{K}| = R$. The maximum of all achievable SK rates is called the key capacity of given model W.

SK, PK, and WSK Capacities. In all cases, the adversary (Eve) has access to all public messages, denoted by $\mathbf{F}$. In addition to $\mathbf{F}$, Eve might have side information about $V_{\mathcal{M}}^n$. When the adversary has no side information about $V_{\mathcal{M}}^n$, the capacity is called *SK capacity*, and denoted by $C_{SK}^{\mathcal{A}}(W)$. In this case, there is no Z^n variable for Eve, and thus equivalently we let $Z^n = $ constant (i.e., independent of $V_{\mathcal{M}}^n$). The adversary may compromise a subset of terminals $\mathcal{D} \subset \mathcal{A}^c$, in which case Eve's side information is of the form $Z^n = V_{\mathcal{D}}^n = (V_j^n | \forall j \in \mathcal{D})$. The compromised terminals are cooperative in the SKA protocol (it is assumed they publicly reveal $V_{\mathcal{D}}^n$ to other terminals.) The capacity for this case is called *PK capacity* and is denoted by $C_{PK}^{\mathcal{A}|\mathcal{D}}(W)$. In the most general sense, if Eve has access to side information Z^n,

which is correlated with $V_{\mathcal{M}}^n$, the key capacity is called *WSK (wiretap secret key) capacity* and is denoted by $C_{WSK}^{\mathcal{A}}(W)$.

Fractional Partition. The following definition will be used a lot for the rest of this chapter.

Definition 5.3 (Fractional Partition, [21, 22]). Consider a finite set $\mathcal{M} = [m] = \{1, 2, \ldots, m\}$. For a subset $\mathcal{A} \subseteq \mathcal{M}$, and $\mathcal{D} \subset \mathcal{A}^c$, define $\Upsilon(\mathcal{A}|\mathcal{D})$ as the family of all nonempty sets $\mathcal{B} \subset \mathcal{D}^c$ such that, $\mathcal{A} \nsubseteq \mathcal{B}$. A fractional partition of $\mathcal{M}$ with respect to $\mathcal{A}$ and $\mathcal{D}$, denoted by $\lambda = (\lambda_{\mathcal{B}}|\ \mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}))$, is a vector of length $|\Upsilon(\mathcal{A}|\mathcal{D})|$ with components $\lambda_{\mathcal{B}} \in [0, 1]$, such that for each $j \in \mathcal{D}^c$

$$\sum_{\substack{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}) \\ \text{s.t. } j \in \mathcal{B}}} \lambda_{\mathcal{B}} = 1. \tag{5.4}$$

We denote by $\Lambda(\mathcal{A}|\mathcal{D})$ the set of all fractional partitions of $\mathcal{M}$ with respect to $\mathcal{A}$ and $\mathcal{D}$. Subset $\mathcal{D}$ can be empty, in which case we simplify our notation to $\Lambda(\mathcal{A})$ and $\Upsilon(\mathcal{A})$.

Remark 5.1. A fractional partition defined by $\lambda \in \Lambda(\mathcal{A}|\mathcal{D})$ allows a terminal j to "fractionally" belong to multiple subsets of $\mathcal{D}^c$, whereas in a partition $\mathcal{P} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_\alpha\}$ of $\mathcal{D}^c$, a terminal j belongs only to one of the parts of the partition. For each subset $\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})$, the component of the λ vector corresponding to $\mathcal{B}$, $\lambda_{\mathcal{B}}$, can be regarded as the "fractional ownership" of $\mathcal{B}$ over the terminals j that are in $\mathcal{B}$. Therefore, for any $j \in \mathcal{D}^c$, the sum of fractions that j belongs to different subsets $\mathcal{B}$ that contain j must be one. This is indeed the defining condition given in Definition 5.3. Hence, it is easy to see that for any given arbitrary partition $\mathcal{P} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_\alpha\}$ of $\mathcal{D}^c$, the λ vector give by $\lambda = (\lambda_{\mathcal{B}}|\ \mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}))$ such that

$$\lambda_{\mathcal{B}} = \begin{cases} 1 & \mathcal{B} = \mathcal{B}_j \text{ for some } j \in \{1, \ldots, \alpha\} \\ 0 & \text{otherwise,} \end{cases}$$
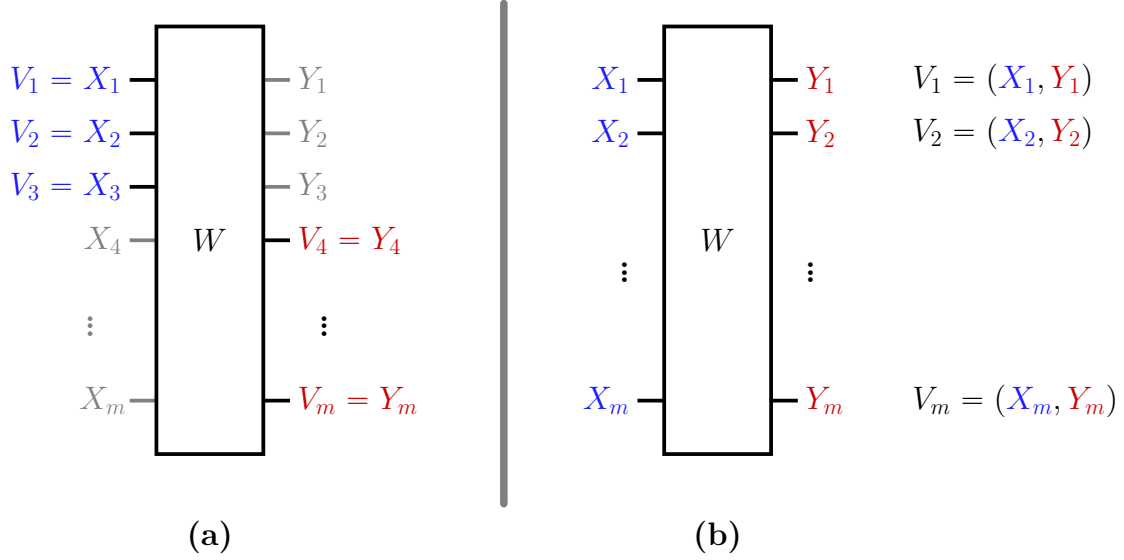
Figure 5.2: (a) The multiaccess channel model of Ref. [23] where $P_{V_{\mathcal{M}\setminus\mathcal{I}}|V_{\mathcal{I}}}$ is the probability transition matrix of DMC W. (b) Our proposed general channel model, where $P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ denotes the transition matrix and for each transceiver terminal j, we have $V_j = (X_j, Y_j)$. Eve's side information is assumed to be $Z =$ constant for both examples here.

characterizes the partition $\mathcal{P}$, and $\lambda \in \Lambda(\mathcal{A}|\mathcal{D})$.

Another important property of the fractional partition is given in the following.

Proposition 5.1. *Consider a finite set $\mathcal{M} = [m] = \{1, 2, \ldots, m\}$. Assign to each $j \in \mathcal{M}$ a random variable V_j. Assume that all V_j RV's are mutually independent. Then, for any $\mathcal{B} \subseteq \mathcal{M}$, $H(V_{\mathcal{B}}) = \sum_{j \in \mathcal{B}} H(V_j)$. Thus, for any given subsets $\mathcal{A} \subseteq \mathcal{M}$, and $\mathcal{D} \subset \mathcal{A}^c$, the following holds for any fractional partition $\lambda \in \Lambda(\mathcal{A}|\mathcal{D})$*

$$\sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}) = \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} \sum_{j \in \mathcal{B}} H(V_j) = \sum_{j \in \mathcal{D}^c} \sum_{\substack{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}) \\ s.t. \ j \in \mathcal{B}}} \lambda_{\mathcal{B}} H(V_j) = \sum_{j \in \mathcal{D}^c} H(V_j) = H(V_{\mathcal{D}^c}).$$

In the proof above we used the mutual independence of V_j's and Equation (5.4).

172

5.2.3 The Relation with Multiaccess Channel Model

The multiaccess channel model, was introduced in [23]. In the multiaccess model, there is a set of m terminals denoted by $\mathcal{M} = [m] = \{1, \ldots, m\}$. A subset of terminals $\mathcal{I} \subset \mathcal{M}$ are called *input terminals*, the rest of terminals in $\mathcal{M} \backslash \mathcal{I}$ are called *output terminals*. There exists a secure noisy DMC between input terminals and output terminals. Input terminals supply input symbols V_j $j \in \mathcal{I}$ to the DMC, and output terminals observe respective output symbols of the DMC. The underlying noisy DMC is called a multiaccess channel and is denoted by $W = (\mathcal{V}_{\mathcal{I}}, P_{V_{\mathcal{M} \backslash \mathcal{I}} | V_{\mathcal{I}}}, \mathcal{V}_{\mathcal{M} \backslash \mathcal{I}})$, where

$$P_{V_{\mathcal{M} \backslash \mathcal{I}} | V_{\mathcal{I}}} : \bigotimes_{j \in \mathcal{I}} \mathcal{V}_j \to \bigotimes_{j \in \mathcal{M} \backslash \mathcal{I}} \mathcal{V}_j.$$

In the multiaccess model of [23], Eve does not have any information about transmission over the DMC. The SK and PK capacities for multiaccess channel model are defined similarly. General upper bounds and lower bounds were proved in [23] for the SK and PK capacities of the multiaccess channel model.

Note that the multiaccess channel model is a special case of the channel model of transceivers by taking $Z = \text{constant}$, $V_j = X_j$ $\forall j \in \mathcal{I}$, and $V_j = Y_j$ $\forall j \in \mathcal{M} \backslash \mathcal{I}$. See Figure 5.2 for a pictorial comparison between the channel model of transceivers, and the multiaccess channel model of [23]. The channel model of [22] is a special case of the multiaccess model of [23] for $|\mathcal{I}| = 1$, and so a special case of our proposed model.

Example 5.1. We give three simple non-wiretapped example models, to compare our transceiver model and the channel models of [22] and [23]. These models are depicted in Figure 5.3.

The DMC of the model in Figure 5.3.(a) is a combination of two independent channels, W_1 and W_2, where W_1 is a point-to-point channel from terminal 1 to 2, and W_2 is a point-to-point channel from terminal 1 to 3. Here, terminal 1 is an input terminal and $V_1 = X_1 = (X_{1,2}, X_{1,3})$. Terminals 2, and 3 are output terminals with $V_2 = Y_2$, and $V_3 = Y_3$. This
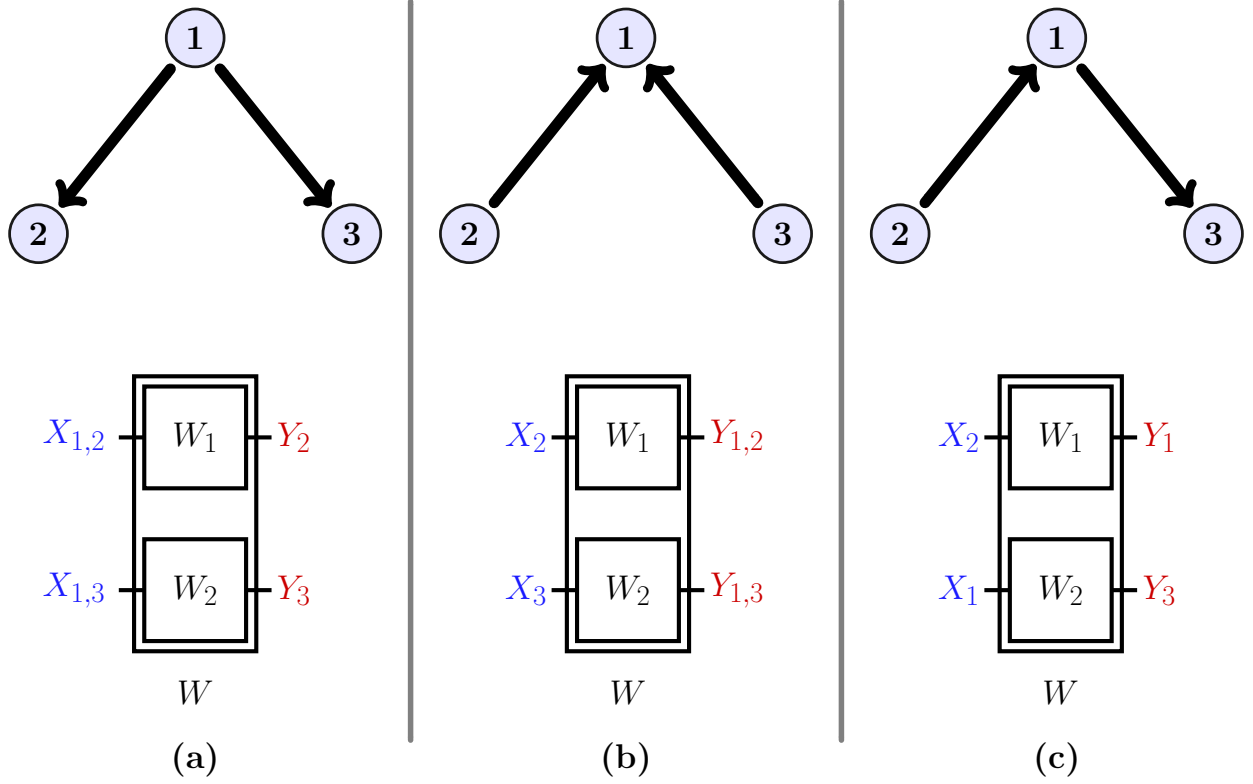
Figure 5.3: Three examples of the transceiver model. The solid arrows show the point-to-point channels between terminals. See Example 5.1 for details.

channel model is a simple example of the single-input multi-output DMC of [22]. The SK and PK capacities of this model can be calculated by [22, Thoerem 4.1].

The DMC of the model in Figure 5.3.(b) is a combination of two independent channels, W_1 and W_2, where W_1 is a point-to-point channel from terminal 2 to 1, and W_2 is a point-to-point channel from terminal 3 to 1. Here, terminal 1 is an output terminal and $V_1 = Y_1 = (Y_{1,2}, Y_{1,3})$. Terminals 2, and 3 are input terminals with $V_2 = X_2$, and $V_3 = X_3$. This channel model is a simple example of the multiaccess DMC of [23]. Upper and lower bounds on SK and PK capacities of this model were given in [22, Thoerem 4 and Theorem 6].

The DMC of the model in Figure 5.3.(c) is a combination of two independent channels, W_1 and W_2, where W_1 is a point-to-point channel from terminal 2 to 1, and W_2 is a point-to-point channel from terminal 1 to 3. Here, terminal 1 is a "transceiver" terminal and

$V_1 = (X_1, Y_1)$. Terminal 2, is an input terminal with $V_2 = X_2$, and Terminal 3, is an output terminal with $V_3 = Y_3$. This channel model and also the two previous models are simple examples of our transceiver model.

More precisely, all of these channel models, are examples of a special class of transceiver model we call the Polytree-PIN model (see Chapter 6). In this chapter, we give upper and lower bounds on the SK, PK, and WSK capacities of such models, and show that these bounds can be tight under certain conditions.

5.3 General Lower and Upper Bounds

In this section, we give general lower and upper bounds for the SK, PK and WSK capacities of the general channel model of transceivers. Later, in Sections 5.4 and then in Chapter 6, we give capacity results for specific channel models by using these lower and upper bounds.

The proof of our lower bound relies on the application of a specific approach to SKA protocols, namely the *source model* SKA. Source model SKA protocols where introduced to achieve the capacity of source models [21], but they also can be used in channel models. Such protocols are important and they have a lot of applications in practice (see [116, 126] and references therein). Therefore, before stating our lower bound result we first recall the general source model of SKA, and the single-letter characterization of its PK capacity as given in [21].

5.3.1 The Multiterminal Source Model

The general multiterminal source model, was introduced in [21]. In this model, there is a set of m terminals denoted by $\mathcal{M} = [m] = \{1, \ldots, m\}$. Each terminal $j \in [m]$ has access to a random variable V_j. Let $V_{\mathcal{M}} = (V_1, \ldots, V_m)$ denote the set of all variables accessible to all terminals. After n IID sampling from $V_{\mathcal{M}}$, terminals use a public channel, that is reliable and authenticated, for a finite number of rounds. A message that is sent by terminal j is

a function of the terminal's IID samples (observations) V_j^n, local randomness, and previous public messages. We denote by $\mathbf{F}$ the set of all messages sent over the public channel. A source model is characterized by its associated public joint probability distribution $P_{V_{\mathcal{M}}}$. The channel models of SKA [22, 23] where in fact introduced as generalizations of the source model.

The SK, PK, and WSK capacities for source model are defined similarly as were defined for the channel models given in Definition 5.2. Note that the key capacity notations, $C_{SK}^{\mathcal{A}}(W)$, $C_{PK}^{\mathcal{A}|\mathcal{P}}(W)$, and $C_{WSK}^{\mathcal{A}}(W)$, refer to a source model capacity if W is a joint distribution, and to a channel model capacity, if W is a conditional distribution.

The following theorem gives an alternative formulation of the PK capacity than the expression presented in Theorem 4.1.

Theorem 5.2 (Source model PK Capacity, see Theorem 3.1 of [22]). *In a given source model $V_{\mathcal{M}}$ described by $P_{V_{\mathcal{M}}}$, for sharing a secret key among terminals in $\mathcal{A} \subseteq \mathcal{M}$, with compromised terminals $\mathcal{D} \subseteq \mathcal{A}^c$, the PK capacity is*

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{V_{\mathcal{M}}}) = \min_{\lambda \in \Lambda(\mathcal{A}|\mathcal{D})} \{ H(V_{\mathcal{M}}|V_{\mathcal{D}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}) \}. \tag{5.5}$$

One obvious situation in a source model for which key agreement is impossible is when the RV's accessible to each terminal $j \in \mathcal{M}$ are mutually independent – i.e., $P_{V_{\mathcal{M}}} = \prod_{j \in \mathcal{M}} P_{V_j}$. In such cases, all terminals are statistically uncorrelated. Note that the PK capacity of such source models as given by Theorem 5.2 is

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{V_{\mathcal{M}}}) = \min_{\lambda \in \Lambda(\mathcal{A}|\mathcal{D})} \{ H(V_{\mathcal{D}^c}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}) \} = \min_{\lambda \in \Lambda(\mathcal{A}|\mathcal{D})} \{ H(V_{\mathcal{D}^c}) - H(V_{\mathcal{D}^c}) \} = 0,$$

where in the second equality we used Proposition 5.1.

Equation (5.5) implies the SK capacity when $\mathcal{D} = \varnothing$. The achievability result is based on a source model SKA protocol in which first, the compromised terminals (that are assumed to

be cooperative) reveal their observed random variables, and then the rest of the terminals in $\mathcal{D}^c$ communicate over the public channel to attain omniscience (i.e., the state that terminals in $\mathcal{D}^c$ learn each other's initial observations). Finally, terminals in $\mathcal{A}$ extract the key from the common shared randomness $V_{\mathcal{M}}^n$. It was also showed that the public communication required to obtain this PK capacity can be noninteractive, meaning that $\mathbf{F} = \mathbf{F}^n = (F_1, \ldots, F_m)$, where $F_j = V_j^n$ for all $j \in \mathcal{D}$ and $F_j = f(V_{\mathcal{D}}^n, V_j^n)$ for all $j \in \mathcal{D}^c$. See the achievablity part of the proof of Theorem 2, in Section IV of [21].

We prove our channel model lower bound based on the *source emulation* approach of [22, 23], that utilizes the source model SKA protocol explained above.

5.3.2   The Source Emulation Lower Bound

Consider the multiaccess channel model (see Section 5.2.3). The *simple source emulation*, introduced in [22], works as follows. For a known IID input distribution $P_{V_{\mathcal{I}}}$, each input terminal $j \in \mathcal{I}$ samples IID symbols V_j^n and transmits their symbols through the DMC. During these n symbol transmissions, terminals do not engage in public discussion. After the symbol transmissions, all terminals have n IID samples according to the IID distribution given by $P_{V_{\mathcal{M}}} = P_{V_{\mathcal{I}}} P_{V_{\mathcal{M} \setminus \mathcal{I}} | V_{\mathcal{I}}}$. This way, in effect, a source model with a known IID distribution is realized (or *emulated*) among terminals of $\mathcal{M}$. Thus, after the symbol transmission steps, any suitable source model SKA protocol can be utilized for key generation.

The source emulation technique is proved to be capacity achieving for single-input multi-output channels [22]– i.e., when $k = 1$ and $\mathcal{I} = \{1\}$. However, in general, using public discussion during symbol transmission can potentially result in more powerful and tighter lower bounds for the multiaccess channel model. This was proved in affirmative for some special multiaccess channels in [120, Theorem 4]. The *general source emulation* is similar to the simple source emulation, and was introduced in [23]. We use the general source emulation approach and prove the following theorem.

Theorem 5.3 (General Source Emulation Lower Bounds). *For a non-wiretapped*

177

channel model of m transceivers $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$, and for any publicly known random variable X' satisfying $P_{X'X_{\mathcal{M}}} = P_{X'} \prod_{j \in \mathcal{M}} P_{X_j|X'}$, we have

$$C_{SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \geqslant C_{PK}^{\mathcal{A}|\{0\}}(P_{X'X_{\mathcal{M}}}P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}),$$

$$= \min_{\lambda \in \Lambda(\mathcal{A}|\{0\})} \left\{ H(V_{\mathcal{M}}|X') - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\{0\})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}) \right\} \tag{5.6}$$

and

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \geqslant C_{PK}^{\mathcal{A}|\mathcal{D}'}(P_{X'X_{\mathcal{M}}}P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}),$$

$$= \min_{\lambda \in \Lambda(\mathcal{A}|\mathcal{D}')} \left\{ H(V_{\mathcal{M}}|V_{\mathcal{D}'}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}')} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}) \right\} \tag{5.7}$$

where $C_{PK}^{\mathcal{A}|\mathcal{D}'}(P_{X'X_{\mathcal{M}}}P_{Y_{\mathcal{M}}|X_{\mathcal{M}}})$ denotes the emulated source model PK capacity of an associated model with $m+1$ terminals, $\mathcal{M}' = \{0, 1, \ldots, m\}$, where $\mathcal{D}' = \mathcal{D} \cup \{0\}$, $V_0 = X'$, $V_j = V_j \; \forall j > 0$, and an underlying source distribution $P_{V_{\mathcal{M}'}} = P_{X'}(\prod_{j \in \mathcal{M}} P_{X_j|X'})P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$. The single letter expressions given in Equation (5.6) and (5.7) for the source model PK capacity are due to [21] – see Theorem 5.2.

Furthermore, for a wiretapped channel model of m transceivers $W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$, and for any publicly known random variable X' satisfying $P_{X'X_{\mathcal{M}}} = P_{X'} \prod_{j \in \mathcal{M}} P_{X_j|X'}$, we have

$$C_{WSK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \geqslant C_{WSK}^{\mathcal{A}}(P_{X'X_{\mathcal{M}}}P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}), \tag{5.8}$$

where $C_{WSK}^{\mathcal{A}}(P_{X'X_{\mathcal{M}}}P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}})$ denotes the emulated source model WSK capacity of an associated model in which Eve's variable is of the form (Z, X'). A single-letter expression for the source model WSK capacity is not known.

We refer to all of the above lower bounds as to general source emulation lower bounds, and we refer to them as to simple source emulation lower bounds when we set $X' = \text{constant}$.

Proof of Theorem 5.3: We show that for a transceiver channel model for terminal
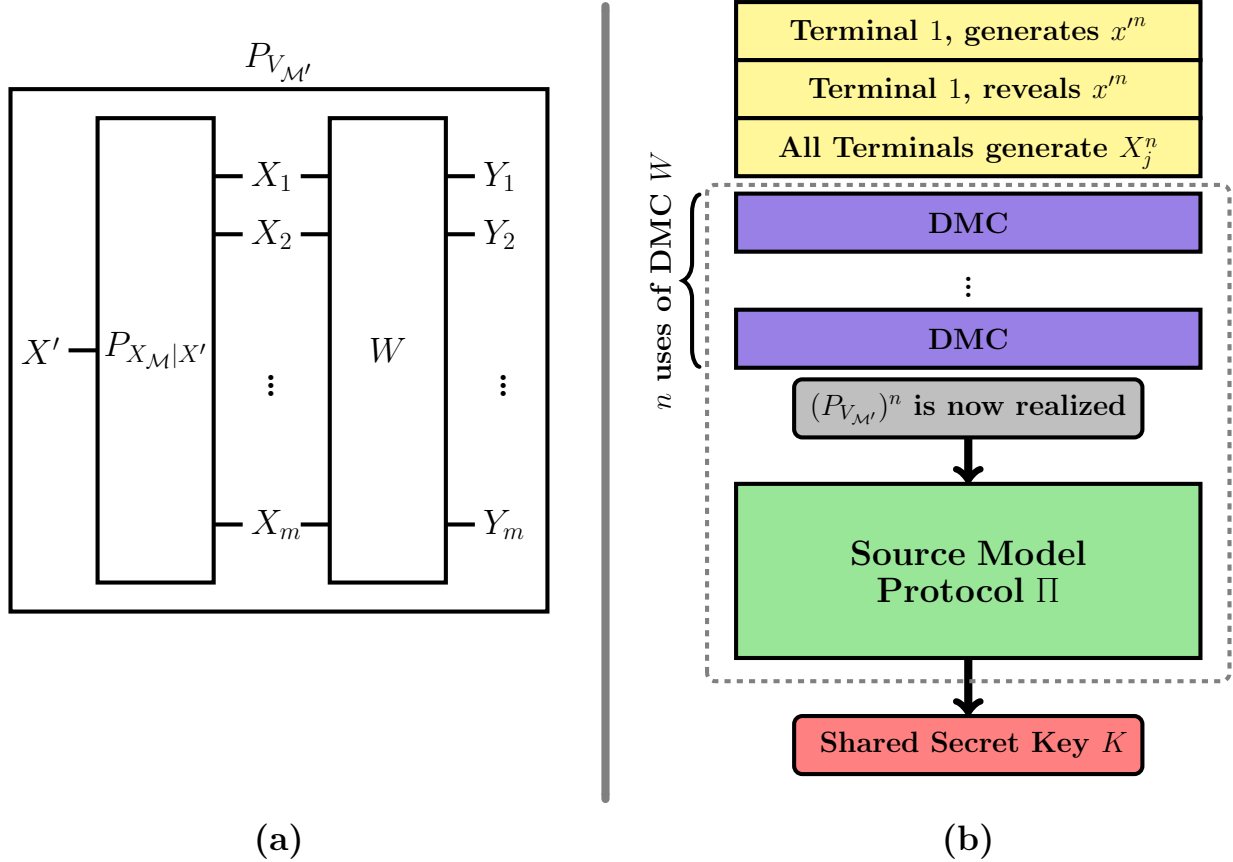
Figure 5.4: (a) The associated source model $P_{V_{\mathcal{M}'}}$ used in the lower bound of Theorem 5.3. (b) The steps of the source emulation SKA protocol that achieves the lower bound of Theorem 5.3.

set $\mathcal{M}$, one can construct a source model for terminal set $\mathcal{M}' = \{0\} \cup \mathcal{M}$, and use capacity achieving source model protocols (e.g., protocol of [21]) in the latter model to obtain a channel model SKA protocol in the transceiver model. This leads to a lower bound on the key capacity of the transceiver model. The case of SK capacity is implied from the argument with $\mathcal{D} = \varnothing$ and $Z = $ constant. The case of PK capacity is implied from the argument with $Z = $ constant; and the WSK capacity is implied from the argument with $\mathcal{D} = \varnothing$.

For a given transceiver channel model $P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ for terminal set $\mathcal{M}$ define an associated source model $P_{V_{\mathcal{M}'}}$ defined over $\mathcal{M}' = \{0\} \cup \mathcal{M}$, where $\{0\}$ is a new terminal added to the terminal set. Let X' denote the random variable of terminal 0. The distribution of this source model is given by $P_{V_{\mathcal{M}'}} = P_{X'} P_{V_{\mathcal{M}}} = P_{X'} (\prod_{j \in \mathcal{M}} P_{X_j|X'}) P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$, where $P_{X'}$

and $P_{X_j|X'}$'s are arbitrary distributions that together generate a distribution on the input symbols of the transceiver channel, which is conditionally independent given X' – i.e., $P_{X_\mathcal{M}|X'} = \prod_{j\in\mathcal{M}} P_{X_j|X'}$. Thus the distribution of $P_{V_\mathcal{M}}$ can be viewed as obtained from symbol transmission over a single-input multi-output channel $P_{X_\mathcal{M}|X'} \cdot W$, where X' is the input symbol, and $V_\mathcal{M} = (X_\mathcal{M}, Y_\mathcal{M})$ denotes output symbols. See Figure 5.4 (a). Terminal 0 is assumed compromised, i.e., $\mathcal{D}' = \{0\} \cup \mathcal{D}$.

First, we emulate (realize) the source model $P_{V_{\mathcal{M}'}}$. Let K be a secret key generated for terminals in $\mathcal{A}$ by the protocol Π that achieves the *source model* key capacity. In Π, the public message of terminal j is a function of X_j^n and Y_j^n. The key K is a function of $V_{\mathcal{M}'}^n$ and $\mathbf{F}$. The protocol Π defines a protocol Π' for the transceiver model, using the following steps. Note that $P_{X'X_\mathcal{M}} = P_{X'}(\prod_{j\in\mathcal{M}} P_{X_j|X'})$ is known. One of the terminals, Terminal 1 for example, generates $x'^n = (x'_1, \ldots, x'_n)$ that is a realization of X'^n, and reveals it to all terminals over the public channel. Each input symbol X_j^n is generated independently (given x'^n) according to $P_{(X_j)t} = P_{X_j|X'=x'_t}$ for all $t \in [n]$. In n consecutive rounds, terminals use the DMC $W = P_{Y_\mathcal{M}|X_\mathcal{M}}$, without using the public communication channel. Thus, after symbol transmission, source model $P_{V_{\mathcal{M}'}^n}$ is emulated for terminals in $\mathcal{M}$. That is each terminal j has access to IID random variables X_j^n and Y_j^n distributed according to the source distribution $P_{V_{\mathcal{M}'}^n}$. Now, terminals in $\mathcal{M}$ can run the source model SKA Π. Compromised terminals send their samples $V_\mathcal{D}^n$ over the public channel. The samples of terminal 0 is also accessible to the rest of the terminals. Messages of terminals in $\mathcal{M}\backslash\mathcal{D}'$ are generated according to Π. Then, all terminals in $\mathcal{A}$ can agree on the common randomness $V_{\mathcal{M}'}$, and extract their secret key. See Figure 5.4 (b). Thus, at the end of Π' the same key K of Π will be established for $\mathcal{A}$, and Π' provides a lower bound on the key capacity of the transceiver channel model. The key rate of Π' is the same as the key rate of Π which can be as large as the emulated source model key capacity. $\blacksquare$

Corollary 5.3.1. *For a wiretapped channel model of m transceivers $W = P_{ZY_\mathcal{M}|X_\mathcal{M}}$, and for*

any publicly known random variable X' satisfying $P_{X'X_{\mathcal{M}}} = P_{X'} \prod_{j \in \mathcal{M}} P_{X_j|X'}$, we have

$$C^{\mathcal{A}}_{WSK}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \geqslant \sum_{i=1}^{\theta} \min_{j \in \mathcal{M}} I(\tilde{V}_i; V_j | \tilde{V}_{[i-1]}) - I(\tilde{V}_i; (Z, X') | \tilde{V}_{[i-1]}) \qquad (5.9)$$

where the lower bound is due to the general source emulation lower bound of (5.8) and the interactive source model SKA protocol of [93, Theorem 7], which holds for any arbitrary integer θ and RV's $\tilde{V}_1, \tilde{V}_2, \ldots, \tilde{V}_\theta$ satisfying

$$\Pr\left\{\tilde{V}_{[\theta]}|V_{\mathcal{M}}Z\right\} = \prod_{i=1}^{\theta} \Pr\left\{\tilde{V}_i|\tilde{V}_{[i-1]}V_{i \mod m}\right\}.$$

Adaptive input symbols VS. The source emulation approach. We emphasis again that the source emulation approach is not always the best approach for SKA, and it is not capacity achieving in general, as such protocols do not employ the possibility of sending adaptive input symbols based on public feedback (which is a function of the received symbols $Y_{\mathcal{M}}$) in between each use of the underlying DMC. In some scenarios, using adaptive input symbols (and public feedback in between each use of the DMC) is strictly required for higher SK rates. An example of such scenario is given in [120, Theorem 4], where the SKA with adaptive inputs outperforms the source emulation technique and achieves higher key rates. However, in some special cases the source emulation approach *is* capacity achieving, see e.g., Theorem 5.5, Theorem 6.1, and [120, Theorem 5].

5.3.3 Upper Bound

In this section, we prove a general upper bound on the SK and PK capacities of the transceiver model. This is by associating a multiaccess channel model [23] to a transceiver model as described below. Given a transceiver channel model $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ over terminal set $\mathcal{M} = [m]$, we define an associated (related) multiaccess channel model $\overline{W}$ over $2m$ terminals denoted by $\overline{\mathcal{M}}$. Consider the original terminal set of W to be the output terminal
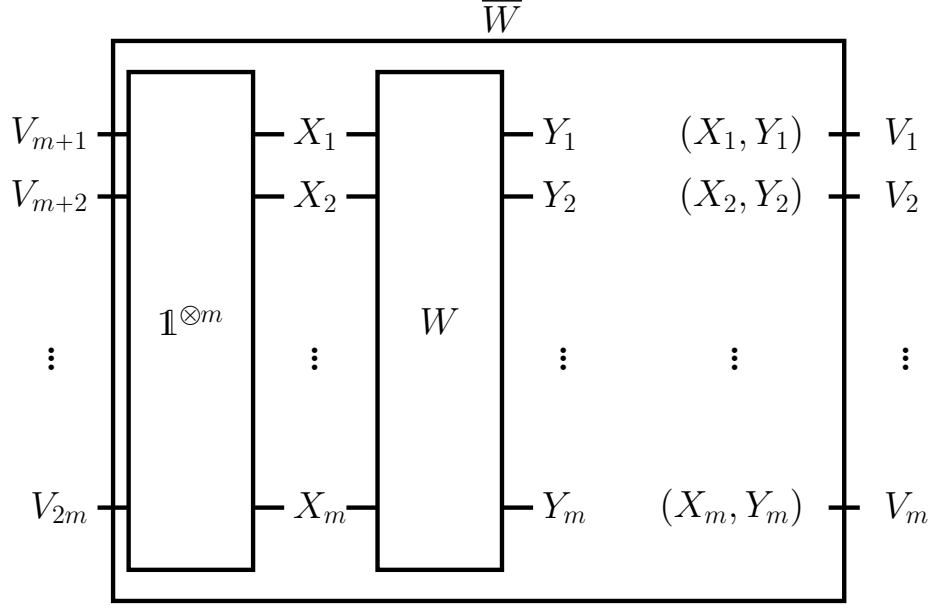
Figure 5.5: The associated multiaccess channel model $\overline{W}$ used in the proof of the upper bound in Theorem 5.4. The set of output terminals is $\mathcal{M} = [m] = \{1, \ldots, m\}$ and the set of input terminals is given by $\mathcal{M}' = \{m+1, \ldots, 2m\}$. The conditional probability distribution of $\overline{W} = P_{V_\mathcal{M}|V_{\mathcal{M}'}}$ is given by $\overline{W} = P_{X_\mathcal{M}|X_{\mathcal{M}'}} \cdot W = \left(\prod_{i \in \mathcal{M}} \mathbb{1}(X_i = X_{i+m}) \right) \cdot W$.

set of $\overline{W}$ and let $\mathcal{M}' = \{m+1, \ldots, 2m\}$ be the set of new input terminals introduced for $\overline{W}$. Thus, $\overline{\mathcal{M}} = \{1, \ldots, m, m+1, \ldots 2m\} = \mathcal{M}' \cup \mathcal{M}$. Input terminals of $\overline{W}$ have RVs that are of the form $V_j = X_j \ \forall j \in \mathcal{M}' = \{m+1, \ldots, 2m\}$, and output terminals of $\overline{W}$ are defined as per the given transceiver model, i.e., their RVs have two components given by $V_j = (X_j, Y_j) \ \forall j \in \mathcal{M} = \{1, \ldots, m\}$. The conditional probability distribution of the multiaccess channel $\overline{W}$ is given by,

$$\overline{W} = P_{V_\mathcal{M}|V_{\mathcal{M}'}} = P_{Y_\mathcal{M} X_\mathcal{M}|X_{\mathcal{M}'}} = P_{X_\mathcal{M}|X_{\mathcal{M}'}} P_{Y_\mathcal{M}|X_\mathcal{M}} = P_{X_\mathcal{M}|X_{\mathcal{M}'}} W,$$

where $P_{X_\mathcal{M}|X_{\mathcal{M}'}}$ is a collection of noiseless DMC's, given by

$$P_{X_\mathcal{M}|X_{\mathcal{M}'}} = \prod_{i \in \mathcal{M}} P_{X_i|X_{i+m}} = \prod_{i \in \mathcal{M}} \mathbb{1}(X_i = X_{i+m}).$$

Note that the RVs of the input terminals in $\overline{W}$ have the special property that $V_j = X_{j-m}$ for all $j \in \mathcal{M}'$. See Figure 5.5.

Theorem 5.4 (General Upper Bounds). *The channel model SK capacity and the channel model PK capacity of any given non-wiretapped transceiver model $W = P_{Y_\mathcal{M}|X_\mathcal{M}}$ for any $\mathcal{D} \subset \mathcal{M} = [m]$, and any $\mathcal{A} \subseteq \mathcal{D}^c$ are upper bounded by*

$$C_{SK}^{\mathcal{A}}(P_{Y_\mathcal{M}|X_\mathcal{M}}) \leqslant C_{SK}^{\mathcal{A}}(P_{V_\mathcal{M}|V_{\mathcal{M}'}}), \tag{5.10}$$

and

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{Y_\mathcal{M}|X_\mathcal{M}}) \leqslant C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{V_\mathcal{M}|V_{\mathcal{M}'}}), \tag{5.11}$$

where the right hand side of the above inequalities are the SK and PK capacities of the associated model $\overline{W}$.

Proof of Theorem 5.4: We show that the PK capacity of W gives an achievable lower bound for $\overline{W}$. Let $K \in \mathcal{K}$ be a private key established for W by SKA protocol Π such that $\log |\mathcal{K}| \leqslant nC_{PK}^{\mathcal{A}|\mathcal{D}}(P_{Y_\mathcal{M}|X_\mathcal{M}})$. We use Π to generate a key $K' \in \mathcal{K}$ in $\overline{W}$. First note that in the associated multiaccess channel model of $\overline{W}$, after each symbol transmission each terminal $j \in \mathcal{M}$ has access to the same variable(s) of the input terminal $j + m \in \mathcal{M}'$. Therefore, terminals of $\mathcal{M}'$ can always remain silent (not sending public messages), and all public messages can be generated by terminals in $\mathcal{M}$, the output terminals of W. Thus, helper terminals of $\mathcal{M}'$ are dummy terminals, and their presence can only help with the key generation. Let Π be such that in each round $t \leqslant n$, terminals generate and send input symbols $X_{jt} = \tilde{X}_{jt}$'s to W and receive corresponding output symbols $Y_{jt} = \tilde{Y}_{jt}$. Then terminals engage in a public discussion $\tilde{\mathbf{F}}^t$. Let Π' be the protocol for SKA in $\overline{W}$ which works as follows. In each round t, input terminals $j + m \in \mathcal{M}'$ generate and send input symbols $V_{(j+m)t} = \tilde{X}_{jt}$'s to $\overline{W}$. Note that every terminal $j \in \mathcal{M}$, receives (as output symbols of $\overline{W}$) the RVs $X_{jt} = \tilde{X}_{jt}$ and $Y_{jt} = \tilde{Y}_{jt}$. Then, input terminals in $\mathcal{M}'$ remain silent and output terminals of multiaccess channel $\overline{W}$ in $\mathcal{M}$ invoke public discussion $\mathbf{F}^t = \tilde{\mathbf{F}}^t$. Following the

same instructions of Π, at the end of round n, terminals in $\mathcal{M}$ can agree on a secret key $K' \in \mathcal{K}$. As, in effect, Π and Π' are identical protocols from the view point of $\mathcal{M}$, K' is equal to K. Therefore, $(1/n) \log |\mathcal{K}|$ is also an achievable key rate for the multiaccess model of $\overline{W}$. The maximum rate of such key is given by the PK capacity of W. The argument for SK capacity is the same with $\mathcal{D} = \varnothing$. ∎

Remark 5.2. We note that using a similar argument, one can prove Theorem 5.4 in the reverse direction – implying (5.10) and (5.11) to be equalities. That is any SKA protocol that achieves the capacity of any given multiaccess model, can be used for establishing a key in any transceiver model. Therefore, one can for example, indirectly prove lower bounds on transceiver model via general lower bounds given for any multiaccess model. To our knowledge the only general lower bound known for multiaccess model is the source emulation lower bound [23, 74], which would imply the same lower bound as in Theorem 5.3 which we easily directly proved in the previous subsection. However, we note that the associated multiaccess model used in Theorem 5.4 has more utility for proving technical upper bounds. This is what we will do in the next two sections. Please note that, Theorem 5.4 in reverse direction does not give lower bounds for transceiver model via special case lower bounds (e.g., the interactive achievability lower bound of [120]) that are proved only for special cases of multiaccess models.

Corollary 5.4.1. *The channel model WSK capacity of any given wiretapped transceiver model $W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$ for any $\mathcal{A} \subseteq \mathcal{M}$ is upper bounded by*

$$C_{WSK}^{\mathcal{A}}(P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant C_{PK}^{\mathcal{A}|\{m+1\}}(P_{Y_{m+1}Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant C_{PK}^{\mathcal{A}|\{m+1\}}(P_{V_{\mathcal{M}}|V_{\mathcal{M}'}}), \qquad (5.12)$$

where $C_{PK}^{\mathcal{A}|\{m+1\}}(P_{Y_{m+1}Y_{\mathcal{M}}|X_{\mathcal{M}}})$ denotes the key capacity of a with $m+1$ terminals where terminal $m+1$ is assumed compromised and models the adversary by $Y_{m+1} = Z$, i.e., Z is assumed to be known. Also $C_{PK}^{\mathcal{A}|\{m+1\}}(P_{V_{\mathcal{M}}|V_{\mathcal{M}'}})$ denotes the PK capacity of the multiaccess model $\overline{W}$ (as per Theorem 5.4,) associated with the aforementioned model where Z is assumed to be

publicly known.

Proof of Corollary 5.4.1: Proof follows directly by the upper bound of (5.11) and Lemma 5.1 of [22] which states that by knowing the adversary's variable Z, the largest achievable SK rate can only increase. ■

We later use Corollary 5.4.1 in Chapter 6 to prove the WSK capacity of a special subclass of transceiver models, namely the wiretapped Polytree-PIN – see Theorem 6.1.

Theorem 5.4 also implies that an upper bound for the SK (or PK) capacity of $\overline{W}$, including the upper bounds of [23, Theorem 6], is an upper bound for $C_{SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}})$ (or $C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}})$). See Corollary below – which immediately follows from Theorem 5.4 and [23, Theorem 6].

Corollary 5.4.2. *Consider a non-wiretapped transceiver model $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ with $\mathcal{D} \subset \mathcal{M}$, and $\mathcal{A} \subseteq \mathcal{D}^c$. Then, for an arbitrary RV X'', and any $\lambda \in \Lambda(\mathcal{A}|\mathcal{D})$ define*

$$g_\lambda(V_{\mathcal{M}}|V_{\mathcal{D}}X'') := H(V_{\mathcal{M}}|V_{\mathcal{D}}X'') - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}X''),$$

and

$$g_\lambda(X_{\mathcal{M}}|V_{\mathcal{D}}X'') := H(X_{\mathcal{M}}|V_{\mathcal{D}}X'') - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D})} \lambda_{\mathcal{B}} H(X_{\mathcal{M}\cap\mathcal{B}}|X_{\mathcal{M}\cap\mathcal{B}^c}V_{\mathcal{D}}X'').$$

Recall that $V_{\mathcal{D}} = constant$ when $\mathcal{D} = \varnothing$. Then

$$C_{SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant \sup_{P_{X''X_{\mathcal{M}}}} \inf_{\lambda \in \Lambda(\mathcal{A})} \left\{ g_\lambda(V_{\mathcal{M}}|X'') - g_\lambda(X_{\mathcal{M}}|X'') \right\}, \tag{5.13}$$

and for any $i \in \mathcal{D}^c$

$$C_{PK}^{\mathcal{A}|\mathcal{D}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant \sup_{P_{X''X_{\mathcal{M}}}} \inf_{\lambda \in \Lambda(\mathcal{A}|\mathcal{D})} \left\{ g_\lambda(V_{\mathcal{M}}|V_{\mathcal{D}}X'') - g_\lambda(X_{\mathcal{M}}|V_{\mathcal{D}}X'') \right.$$
$$\left. + \sum_{\substack{\mathcal{B} \in \Upsilon(\mathcal{A}|\mathcal{D}) \\ s.t. \ i \notin \mathcal{B}}} \lambda_{\mathcal{B}} I(V_{\mathcal{D}}; X_{\mathcal{M}\cap\mathcal{B}}|X_{\mathcal{M}\cap\mathcal{B}^c}X'') \right\}, \tag{5.14}$$

where the variable X'' can be arbitrarily correlated with input RV's $X_{\mathcal{M}}$ without any condition.

5.4   The Non-adaptive SK Capacity

For two-party SKA, the noninteractive one-way SK capacity has been extensively studied in the past; see Chapter 3. Noninteractive SKA protocols are preferred in practice as they are more efficient in terms of public communication cost, and implementation complexity. In this section, we define the non-adaptive SK capacity for the multiterminal channel model of transceivers, and show that the source emulation lower bound of Section 5.3.2 with the noninteractive public communication protocol of [21] achieves this capacity.

Definition 5.4 (Non-adaptive SKA). Consider the following limitations imposed on an SKA channel model:

(a) No Feedback: The protocol starts with symbol transmission over DMC, and after its completion, terminals engage in a (possibly interactive) public discussion phase.

(b) Independent IID Inputs: Terminals are locally controlling their input variables, and the input variables are independent and IID, i.e., $P_{X_{\mathcal{M}}} = \prod_{j \in \mathcal{M}} P_{X_j}$.

SKA protocols satisfying (a) and (b) are called non-adaptive. The non-adaptive secret key capacity, is defined as the largest achievable key rate of all non-adaptive SKAs, and is denoted by $C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}})$.

These are commonly used assumptions that hold in many real-life settings. Next theorem proves that if using public communication in between noisy channel uses is not allowed, the source emulation technique is the optimum SKA approach, as it achieves the non-adaptive secret key capacity.

Theorem 5.5 (Non-adaptive SK Capacity). *For a transceiver model, and a subset $\mathcal{A} \subseteq \mathcal{M}$, the non-adaptive SK capacity is given by*

$$C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}) = \max_{P_{X_{\mathcal{M}}}} C_{SK}^{\mathcal{A}}(P_{X_{\mathcal{M}}} P_{Y_{\mathcal{M}}|X_{\mathcal{M}}). \tag{5.15}$$

Proof of Theorem 5.5:   First, we prove that the right hand side of the above equation

is an upper bound on the non-adaptive capacity (i.e., the converse). Consider an associated multiaccess channel $\overline{W}$ with $2m$ terminals denoted by $\overline{\mathcal{M}} = \{1, \ldots, m, m+1, \ldots 2m\}$, where $V_j = (X_j, Y_j) \ \forall j \in \mathcal{M} = \{1, \ldots, m\}$, are output variables of the multiaccess DMC and $V_j \ \forall j \in \mathcal{M}' = \{m+1, \ldots, 2m\}$ are input variables of DMC, satisfying $P_{X_{\mathcal{M}}|V_{\mathcal{M}'}} = \prod_{j \in [m]} P_{X_j|V_{j+m}}$ and $P_{X_j|V_{j+m}} = \mathbb{1}(X_j = V_{j+m})$. By Theorem 5.4, we have

$$C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant C_{NA-SK}^{\mathcal{A}}(P_{X_{\mathcal{M}}|V_{\mathcal{M}'}}P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}),$$

and thus any upper bound on the non-adaptive SK capacity of multiaccess model $\overline{W}$ is also an upper bound on the non-adaptive SK capacity of the transceiver model W. An upper bound is given for the SK capacity of any multiaccess channel model in [23].

Lemma 5.6 ([23]). *Let $\overline{W}$ be a multiaccess channel, for which $\mathcal{M}'$ is the set of input terminals (transmitters), $\mathcal{M}$ is the set of output terminals (receivers), and $\mathcal{D} = \varnothing$. For any $\mathcal{A} \subseteq \mathcal{M}$ and any $\lambda \in \Lambda(\mathcal{A})$ (as defined in Definition 5.3), any achievable secret key K with alphabet $\mathcal{K}$ satisfies*

$$\frac{1}{n} \log |\mathcal{K}| \leqslant \frac{\alpha_n}{n} E_n + \beta_n, \tag{5.16}$$

where $\alpha_n \to 1$ and $\beta_n \to 0$, as $n \to \infty$; and

$$E_n = \sum_{t=1}^{n} \left[\left(H(V_{\mathcal{M}t}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}t}|V_{\mathcal{B}^c t}) \right) \right.$$
$$\left. - \left(H(V_{\mathcal{M}'t}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{(\mathcal{B} \cap \mathcal{M}')t}|V_{(\mathcal{B}^c \cap \mathcal{M}')t}) \right) \right].$$

The proof of this Lemma is given in [23, Appendix A] (See Equation (A8)). Note that under $n \to \infty$ the right hand side of (5.16) gives an upper bound on the SK capacity of the multiaccess model $\overline{W}$ which, because of Theorem 5.4, implies an upper bound on the SK capacity of W. The general upper bound in (5.16) holds even if assumptions (a) and (b) are not satisfied. However, considering assumptions (a) and (b) we can simplify the expression

of E_n in (5.16). Due to no feedback assumption (a), we have

$$
\begin{aligned}
E_n \;=\; & \Big(H(V_{\mathcal{M}}^n) - \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}^n | V_{\mathcal{B}^c}^n) \Big) \\
& - \Big(H(V_{\mathcal{M}'}^n) - \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}\cap\mathcal{M}'}^n | V_{\mathcal{B}^c\cap\mathcal{M}'}^n) \Big).
\end{aligned}
$$

By the independence of the inputs assumption (b), for any $\mathcal{B}\subseteq\mathcal{M}'$ we have $H(V_{\mathcal{B}}) = \sum_{j\in\mathcal{B}} H(V_j)$, and by properties of λ vectors (see Definition 5.3 and Proposition 5.1) we get

$$
\sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}\cap\mathcal{M}'} | V_{\mathcal{B}^c\cap\mathcal{M}'}) = \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \sum_{j\in\mathcal{B}\cap\mathcal{M}'} H(V_j) = \sum_{j\in\mathcal{M}'} \sum_{\substack{\mathcal{B}\in\Upsilon(\mathcal{A}) \\ \text{s.t. } j\in\mathcal{B}}} \lambda_{\mathcal{B}} H(V_j) = H(V_{\mathcal{M}'}),
$$

and since V_j^n's are IID due to assumptions (a) and (b), we have

$$
E_n = n\Big(H(V_{\mathcal{M}}) - \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}} | V_{\mathcal{B}^c}) \Big).
$$

Lemma 5.6 holds for any λ and any distribution $P_{V_{\mathcal{M}}}$ (that is a function of $P_{X_{\mathcal{M}}}$ as $P_{V_{\mathcal{M}}} = P_{X_{\mathcal{M}}} P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$). Therefore, for every $\lambda \in \Lambda(\mathcal{A})$ the largest upper bound on the non-adaptive SK capacity is given by

$$
C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant \max_{P_{X_{\mathcal{M}}}} \limsup_{n\to\infty} \Big(\frac{\alpha_n}{n} E_n + \beta_n \Big) = \max_{P_{X_{\mathcal{M}}}} g_{\lambda}(V_{\mathcal{M}}),
$$

where we define $g_{\lambda}(V_{\mathcal{M}}) = H(V_{\mathcal{M}}) - \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}} | V_{\mathcal{B}^c})$. As the above inequality holds for every $\lambda \in \Lambda(\mathcal{A})$, the non-adaptive SK capacity is upper bounded by the smallest upper bound as a function of λ (that is a variable independent of how any SKA protocol is executed). Hence,

$$
C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant \min_{\lambda\in\Lambda(\mathcal{A})} \max_{P_{X_{\mathcal{M}}}} g_{\lambda}(V_{\mathcal{M}}).
$$

Function $g_{\lambda}(V_{\mathcal{M}})$ is a concave function of $P_{X_{\mathcal{M}}} = \prod_{j\in\mathcal{M}} P_{X_j}$ and affine as a function of λ

(see Appendix – Section 5.6.) So the minimax theorem [127, Section 5.5] implies that

$$C_{NA-SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) \leqslant \max_{P_{X_{\mathcal{M}}}} \min_{\lambda \in \Lambda(\mathcal{A})} g_{\lambda}(V_{\mathcal{M}}).$$

This completes the proof of the converse. The achievability proof is simple. By Theorem 5.2 we know that for the source model $Q = P_{X_{\mathcal{M}}} P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ and for $\mathcal{D} = \varnothing$ we have

$$C_{SK}^{\mathcal{A}}(Q) = \min_{\lambda \in \Lambda(\mathcal{A})} \left\{ H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}) \right\} = \min_{\lambda \in \Lambda(\mathcal{A})} g_{\lambda}(V_{\mathcal{M}}).$$

Thus, for any $P_{X_{\mathcal{M}}}$ the source emulation SKA protocol of Theorem 5.3 with $X' = \text{constant}$ achieves a key rate lower bounded by

$$\frac{1}{n} \log |\mathcal{K}| \geqslant \min_{\lambda \in \Lambda(\mathcal{A})} g_{\lambda}(V_{\mathcal{M}}) - \xi,$$

for any arbitrary $\xi > 0$. By maximizing $P_{X_{\mathcal{M}}}$ and since ξ can be arbitrarily small the upper bound proved in the converse can be asymptotically achieved. ∎

5.5   Conclusion

We introduced a new general channel model of transceivers for multiterminal secret key agreement and showed that the models in [22] and [23] are special cases of the new model. We gave lower bounds and upper bounds for the SK, PK and WSK capacities of the transceiver model. Then, we studied the problem of non-adaptive secret key agreement and gave the non-adaptive SK capacity of the transceiver model. This result is important because of the ease of implementation of non-adaptive (and noninteractive) protocols. Future research directions include finding tighter bounds for the key capacities of the general transceiver model, and construction of a capacity achieving SKA protocol for wiretapped and non-wiretapped transceiver models.

5.6 Appendix

Lemma 5.7. *Consider $\mathcal{M} = [m]$ and a multi-input multi-output DMC $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$, where $X_{\mathcal{M}} = (X_j|\ j \in \mathcal{M})$, $Y_{\mathcal{M}} = (Y_j|\ j \in \mathcal{M})$, and define $V_{\mathcal{M}} = ((X_j, Y_j)|\ j \in \mathcal{M})$. Let $\mathcal{D} = \varnothing$. For any $\mathcal{A} \subseteq \mathcal{M}$, every fractional partition $\lambda \in \Lambda(\mathcal{A})$, and any given distribution $P_{X_{\mathcal{M}}}$ define $g_\lambda(V_{\mathcal{M}}) = H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c})$. If the probability distribution of $X_{\mathcal{M}}$ satisfies $P_{X_{\mathcal{M}}} = \prod_{j \in \mathcal{M}} P_{X_j}$, then $g_\lambda(V_{\mathcal{M}})$ is a concave function of $P_{X_{\mathcal{M}}}$.*

Proof of Lemma 5.7: Given that $P_{X_{\mathcal{M}}} = \prod_{j \in \mathcal{M}} P_{X_j}$, for any $\mathcal{B} \subseteq \mathcal{M}$ we have $H(X_{\mathcal{M}}) = H(X_{\mathcal{B}}) + H(X_{\mathcal{B}^c})$, and $H(X_{\mathcal{B}}) = \sum_{j \in \mathcal{B}} H(X_j)$. By simple manipulation of the entropic quantities, we can rewrite $g_\lambda(V_{\mathcal{M}})$ as

$$
\begin{aligned}
g_\lambda(V_{\mathcal{M}}) &= H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}|V_{\mathcal{B}^c}) \\
&= H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \Big[H(V_{\mathcal{B}}, V_{\mathcal{B}^c}) - H(V_{\mathcal{B}^c}) \Big] \\
&= H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \Big[H(X_{\mathcal{M}}) + H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - H(X_{\mathcal{B}^c}) - H(Y_{\mathcal{B}^c}|X_{\mathcal{B}^c}) \Big] \\
&= H(V_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \Big[H(X_{\mathcal{B}}) + H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - H(Y_{\mathcal{B}^c}|X_{\mathcal{B}^c}) \Big] \\
&= \Big(H(X_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(X_{\mathcal{B}}) \Big) \\
&\quad + H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \Big[H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - H(Y_{\mathcal{B}^c}|X_{\mathcal{B}^c}) \Big] \\
&= H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} \Big[H(Y_{\mathcal{M}}|X_{\mathcal{M}}) - H(Y_{\mathcal{B}^c}|X_{\mathcal{B}^c}) \Big].
\end{aligned}
$$

The conditional entropy $H(Y_{\mathcal{M}}|X_{\mathcal{M}})$ is an affine function of $P_{X_{\mathcal{M}}}$, and $H(Y_{\mathcal{B}^c}|X_{\mathcal{B}^c})$ is an affine function of $P_{X_{\mathcal{B}^c}}$ and a concave function of $P_{X_{\mathcal{B}}}$. Thus $g_\lambda(V_{\mathcal{M}})$ is a concave function

of $P_{X_\mathcal{M}}$. Note that in the second last equality we used the fact that

$$\sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_\mathcal{B} H(X_\mathcal{B}) = \sum_{\mathcal{B}\in\Upsilon(\mathcal{A})} \lambda_\mathcal{B} \sum_{j\in\mathcal{B}} H(X_j) = \sum_{j\in\mathcal{M}} \sum_{\substack{\mathcal{B}\in\Upsilon(\mathcal{A}) \\ \text{s.t. } j\in\mathcal{B}}} \lambda_\mathcal{B} H(X_j) = \sum_{j\in\mathcal{M}} H(X_j) = H(X_\mathcal{M}),$$

which is due to the independence property that $P_{X_\mathcal{M}} = \prod_{j\in\mathcal{M}} P_{X_j}$. See Proposition 5.1. ∎

The above lemma can be regarded as a generalization of Lemma A.1 in [22].

Chapter 6

Secret Key Capacity of Wiretapped Polytree-PIN

Abstract. In secret key agreement (SKA) in multiterminal channel model, terminals are connected by a noisy discrete memoryless channel (DMC) with multiple input and multiple outputs. Terminals can use the DMC to obtain correlated randomness, and communicate over a noiseless public channel to establish a shared secret key among a designated subset of terminals. We focus on a special class of multiterminal channel models, called wiretapped Polytree-PIN, in which the noisy channel consists of a set of independent point-to-point channels whose underlying undirected connectivity graph forms a tree. We consider a wiretap setting, where the output of each point-to-point channel is partially leaked to a passive wiretapper adversary, Eve, through a second independent noisy channel. A secure SKA protocol generates a group secret key such that Eve has no information about it. In this chapter, we derive the wiretap secret key capacity, which is the largest achievable secret key rate, of the wiretapped Polytree-PIN model. Our result also implies the key capacity of the non-wiretapped Polytree-PIN model, that is the case when there is no leakage from point-to-point channels to Eve.

6.1 Introduction

Two-party secret key agreement (SKA) with information-theoretic security was first intro-
duced and studied in [19], and [20]. The SKA problem was generalized to the case of multiple
terminals in [21, 22] and has been well studied ever since [74]. In the multiterminal channel
models studied in [22, 23, 75, 120] it was assumed that the set of terminals is partitioned
into two disjoint subsets: a subset of terminals that supply input symbols to the underly-
ing discrete memoryless channel (DMC) and a (non-overlapping) subset of terminals that
observe individual outputs of the DMC. The *transceivers channel model* introduced in [36]
(see Chapter 5) generalized previous models to the case of multiple transceivers; where each
terminal is capable of simultaneously sending to, and receiving from, the noisy DMC.

The multiterminal transceivers channel model is defined by an underlying multi-input
multi-output noisy discrete memoryless channel (DMC). There are m terminals denoted by
$\mathcal{M} = \{1, \ldots, m\}$, and a subset of terminals have control over the input variables of the noisy
DMC, and another (possibly overlapping) subset of terminals observe the corresponding
output variables of the DMC. The goal of an SKA protocol is to establish a shared secret
key among a subset of terminals $\mathcal{A} \subseteq \mathcal{M}$. Terminals are allowed to send symbols over the
noisy DMC to generate correlation. This DMC, however, is *wiretapped* and the transmitted
symbols partially leak to a passive wiretapper adversary, Eve. Terminals have access to a
noiseless authenticated public channel which they can use to send public messages (interac-
tively) before, in between, and after symbol transmissions over the noisy DMC. In addition
to the leaked information from the noisy DMC, Eve observes all public messages sent by ter-
minals over the public channel. Utilizing the noisy DMC and the noiseless public channel,
at the end of the SKA protocol, terminals in $\mathcal{A}$ agree on a secret key (SK), such that Eve
has no information about it. The key may or may not be known to the *helper terminals*, i.e.,
terminals in $\mathcal{A}^c = \mathcal{M} \backslash \mathcal{A}$.

The "*key capacity*" of a model is the highest achievable secret key rate of the model
where the rate is defined as the number of key bits that can be established for each symbol

that is transmitted over the DMC. Depending on the Eve's side information, three notions of key capacity have been defined [21, 22]. If Eve has no side information, the capacity is called, the *secret key (SK) capacity*. If a subset of helper terminals $\mathcal{D} \subseteq \mathcal{A}^c$ are compromised by Eve, the key capacity is called the *private key (PK) capacity*. In the most general sense, if Eve has some side information represented by a variable Z, about other terminals' private variables, key capacity is called the *wiretap secret key (WSK) capacity*. In this chapter, we focus on the latter notion of key capacity.

The WSK capacity is known only for few special cases [45, 74], and for the general case, even for two-party SKA ($m = 2$), remains an open problem. The two-party WSK capacity when public communication is one-way (noninteractive messages from input terminal to the output terminal) is shown [20, Theorem 2] to be equal to Wyner's wiretap secrecy capacity [57] of the underlying DMC. This is also called the forward WSK capacity. However, when interaction is allowed, it is proved [20, Theorem 2] that

$$C_{WSK} \leqslant \max_{P_X} I(X; Y|Z),$$

where, X, Y, and Z are, the input variable of terminal 1, output variable of terminal 2, and output variable (side information) of Eve, respectively. When a Markov relation holds between (X, Y, Z) in any order, this bound is tight (gives WSK capacity) and can be achieved with a one-way two-party SKA protocol. The best known upper and lower bounds on multiterminal WSK capacity are due to [75].

6.1.1 Our Work

We study the wiretapped Polytree-PIN model with independent leakage, a special class of transceivers channel model [36], in which terminals are connected by a set of mutually independent point-to-point (directed) channels, and Eve has access to a noisy version of each output variable of each channel. See Figure 6.1 and the description in Section 6.2. For the

case of $m = 2$, the wiretapped Polytree- PIN model is the same as the two-party model of [20, Theorem 2] when the Markov relation $X - Y - Z$ holds. We focus on this special type of wiretap model, which does not fit into other multiterminal channel models [22, 23, 75] and so its key capacity cannot be directly using previously known results in [23, 75, 120]. This special case can however be seen as an instance of the transceivers model [36]. The Polytree-PIN model can also be viewed as the channel model counterpart of the Tree-PIN model of [34], which is a special class of pairwise independent network (PIN) source models [55, 60].

The SK capacity of non-wiretapped Polytree-PIN was derived under the constraints that the input variables are independently generated, and the public communication is noninteractive [36]. In this work, we consider no restriction on the terminals use of the public channel, and prove the WSK capacity of wiretapped Polytree-PIN.

6.1.2 Related Works

Multiterminal SKA problem in channel model has been studied extensively; e.g., see [22, 23, 36, 75, 120]. Single-input multi-output DMC's where studied in [22, 75]. The SK and PK capacities are derived in [22] for the case when terminals are allowed to send public messages without any specific restriction. These capacities are shown to be achievable by an SKA approach called *source emulation*. We also employ this approach and show that source emulation is capacity achieving for the case of wiretapped Polytree-PIN. The WSK capacity of a special class of models (called Markov Tree) was derived in [22, Theorem 5.1] in which Eve gets side information only about the variable associated with *one* of the terminals[1]. The case of multi-input multi-output *(the multiaccess)* channel was studied in [23, 120]. The SK and PK capacities of the multiaccess model are not know, but general upper and

---

[1]We emphasize that while our model of wiretapped Polytree-PIN resembles the wiretapped Markov Tree of [22], they are different in some aspects. Eve in our model is more powerful in the sense that it observes side information about all variables of terminals, but our model is more restrictive as it has the structure of pairwise independent channels.

lower bounds were proved in [23]. The SK capacity is proved for the *symmetric*[2] multiaccess channel with a single output terminal under the constraint that only the output terminal sends public messages [120].

In all these SKA channel models, it was assumed that a terminal is either providing an input to the DMC, or is receiving an output from the DMC. The transceivers model, introduced in Chapter 5 (see also[36]), generalizes the multiaccess model of [23] by allowing terminals to send to, and receive from the DMC. For this model of transceivers, general upper and lower bounds were proved for the SK, PK, and WSK capacities, and the SK capacity was derived when the DMC's input variables are IID and generated independently and non-adaptively.

To the best of our knowledge, except for the result of [22, Theorem 5.1], our result is the only other channel model WSK capacity result for a family of multiterminal models.

6.1.3   Organization

We define the wiretapped Polytree-PIN, and state our main result in Section 6.2. We then prove our result in Section 6.3, and conclude the chapter in Section 6.4.

6.2   Problem Formulation and Main Result

6.2.1   The Model

A Polytree-PIN transceiver model consists of a set of m terminals denoted by $\mathcal{M} = \{1, \ldots, m\}$, and a noisy DMC which is defined by a *polytree* $G = (\mathcal{M}, \mathcal{E})$, that is a directed acyclic graph for which the undirected version is a tree. All terminals have access to a noiseless authenticated public channel as well. In G, each directed edge $e_{ij} \in \mathcal{E}$ is unique and can be represented by an arrow from terminal i to terminal j. When $e_{ij} \in \mathcal{E}$ exists then $e_{ji} \notin \mathcal{E}$, and there are no

---

[2]A two input single out put channel is called symmetric if the conditional distribution of the channel satisfies $P_{V_3|V_1V_2} = P_{V_3|V_2V_1}$.
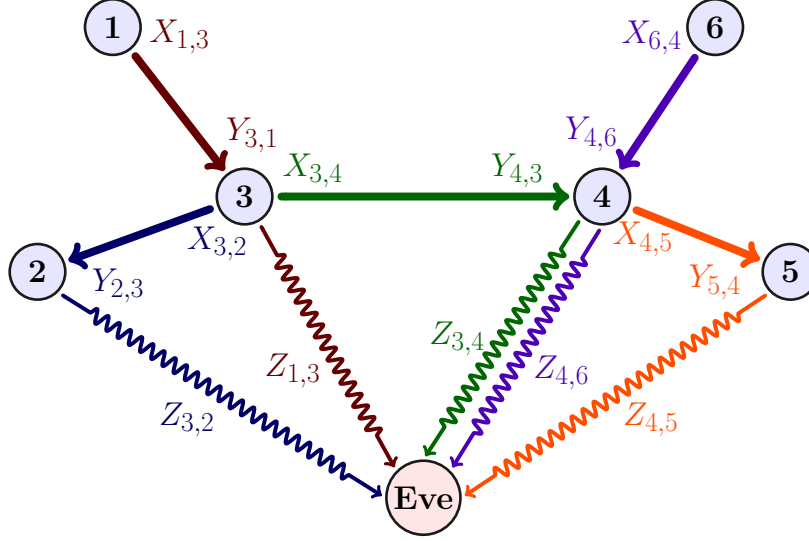
Figure 6.1: An example wiretapped Polytree-PIN. The solid arrows (directed edges) show the independent point-to-point channels of the model, and the curly arrows (with the same color) show the corresponding wiretapping RV's of Eve. With respect to each directed edge $e_{ij} \in \mathcal{E}$ we have $X_{ij} - Y_{ji} - Z_{ij}$, and Eve's side information is $Z = (Z_{ij}|\ e_{ij} \in \mathcal{E})$.

self loops, that is $e_{ii} \notin \mathcal{E}$. Each directed edge $e_{ij} \in \mathcal{E}$ of $G = (\mathcal{M}, \mathcal{E})$ that connects terminal i to terminal j corresponds to an independent point-to-point channel where X_{ij} and Y_{ji} are respectively its input and output RVs. Thus, the RVs of each terminal $i \in \mathcal{M}$ are of the form $V_i = (X_i, Y_i)$, where $X_i = (X_{ij}|\ e_{ij} \in \mathcal{E})$, $Y_i = (Y_{ij}|\ e_{ji} \in \mathcal{E})$. let $X_{\mathcal{M}} = (X_1, \dots, X_m)$, $Y_{\mathcal{M}} = (Y_1, \dots, Y_m)$, and $V_{\mathcal{M}} = (V_1, \dots, V_m)$ be the random vectors representing all input RV's, all output RV's, and all terminals' RV's, respectively. The DMC of a *wiretapped* polytree-PIN is given by $W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}} P_{Z|X_{\mathcal{M}}Y_{\mathcal{M}}}$, with $P_{Y_{\mathcal{M}}|X_{\mathcal{M}}} = \prod_{e_{ij} \in \mathcal{E}} P_{Y_{ji}|X_{ij}}$ where $P_{Y_{ji}|X_{ij}}$ corresponds to the point-to-point channel between X_{ij} and Y_{ji}.

A polytree-PIN is called with *independent leakage* if the wiretapper's RV is of the form $Z = (Z_{ij}|\ e_{ij} \in \mathcal{E})^3$, where the Markov relation $X_{ij} - Y_{ji} - Z_{ij}$ holds for all $e_{ij} \in \mathcal{E}$. For this case, the DMC is given by

$$W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} = \prod_{i \in \mathcal{M}} \prod_{\substack{j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}}} P_{Y_{ji}|X_{ij}} P_{Z_{ij}|Y_{ji}},$$

---

[3] With respect to some fixed arbitrary order over the edges in $\mathcal{E}$.

where $P_{Z_{ij}|Y_{ji}}$ represents the point-to-point channel between Y_{ji} and Z_{ij} (Eve). See Figure 6.1. When $Z = $ constant, that is when Eve has no side information, the model is called *non-wiretapped*, and is denoted by $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$.

The goal of an SKA protocol is for terminals in $\mathcal{A} \subseteq \mathcal{M}$ to share a secret key K by using cooperation of terminals in $\mathcal{M}$. All terminals have access to a public channel, and public messages sent by a terminals will be received by all terminals and the passive adversary Eve, who will not interfere with the public communication. There is no initial correlation between the terminals. Before starting the SKA protocol, terminals are allowed to use the public channel for initialization (e.g., agreeing on public parameters). An SKA protocol consists of n rounds, where each round consists of one invocation of the noisy channel, followed by public communication by terminals in $\mathcal{M}$ over the public channel. Let $\mathbf{F}^t$ denote the random variable representing all public messages of the m terminals in round $1 \leqslant t \leqslant n$, and let $\mathbf{F} = (\mathbf{F}^1, \ldots, \mathbf{F}^n)$ denote the entire public communication. Each public message of terminal j in round $1 \leqslant t \leqslant n$ is a function of all previous samples $V_{j1}, \ldots, V_{jt}$, its local randomness, public messages of the previous rounds $\mathbf{F}^1, \ldots, \mathbf{F}^{t-1}$, and previous public message sent in round t. Each input symbol X_{jt} of round $t \geqslant 2$ may depend on previous public discussions $\mathbf{F}^1, \ldots, \mathbf{F}^{t-1}$, and previous samples $V_{j1}, \ldots, V_{j(t-1)}$. After the n rounds of the SKA protocol, the RV associated to each terminal j is given by $V_j^n = (X_j^n, Y_j^n)$. Also, let $V_{\mathcal{M}}^n$ denote the collection of all RVs accessible to all terminals after round n. The protocol ends when each terminal j computes their version of the key $K_j(V_j^n, \mathbf{F})$ that is a function of all the symbols they send to, or receive from, the noisy DMC (V_j^n), and all of the exchanged public messages $(\mathbf{F})$. Eve has access to all public messages, $\mathbf{F}$, and the side information Z^n, which is correlated with $V_{\mathcal{M}}^n$.

6.2.2 Definitions

Definition 6.1. For a set of terminals $\mathcal{M}$, let $\mathcal{A} \subseteq \mathcal{M}$ denote the subset of terminals that want to obtain a shared key K with alphabet $\mathcal{K}$. Let Z^n denote Eve's side information about

$V_{\mathcal{M}}^n$. The key K is an (ϵ, σ)-Secret Key (in short (ϵ, σ)-SK) for $\mathcal{A}$, if there exists an SKA protocol with public communication $\mathbf{F}$, and output RVs $\{K_j\}_{j \in \mathcal{A}}$ for each terminal, such that

$$\text{(reliability)} \quad \Pr\{K_j = K\} \geqslant 1 - \epsilon, \quad \forall j \in \mathcal{A}, \tag{6.1}$$

$$\text{(secrecy)} \quad \mathbf{SD}\left((K, \mathbf{F}, Z^n), (U, \mathbf{F}, Z^n)\right) \leqslant \sigma, \tag{6.2}$$

where $\mathbf{SD}$ denotes the statistical distance and U is the uniform probability distribution over alphabet $\mathcal{K}$.

Definition 6.2 (Key Capacity – see Definition 17.16 of [90]). Let Z^n denote Eve's side information about $V_{\mathcal{M}}^n$. For a given channel model W, where W is the conditional distribution of the underlying DMC, a real number $R \geqslant 0$ is an achievable SK rate if there exists an SKA protocol that for every n establishes an (ϵ_n, σ_n)−SK $K \in \mathcal{K}$ where $\lim_{n \to \infty} \epsilon_n = 0$, $\lim_{n \to \infty} \sigma_n = 0$, and $\liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{K}| = R$. The maximum of all achievable SK rates is called the key capacity of the given model W.

SK, PK, and WSK Capacities. Eve has access to all public messages, denoted by $\mathbf{F}$, and might wiretap some side information about $V_{\mathcal{M}}^n$, denoted by RV Z^n. When Eve has no side information about $V_{\mathcal{M}}^n$, then $Z^n = $ constant (i.e., independent of $V_{\mathcal{M}}^n$), and the capacity is called *SK capacity*, denoted by $C_{SK}^{\mathcal{A}}(W)$. Eve may compromise a subset of terminals $\mathcal{D} \subset \mathcal{A}^c$, in which case $Z^n = V_{\mathcal{D}}^n = (V_j^n | \, \forall j \in \mathcal{D})$. The compromised terminals remain cooperative in the SKA protocol (it is assumed that they reveal $V_{\mathcal{D}}^n$ to other terminals through the public channel.) The capacity for this case is called *PK capacity* and is denoted by $C_{PK}^{\mathcal{A}|\mathcal{D}}(W)$. In the most general sense, if Eve has access to side information Z^n, which is correlated with $V_{\mathcal{M}}^n$, the model is called "wiretapped", and the key capacity is called *WSK capacity*, denoted by $C_{WSK}^{\mathcal{A}}(W)$.

6.2.3 WSK Capacity of Polytree-PIN

The main contribution of this chapter is deriving the WSK capacity of wiretapped Polytree-PIN with independent leakage. Here, we state the claimed result and give the proof in the next section.

Theorem 6.1 (WSK capacity of Polytree-PIN). *The WSK capacity of a wiretapped Polytree-PIN with independent leakage defined by $G = (\mathcal{M}, \mathcal{E})$ is*

$$C_{WSK}^{\mathcal{A}}(P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}) = \max_{P_{X_{\mathcal{M}}}} \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(X_{ij}; Y_{ji}|Z_{ij}), \tag{6.3}$$

where $G_{\mathcal{A}} = (\mathcal{M}_{\mathcal{A}}, \mathcal{E}_{\mathcal{A}})$ is the subgraph of G with the smallest number of edges that spans all terminals of $\mathcal{A}$. Moreover, this key capacity is achievable by the simple source emulation approach.

The above theorem also implies the SK capacity of the non-wiretapped Polytree-PIN model, with the choice of $Z = $ constant. That is,

$$C_{SK}^{\mathcal{A}}(P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}) = \max_{P_{X_{\mathcal{M}}}} \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(X_{ij}; Y_{ji}), \tag{6.4}$$

and that this SK capacity is achievable by the simple source emulation approach.

Note than the implied non-adaptive SK capacity formulation of Theorem 5.5 for the case of non-wiretapped Polytree-PIN is equal to the right hand side of Equation (6.4); however, the statement of Equation (6.4) is stronger than the claim of Theorem 5.5. While in Equation (6.4) the model is *not* restricted to non-adaptive SKA, we prove that the non-adaptive SKA approach of source emulation *is capacity achieving* for the case of wiretapped (and non-wiretapped) Polytree-PIN.

6.3 Proof of Theorem 6.1

In this section, we provide the proof for Theorem 6.1. The proof has two parts. In the direct part, to prove a lower bound on $C_{WSK}^{\mathcal{A}}(W)$, we use the source emulation approach of [22] and the source model Tree-PIN SKA protocol of [34]. In the converse part, we prove an upper bound on $C_{WSK}^{\mathcal{A}}(W)$ using a combination of techniques from [23] and [36]. The novelty of the upper bound proof lies in exploiting the induced Markov relations among all variables (within $V_{\mathcal{M}}^n, \mathbf{F}$), without imposing additional limitations on the model, to show that the upper bound is tight in general for Polytree-PIN, and is achievable by the source emulation lower bound. Theorem 6.1 also implies the SK capacity by letting $Z = $ constant. We first review the lemmas that are used in the converse, and then we review the source emulation technique.

6.3.1 Converse Techniques

The following lemmas will be used for the converse part of our proof.

Lemma 6.2 (Lemma 5.1 of [22]). *For a wiretapped channel model W, where Eve's RV is not constant, let $C_{PK}^{\mathcal{A}|\{|\mathcal{M}|+1\}}(\widetilde{W})$ be the PK capacity of an associated model $\widetilde{W}$ that is the same as W except that Eve is assumed to be a new compromised terminal – that is $\widetilde{\mathcal{M}} = \mathcal{M} \cup \{|\mathcal{M}| + 1\}$, and $V_{|\mathcal{M}|+1} = Z$. By definition of the PK and WSK capacities we have, $C_{WSK}^{\mathcal{A}}(W) \leqslant C_{PK}^{\mathcal{A}|\{|\mathcal{M}|+1\}}(\widetilde{W})$.*

The multiaccess model of [23], is a special case of transceivers model, in which a subset of terminals only have access to the input RVs and the rest of the terminals only have access to the output RVs of the DMC.

Lemma 6.3 (Theorem 5.4 of Chapter 5, see also Theorem 3 of [36]). *Consider a non-wiretapped transceivers model $W = P_{Y_{\mathcal{M}}|X_{\mathcal{M}}}$ where $\mathcal{M} = \{1, \ldots, m\}$. Define the associated multiaccess model $\overline{W}$ over the terminal set $\overline{\mathcal{M}} = \mathcal{M}' \cup \mathcal{M}$, where $\mathcal{M}' = \{m + 1, \ldots, 2m\}$ is a new subset of terminals that only provide input symbols $V_i = \overline{X}_i \ \forall i \in \mathcal{M}'$*

to the underlying multiaccess DMC of $\overline{W}$. The original terminal set of W (i.e., $\mathcal{M}$) is now the (nonoverlapping) subset of terminals that observe output RV's $\overline{Y}_j$ $\forall j \in \mathcal{M}$, composed of two components $V_j = \overline{Y}_j = (X_j, Y_j)$. The multiaccess DMC of $\overline{W} = P_{V_\mathcal{M}|V_{\mathcal{M}'}} = P_{\overline{Y}_\mathcal{M}|\overline{X}_{\mathcal{M}'}}$ is defined based on the DMC of W as

$$\overline{W} = P_{X_\mathcal{M}|\overline{X}_{\mathcal{M}'}} P_{Y_\mathcal{M}|X_\mathcal{M}} = \big(\prod_{i \in \mathcal{M}} \mathbb{1}(X_i = \overline{X}_{i+m}) \big) W,$$

where, for any $i \in \mathcal{M}$, the connection between input symbols $\overline{X}_{i+m}$ and X_i component of output RV's is given by noiseless DMC's $\mathbb{1}(X_i = \overline{X}_{i+m})$. Then, for any $\mathcal{D} \subset \mathcal{M}$ and $\mathcal{A} \subseteq \mathcal{D}^c$ we have $C_{PK}^{\mathcal{A}|\mathcal{D}}(W) \leqslant C_{PK}^{\mathcal{A}|\mathcal{D}}(\overline{W})$.

6.3.2 Source Emulation

For the achievability, we use the source emulation approach [22, 23, 75]. In the multiterminal source model [21], there are m terminals each having access to n IID copies of V_j ($\forall j \leqslant m$) that are used for SKA by using public communication among terminals. The SK, PK, and WSK capacities have been defined for source model in [21] (similar to Definition 6.2).

A source model SKA protocol, gives a lower (achievability) bound on the channel model, in the following way. Let terminals of a given channel model $W = P_{ZY_\mathcal{M}|X_\mathcal{M}}$, use the underlying DMC n times, with IID input symbols, and without using feedback over the public channel. The variables that will be held by each terminal at the end of this symbol transmission define a source model that is described by $P_{ZV_\mathcal{M}} = P_{X_\mathcal{M}}W$. This is called source emulation [22]. A secure source model protocol will give a protocol for the channel model by first invoking the symbol transmissions that emulate the source, and then directly using the source model protocol, which immediately implies a channel model lower bound.

Note that source emulation may not achieve the key capacity, since it does not utilize the available public channel between DMC applications, which can be used for example to provide feedback and adaptation of the channel input variables. Theorem 4 of [120]
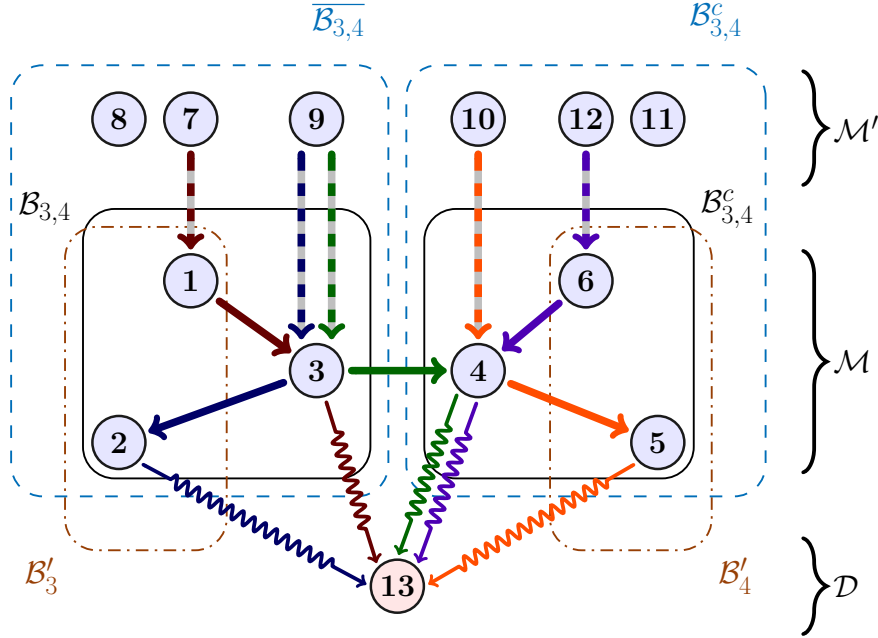
Figure 6.2: An example of Polytree-PIN transceiver model and its associated multiaccess channel model as per the proof of Theorem 6.1. The dashed arrows show the noiseless channels of the multiaccess model, connecting terminals of $\mathcal{M}'$ to the output terminals in $\mathcal{M}$. There is one noiseless DMC per each input RV of the original Polytree-PIN.

shows that source emulation is not capacity achieving for certain channel models. We show, however, that though Polytree-PIN model allows adaptive channel inputs, such SKA method is not necessary for achieving the WSK capacity, and source emulation is sufficient.

6.3.3 The Proof

Proof of Theorem 6.1: For the converse part, we start with Lemma 6.2, and Lemma 6.3. For wiretapped Polytree-PIN $W = P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$ with polytree $G = (\mathcal{M}, \mathcal{E})$ where $\mathcal{M} = \{1, \ldots, m\}$, we prove an upper bound on $C_{WSK}^{\mathcal{A}}(W)$. Define the associated transceiver model $\widetilde{W}$ by considering Eve as a new compromised terminal labeled as $2m + 1$ added to the terminal set; that is, $\mathcal{D} = \{2m + 1\}$, $\widetilde{\mathcal{M}} = \mathcal{M} \cup \mathcal{D}$, $V_{\mathcal{D}} = Y_{\mathcal{D}} = Z = (Z_{ij}|\ e_{ij} \in \mathcal{E})$, and $\widetilde{W} = P_{Y_{\mathcal{D}}Y_{\mathcal{M}}|X_{\mathcal{M}}} = W$. By Lemma 6.2, $C_{WSK}^{\mathcal{A}}(W) \leqslant C_{PK}^{\mathcal{A}|\mathcal{D}}(\widetilde{W})$. Then, we define a multiaccess model $\overline{W}$ by considering the $m + 1$ terminals of $\widetilde{W}$ as its output terminals, and introducing

m new input terminals denoted by $\mathcal{M}' = \{m+1, \ldots, 2m\}$ – i.e., $\overline{\mathcal{M}} = \mathcal{M}' \cup \widetilde{\mathcal{M}}$. The input terminals' RVs of $\overline{W}$ are $\overline{X}_i = (X_{ij}| e_{(i-m)j} \in \mathcal{E}) \; \forall i \in \mathcal{M}'$, and the non-compromised output terminals' RVs are of the form $\overline{Y}_i = (X_i, Y_i) \; \forall i \in \mathcal{M}$, with $X_i = (X_{ij}| e_{ij} \in \mathcal{E})$, and $Y_j = (Y_{ji}| e_{ij} \in \mathcal{E})$. The only compromised terminal $2m+1$ is an output terminal with RV $\overline{Y}_{2m+1} = \overline{Y}_{\mathcal{D}} = Z$. The conditional probability distribution of the multiaccess channel $\overline{W} = P_{\overline{Y}_{\widetilde{\mathcal{M}}}|\overline{X}_{\mathcal{M}'}}$ is given by,

$$P_{\overline{Y}_{\mathcal{D}}\overline{Y}_{\mathcal{M}}|\overline{X}_{\mathcal{M}'}} = P_{ZY_{\mathcal{M}}X_{\mathcal{M}}|\overline{X}_{\mathcal{M}'}} = P_{X_{\mathcal{M}}|\overline{X}_{\mathcal{M}'}} P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$$

where $P_{X_{\mathcal{M}}|\overline{X}_{\mathcal{M}'}}$ is a collection of noiseless DMC's, given by

$$\prod_{i \in \mathcal{M}} P_{X_i|\overline{X}_{i+m}} = \prod_{i \in \mathcal{M}} \prod_{\substack{j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}}} \mathbb{1}(X_{ij} = \overline{X}_{(i+m)j}),$$

and $P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}} = W$. By Lemma 6.2 and Lemma 6.3 we have

$$C^{\mathcal{A}}_{WSK}(W) \leq C^{\mathcal{A}|\mathcal{D}}_{PK}(\widetilde{W}) \leq C^{\mathcal{A}|\mathcal{D}}_{PK}(\overline{W}).$$

Next, we prove an upper bound for the PK capacity of $\overline{W}$. For any edge $e_{ij} \in \mathcal{E}_{\mathcal{A}}$, define $\mathcal{P}_{ij} = \{\mathcal{B}_{ij}, \mathcal{B}^c_{ij}\}$ to be the bi-partition (the cut) of $\mathcal{M}$ that $e_{ij} \in \mathcal{E}_{\mathcal{A}}$ crosses (i.e., either $i \in B_{ij}$ and $j \in B^c_{ij}$, or $j \in B_{ij}$ and $i \in B^c_{ij}$). Also, define $\overline{\mathcal{P}_{ij}} = \{\overline{\mathcal{B}_{ij}}, \overline{\mathcal{B}^c_{ij}}\}$ to be the partition of $\widehat{\mathcal{M}} = \mathcal{M} \cup \mathcal{M}'$ such that e_{ij} crosses $\overline{\mathcal{P}_{ij}}$, and for every $j \leq m$ in one part of $\overline{\mathcal{P}_{ij}}$, $j + m$ belongs to the same part. Note that $\overline{\mathcal{B}_{ij}} \cap \mathcal{M} = \mathcal{B}_{ij}$ and $\overline{\mathcal{B}^c_{ij}} \cap \mathcal{M} = \mathcal{B}^c_{ij}$. See the example partitions in Figure 6.2.

Let $K \in \mathcal{K}$ be an achievable $(\epsilon, \sigma)-$SK for $\overline{W}$, achieved by an SKA protocol Π with public communication $\mathbf{F}$. We prove that for any achievable $(\epsilon, \sigma)-$SK for $\overline{W}$ we have

$$\log|\mathcal{K}| \leq I(V^n_{\mathcal{B}_{ij}}; V^n_{\mathcal{B}^c_{ij}}|Z^n\mathbf{F}) + \delta, \; \forall e_{ij} \in \mathcal{E}_{\mathcal{A}} \tag{6.5}$$

where $\delta = \delta(\epsilon, \sigma) = \sigma \log \frac{|\mathcal{K}|}{\sigma} + (|\widehat{\mathcal{M}}| + 2)(\epsilon \log |\mathcal{K}| + h(\epsilon))$.

Before proving Equation (6.5), we review some notations. For a subset $\mathcal{A} \subseteq \widehat{\mathcal{M}}$, let $\Upsilon(\mathcal{A})$ be the family of all nonempty sets $\mathcal{B} \subset \widehat{\mathcal{M}}$ such that, $\mathcal{B}$ does not contain $\mathcal{A}$ ($\mathcal{A} \nsubseteq \mathcal{B}$), and let $\Lambda(\mathcal{A})$ be the set of all $|\Upsilon(\mathcal{A})|$-dimensional vectors $\lambda = (\lambda_{\mathcal{B}} \mid \mathcal{B} \in \Upsilon(\mathcal{A}))$ such that $0 \leqslant \lambda_{\mathcal{B}} \leqslant 1$, and for any terminal $j \in \mathcal{D}^c$, all $\lambda \in \lambda(\mathcal{A})$ satisfy $\sum_{\mathcal{B} \in \Upsilon(\mathcal{A}) \text{ s.t. } j \in \mathcal{B}} \lambda_{\mathcal{B}} = 1$. See Definition 5.3. As K is an (ϵ, σ)−SK achieved by $\mathbf{F}$, Equations (6.1) and (6.2) are satisfied. Since $\overline{W}$ is a multiaccess model, [22, Lemma A.2] implies that for every $\lambda \in \Lambda(\mathcal{A})$ we have

$$H(K|\mathbf{F}Z^n) \leqslant H(V_{\widehat{\mathcal{M}}}^n|\mathbf{F}Z^n) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}^n|V_{\mathcal{B}^c}^n \mathbf{F}Z^n) + \nu,$$

where $\nu = (|\widehat{\mathcal{M}}| + 2)(\epsilon \log |\mathcal{K}| + h(\epsilon))$. Also, we define $s = s(K, \mathbf{F}, Z^n) = \log |\mathcal{K}| - H(K|\mathbf{F}, Z^n)$. Then

$$\log |\mathcal{K}| \leqslant H(V_{\widehat{\mathcal{M}}}^n|\mathbf{F}Z^n) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}^n|V_{\mathcal{B}^c}^n \mathbf{F}Z^n) + s + \nu \quad \text{(a)}$$

$$\leqslant H(V_{\widehat{\mathcal{M}}}^n|\mathbf{F}Z^n) - \sum_{\mathcal{B} \in \Upsilon(\mathcal{A})} \lambda_{\mathcal{B}} H(V_{\mathcal{B}}^n|V_{\mathcal{B}^c}^n \mathbf{F}Z^n) + \delta \quad \text{(b)}$$

$$\leqslant I(V_{\mathcal{B}_{ij}}^n; V_{\mathcal{B}_{ij}^c}^n|Z^n \mathbf{F}) + \delta \quad \text{(c)}$$

where (a) is due to definition of $s = s(K, \mathbf{F}, Z^n)$, (b) is due to Lemma 1 of [21], i.e., $s(K, \mathbf{F}, Z^n) = s \leqslant \sigma \log \frac{|\mathcal{K}|}{\sigma}$, (c) is due to choosing $\lambda = (\lambda_{\mathcal{G}}, \mathcal{G} \in \Upsilon(\mathcal{A}))$ as

$$\lambda_{\mathcal{G}} = \begin{cases} 1 & \mathcal{G} = \overline{\mathcal{B}_{ij}} \text{ or } \mathcal{G} = \overline{\mathcal{B}_{ij}^c} \\ 0 & \text{otherwise,} \end{cases}$$

and that we have $\overline{X}_{j+m} = X_j \; \forall j \leqslant m$. This proves Equation (6.5). Next, let $\mathcal{P}_{ij} = \{\mathcal{B}_{ij}, \mathcal{B}_{ij}^c\}$ be the bi-partition of $\mathcal{M}$ such that $i \in B_{ij}$ and $j \in B_{ij}^c$, and that the directed edge e_{ij} is from terminal i to j (i.e., $X_{ij} - Y_{ji} - Z_{ij}$). Define $\mathcal{B}_i' = \mathcal{B}_{ij} \backslash \{i\}$, $\mathcal{B}_j' = \mathcal{B}_{ij}^c \backslash \{j\}$, and also let $\tilde{V}_i = ((X_{ik}|k \neq j), Y_i)$ and $\tilde{V}_j = (X_j, (X_{jk}|k \neq i))$. Note that neither $\mathcal{B}_{ij}$ nor B_{ij}^c contain $\mathcal{A}$

as $e_{ij} \in \mathcal{E}_{\mathcal{A}}$. Thus,

$$\log|\mathcal{K}| \leqslant I(V_{\mathcal{B}_{ij}}^n; V_{\mathcal{B}_{ij}^c}^n | Z^n \mathbf{F}) + \delta$$

$$= I(V_i^n; V_{\mathcal{B}_{ij}^c}^n | Z^n \mathbf{F}) + \underline{I(V_{\mathcal{B}_i'}^n; V_{\mathcal{B}_{ij}^c}^n | Z^n \mathbf{F} V_i^n)} + \delta \qquad \text{(a)}$$

$$= I(V_i^n; Y_{ji}^n | Z^n \mathbf{F}) + \underline{I(V_i^n; \tilde{V}_j^n V_{\mathcal{B}_j'}^n | Z^n \mathbf{F} Y_{ji}^n)} + \delta \qquad \text{(b)}$$

$$= I(V_i^n; Y_{ji,n} | Z^n \mathbf{F} Y_{ji}^{n-1}) + I(V_i^n; Y_{ji}^{n-1} | Z^n \mathbf{F}) + \delta \qquad \text{(c)}$$

$$\leqslant I(V_i^n \mathbf{F} (ZY_{ji})^{n-1}; Y_{ji,n} | Z_{ij,n}) +$$

$$\quad I(V_i^{n-1}; Y_{ji}^{n-1} | Z^n \mathbf{F}) + \underline{I(V_{i,n}; Y_{ji}^{n-1} | Z^n \mathbf{F} V_i^{n-1})} + \delta \quad \text{(d)}$$

$$\leqslant I(X_{ij,n}; Y_{ji,n} | Z_{ij,n}) + I(V_i^{n-1}; Y_{ji}^{n-1} | Z^{n-1} \mathbf{F}) +$$

$$\quad \underline{I(\tilde{V}_i^n; Y_{ji,n} | Z^n \mathbf{F} (V_i Y_{ji})^{n-1} X_{ij,n})} + \delta \qquad \text{(e)}$$

$$\leqslant \sum_{t=1}^{n} I(X_{ij,t}; Y_{ji,t} | Z_{ij,t}) + \delta$$

where in (a) the second term cancels to zero as Markov relation $V_{\mathcal{B}_i'}^n - Z^n \mathbf{F} V_i^n - V_{\mathcal{B}_{ij}^c}^n$ holds, (b) is by $V_i^n - Z^n \mathbf{F} Y_{ji}^n - \tilde{V}_j^n V_{\mathcal{B}_j'}^n$, (c) follows from entropy chain rule, (d) holds since conditioning reduces entropy, and that $Z_n V_{i,n} - (ZV_i)^{n-1} \mathbf{F} - Y_{ji}^{n-1}$ holds, and (e) is due to $\tilde{V}_i^n - \mathbf{F}(ZV_i)^{n-1} - X_{ij,n} - Y_{ji,n}$. Recall that $I(J_1; J_3 | J_2) = 0$ if $J_1 - J_2 - J_3$. Also, remember we assume that $\mathbf{F}$ contains Z^n when considering the PK capacity. The above inequality holds for any $e_{ij} \in \mathcal{E}_{\mathcal{A}}$, thus

$$C_{WSK}^{\mathcal{A}}(W) \leqslant C_{PK}^{\mathcal{A}|\mathcal{D}}(\overline{W}) \leqslant \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} \limsup_{n \to \infty} \max_{P_{X_{\mathcal{M}}^n}} \frac{1}{n} \log|\mathcal{K}|$$

$$\leqslant \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} \max_{P_{X_{\mathcal{M}}}} I(X_{ij}; Y_{ji} | Z_{ij})$$

$$= \max_{P_{X_{\mathcal{M}}}} \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} I(X_{ij}; Y_{ji} | Z_{ij}), \qquad (6.6)$$

and the last equality is due to minimax inequality [127, Section 5.5] and that there exists a distribution of the form $P_{X_{\mathcal{M}}} = \prod_{e_{ij} \in \mathcal{E}} P_{X_{ij}}$ that maximizes $I(X_{ij}; Y_{ij} | Z_{ij})$ for all e_{ij}. See Lemma 6.4 of Appendix (Section 6.5). Thus, the maximization can be reduced to maximum

206

over independent IID inputs.

To prove achievability, we prove that simple source emulation approach of Equation (5.8) with $X' = \text{constant}$ (see also [22, Section IV] and [23, Section IV]) achieves the upper bound in (6.6). Each terminal with control over input symbols X_{ij}, independently generates X_{ij}^n according to the distribution $P_{X_{\mathcal{M}}} = \prod_{e_{ij} \in \mathcal{E}} P_{X_{ij}}$. Then by sending input symbols through the DMC, terminals receive the output RVs $Y_{\mathcal{M}}^n$. Thus, n IID copies of a source model $P_{ZV_{\mathcal{M}}} = P_{X_{\mathcal{M}}} P_{ZY_{\mathcal{M}}|X_{\mathcal{M}}}$ is generated amongst the terminals. Note that due to properties of Polytree-PIN transceiver model (i.e., mutually independent point-to-point channels, and $X_{ij} - Y_{ji} - Z_{ij}$ for all $e_{ij} \in \mathcal{E}$), this source model is a Tree-PIN source model[4] and for any arbitrary $\xi > 0$, and any distribution $P_{X_{\mathcal{M}}} = \prod_{e_{ij} \in \mathcal{E}} P_{X_{ij}}$, one can employ the source model SKA protocol 6 of Chapter 4 (see also [34]) to establish a secret key of rate of

$$\frac{1}{n} \log |\mathcal{K}| \geqslant \min_{\substack{i,j \in \mathcal{M} \\ \text{s.t. } e_{ij} \in \mathcal{E}_{\mathcal{A}}}} \{I(X_{ij}; Y_{ij}|Z_{ij})\} - \xi.$$

The protocol works as follows. Corresponding to each e_{ij}, terminals i and j use their RV's X_{ij}^n and Y_{ji}^n to agree on pairwise secret key S_{ij} (using a two-party SKA, e.g., Protocol 4 or 5– see also [32, 79]), where $\text{length}(S_{ij}) = nI(X_{ij}; Y_{ij}|Z_{ij}) - o(n)$. Then, terminals run a public communication protocol that enables all terminals in $\mathcal{M}_{\mathcal{A}} \subseteq \mathcal{D}^c$ to securely agree on one of the pairwise keys as their final key, say for example $K = S_{ij}$ with $j \in \mathcal{M}_{\mathcal{A}}$ and $e_{ij} \in \mathcal{E}_{\mathcal{A}}$. The proof is in the Appendix of Chapter 4. Since ξ was arbitrary, by maximizing over $P_{X_{\mathcal{M}}}$ we obtain the upper bound of Equation (6.6) and hence the capacity. ∎

6.4 Conclusion

Secret key agreement protocols with security against wiretapping adversaries are important in practice as they naturally model leakage of communication to eavesdropping adversaries in

---

[4]A source model described by $P_{ZV_{\mathcal{M}}}$ is a wiretapped Tree-PIN if there exists an undirected tree $G = (\mathcal{M}, \mathcal{E})$ such that $\{(V_{ij}, V_{ji}, Z_{ij})\}_{e_{ij} \in \mathcal{E}}$ are mutually independent, and $V_{ij} - V_{ji} - Z_{ij}$ holds for all $e_{ij} \in \mathcal{E}$, where $V_j = (V_{ji}| e_{ij} \in \mathcal{E})$ is terminal j's RV.

wireless settings. Finding WSK capacity in general is an open problem [45, 74]. We studied the special case of wiretapped Polytree-PIN with independent leakage, and derived its WSK capacity. Our results also proved that, for this model, capacity can be achieved by using the source emulation approach which can be implemented in practice. Finding secret key capacity of other wiretapped transceiver models is an interesting direction for future work. One possible generalization of Polytree-PIN transceiver model is the case of independent point-to-point channels characterized by general directed graphs where the directed graph is not a polytree (contains loops).

6.5   Appendix

Lemma 6.4. *For any given Polytree-PIN $W = P_{ZY_\mathcal{M}|X_\mathcal{M}}$ with independent leakages, define* $R_{ij}(P_{X_\mathcal{M}}) = I(X_{ij}; Y_{ji}|Z_{ij})$. *Then, we have*

$$\max_{P_{X_\mathcal{M}}} \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} R_{ij}(P_{X_\mathcal{M}}) = \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} \max_{P_{X_\mathcal{M}}} R_{ij}(P_{X_\mathcal{M}}).$$

Proof: By minimax inequality [127, Section 5.5] we have

$$\max_{P_{X_\mathcal{M}}} \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} R_{ij}(P_{X_\mathcal{M}}) \leqslant \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} \max_{P_{X_\mathcal{M}}} R_{ij}(P_{X_\mathcal{M}}).$$

Then we note that, for Polytree-PIN W in which point-to-point channels are independent, there exists a joint probability distribution of the form $P^*_{X_\mathcal{M}} = \prod_{e_{ij}\in\mathcal{E}} P^*_{X_{ij}}$ that maximizes $I(X_{ij}; Y_{ij}|Z_{ij})$ for all e_{ij}. That is for any e_{ij}

$$R^*_{ij} := \max_{P_{X_\mathcal{M}}} R_{ij}(P_{X_\mathcal{M}}) = \max_{P_{X_{ij}}} I(X_{ij}; Y_{ji}|Z_{ij}) = R_{ij}(P^*_{X_\mathcal{M}})$$

Therefore, we have

$$\begin{aligned}
\max_{P_{X_\mathcal{M}}} \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} R_{ij}(P_{X_\mathcal{M}}) &\geqslant \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} R_{ij}(P^*_{X_\mathcal{M}}) \\
&= \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} R^*_{ij} \\
&= \min_{\substack{i,j\in\mathcal{M} \\ \text{s.t. } e_{ij}\in\mathcal{E}_\mathcal{A}}} \max_{P_{X_\mathcal{M}}} R_{ij}(P_{X_\mathcal{M}}),
\end{aligned}$$

which completes the proof.   ■

Chapter 7

Conclusion and Future Work

In this thesis, we studied the problem of secret key agreement through the lens of information theory. The key agreement protocols that we considered, guarantee information theoretic security without making computational assumptions about the adversary. Instead, conceptual assumptions are on the amount and structure of the leakage that is available to the eavesdropping adversary; which in practice translate into physical-layer requirements that could be realized, for example, in wireless network environments.

The two main categories of information theoretic SKA models are: (i) source model, and (ii) channel model. Chapters 3 and 4 studied two different topics within the context of source model, and Chapters 5 and 6 explored SKA in channel model. The contributions of these chapters include:

- Proving finite-length upper and lower bounds for the maximum achievable key length of two-party one-way SKA.

- Deriving the WSK capacity of the Tree-PIN source model.

- Introducing the transceiver model and proving upper and lower bounds on its SK, PK, and WSK capacities.

- Deriving the WSK capacity of the Polytree-PIN channel model.

In the following, we outline some of the open research questions that arise from our investigation and will be interesting directions for future research. We list them in order of appearing in the thesis Chapters.

- In Chapter 3 we gave finite-length upper and lower bounds on $S^{\rightarrow}$ – the maximum achievable key length. However, these bounds are not tight. One immediate direction is the quest for a tight pair of finite-length upper and lower bounds that match up to at least the second order terms. Another related question is to find a tight converse for $S^{\rightarrow}$ under the assumption that $X - Y - Z$ holds. Such finite-length upper bound would then be directly comparable with the second order characterization of (interactive) S that was proved in [31]. Information spectrum methods and appropriate spectral entropies can provide powerful tools to tackle these questions.

- In Chapter 3 we proposed a two-party one-way SKA protocol (Protocol 5 $\mathbf{\Pi_{PH}}$) that has computational complexity $\mathcal{O}(n \log n)$ and we proved its finite-length behavior. It would certainly be an intriguing research avenue to find more efficient OW-SKA protocols, that can either perform better than $\mathbf{\Pi_{PH}}$ with respect to either computational complexity, and/or finite-key length.

- We utilized the information spectrum methods, in Chapter 3, to prove a "single-copy" ($n = 1$) upper bound on maximum key length of two-party OW-SKA. A similar direction is to investigate if we can prove single-copy upper bounds on maximum key length of the general multiterminal source model of [21].

- In Chapter 4 we argued that our proposed SKA protocol to achieve the WSK capacity of Tree-PIN is more efficient than the PK capacity achieving protocol of [21] from the perspective of using public communication bits. It remains open whether our protocol is optimum in that sense. Then, a related question also arises which is to find the WSK capacity of Tree-PIN under arbitrary upper limits on the asymptotic pubic communication rate. Recall that we proved WSK of Tree-PIN under the assumption

that terminals have free and *unlimited* access to public discussion. See [115], which studies similar questions for non-wiretapped PIN.

- In Chapter 4 we also discussed about the gap between noninteractive WSK capacity and the general WSK capacity in multiterminal source models. The gap remains open, and it would be interesting to see if the gap can be tightened or closed using either novel converse techniques or new noninteractive SKA protocols. A converse technique which has been often fruitful is the axiomatic method for proving upper bounds – see for example [26, 93]. That is a function of the source distribution is proved to be a legitimate upper bound on key capacity, if it satisfies a set of logical axioms (conditions). It seems that this method falls short of providing a powerful tool for proving a noninteractive converse. Maybe this approach could be improved and then utilized or perhaps new converse methods will be required.

- We studied the transceiver model in Chapter 5 and proved upper and lower bounds for SK, PK, and WSK capacities. A primary future objective is the pursuit of characterizing SK capacity for the multiterminal transceiver model. This goal perhaps requires new converse methods and/or novel SKA constructions. Similarly, finding tighter upper and lower bounds on PK, and WSK capacity is an important open question we leave for future work.

- Our channel model lower bounds in Chapter 5 and 6 are all based on the source emulation approach. Unlike the case of Polytree-PIN, source emulation is not always capacity achieving. Finding more efficient interactive and adaptive SKA protocols is also a valuable research topic. Moreover, an appealing question is to search for necessary and/or sufficient conditions a channel model has to satisfy under which the general or simple source emulation approaches are capacity achieving.

- At the end of Chapter 5 we proved the non-adaptive SK capacity using the converse bound of [23] and simplifying the expression by enforcing the non-adaptive SKA as-

sumptions. It remains open whether a similar tight converse can be proved for PK capacity using the same bound in [23]. This leads also to a call for improving the PK capacity upper bound of [23].

- Similar to the extension of Tree-PIN model that we studied in Chapter 4 Section 4.5.2, it would be interesting to study SKA in the extended model of Polytree-PIN where two Markov relations (with opposite directions) hold with respect to each edge e_{ij}. For such case, our converse techniques presented in Chapter 6 might be useful.

- In this thesis, we focused on the source and channel models of [21–23]. However, modified variants of these models have been studied for key agreement as well. For example, [128–130] consider the problem of secret key agreement for state-dependent channel models, and [122] proposes a new two-party model that combines the source and channel models of SKA. Therefore, extending the transceiver (or Polytree-PIN) model to the case of state-dependent channels or the case when correlated variables of a multiterminal source model are also available at the transceiver terminals are among the exciting problems we leave for future work.

- For both the multiterminal source and channel models, we considered the case where a single adversary is obtaining side information about the variables of the legitimate terminals. Studying SKA models with multiple (cooperating and/or non-cooperating) adversaries with distinct goals, is an interesting future research direction.

Bibliography

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, Oct. 1949. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6769090

[2] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, Nov. 1976. [Online]. Available: http://ieeexplore.ieee.org/document/1055638/

[3] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, Oct. 1997. [Online]. Available: http://epubs.siam.org/doi/10.1137/S0097539795293172http://arxiv.org/abs/quant-ph/9508027http://dx.doi.org/10.1137/S0097539795293172

[4] M. Steiner, G. Tsudik, and M. Waidner, "Key agreement in dynamic peer groups," *IEEE Transactions on Parallel and Distributed Systems*, vol. 11, no. 8, pp. 769–780, Oct. 2000. [Online]. Available: http://ieeexplore.ieee.org/document/877936/

[5] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, p. 120–126, Feb. 1978. [Online]. Available: https://doi.org/10.1145/359340.359342

[6] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-

H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, Q. Zhu, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, "Strong quantum computational advantage using a superconducting quantum processor," *Phys. Rev. Lett.*, vol. 127, p. 180501, Oct. 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.127.180501

[7] E. Gouzien and N. Sangouard, "Factoring 2048-bit rsa integers in 177 days with 13 436 qubits and a multimode memory," *Phys. Rev. Lett.*, vol. 127, p. 140503, Sep. 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.127.140503

[8] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, Oct. 2019. [Online]. Available: http://www.nature.com/articles/s41586-019-1666-5

[9] M. Mosca, "Cybersecurity in an era with quantum computers: Will we be ready?"

IEEE Security & Privacy, vol. 16, no. 5, pp. 38–41, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8490169/

[10] "Post-quantum cryptography standardization," https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization, accessed: 2021-11-10.

[11] "Post-quantum cryptography," https://www.dhs.gov/quantum, accessed: 2021-11-10.

[12] J. Bos, C. Costello, L. Ducas, I. Mironov, M. Naehrig, V. Nikolaenko, A. Raghunathan, and D. Stebila, "Frodo: Take off the ring! practical, quantum-secure key exchange from LWE," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1006–1018. [Online]. Available: https://doi.org/10.1145/2976749.2978425

[13] D. Stebila and M. Mosca, "Post-quantum key exchange for the Internet and the Open Quantum Safe project," in *Proc. 23rd Conference on Selected Areas in Cryptography (SAC) 2016*, ser. LNCS, R. Avanzi and H. Heys, Eds., vol. 10532. Springer, October 2017, pp. 1–24. [Online]. Available: https://openquantumsafe.org

[14] M. Sasaki, "Quantum key distribution and its applications," *IEEE Security & Privacy*, vol. 16, no. 5, pp. 42–48, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8490176/

[15] E. Crockett, C. Paquin, and D. Stebila, "Prototyping post-quantum and hybrid key exchange and authentication in tls and ssh," Cryptology ePrint Archive, Report 2019/858, 2019, https://ia.cr/2019/858.

[16] Y. Liang, H. V. Poor, and S. Shamai, "Information theoretic security," *Foundations and Trends® in Communications and Information Theory*, vol. 5, no. 4–5, pp. 355–

580, Jun. 2008. [Online]. Available: http://www.nowpublishers.com/article/Details/CIT-036

[17] Z. Rezki, M. Zorgui, B. Alomair, and M.-S. Alouini, "Secret key agreement: Fundamental limits and practical challenges," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 72–79, Jun. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7856875/

[18] Y. Wu, A. Khisti, C. Xiao, G. Caire, K.-K. Wong, and X. Gao, "A survey of physical layer security techniques for 5g wireless networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 679–695, Apr. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8335290/

[19] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, May 1993. [Online]. Available: http://ieeexplore.ieee.org/document/256484/

[20] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993. [Online]. Available: http://ieeexplore.ieee.org/document/243431/

[21] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1362897/

[22] I. Csiszár and P. Narayan, "Secrecy capacities for multiterminal channel models," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2437–2452, Jun. 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4529269/

[23] I. Csiszár and P. Narayan, "Secrecy generation for multiaccess channel models," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 17–31, Jan. 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6290395/

[24] T. S. Han, *Information-Spectrum Methods in Information Theory*, ser. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, vol. 50. [Online]. Available: http://link.springer.com/10.1007/978-3-662-12066-8

[25] S. Watanabe and M. Hayashi, "Non-asymptotic analysis of privacy amplification via renyi entropy and inf-spectral entropy," in *2013 IEEE International Symposium on Information Theory*. IEEE, Jul. 2013, pp. 2715–2719. [Online]. Available: http://ieeexplore.ieee.org/document/6620720/

[26] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *11th International Conference on the Theory and Application of Cryptology and Information Security - Advances in Cryptology - ASIACRYPT 2005*, B. Roy, Ed. Chennai, India: Springer Berlin Heidelberg, 2005, pp. 199–216. [Online]. Available: http://link.springer.com/10.1007/11593447_11

[27] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM Journal on Computing*, vol. 17, no. 2, pp. 210–229, Apr. 1988. [Online]. Available: http://epubs.siam.org/doi/10.1137/0217014

[28] J. L. Carter and M. N. Wegman, "Universal classes of hash functions (extended abstract)," in *Proceedings of the ninth annual ACM symposium on Theory of computing - STOC '77*. New York, New York, USA: ACM Press, 1977, pp. 106–112. [Online]. Available: http://portal.acm.org/citation.cfm?doid=800105.803400

[29] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5075875/

[30] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," in *Proceedings of the twenty-first annual ACM symposium on Theory of*

computing - STOC '89.   New York, New York, USA: ACM Press, 1989, pp. 12–24. [Online]. Available: http://portal.acm.org/citation.cfm?doid=73007.73009

[31] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3796–3810, Jul. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7469367/

[32] S. Sharifian, A. Poostindouz, and R. Safavi-Naini, "A capacity-achieving one-way key agreement with improved finite blocklength analysis," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*.   IEEE, Oct. 2020, pp. 407–411, Copyright© 2020 IEICE. [Online]. Available: https://ieeexplore.ieee.org/document/9366148

[33] A. Poostindouz and R. Safavi-Naini, "Second-order asymptotics for one-way secret key agreement," in *2021 IEEE International Symposium on Information Theory (ISIT)*.   IEEE, Jul. 2021, pp. 1254–1259. [Online]. Available: https://ieeexplore.ieee.org/document/9518202/

[34] A. Poostindouz and R. Safavi-Naini, "Wiretap secret key capacity of Tree-PIN," in *2019 IEEE International Symposium on Information Theory (ISIT)*.   IEEE, Jul. 2019, pp. 315–319. [Online]. Available: https://ieeexplore.ieee.org/document/8849553/

[35] A. Poostindouz and R. Safavi-Naini, "Secret key agreement in wiretapped Tree-PIN," *To be submitted to IEEE Transactions on Information Theory*, 2022. [Online]. Available: http://arxiv.org/abs/1903.06134

[36] A. Poostindouz and R. Safavi-Naini, "A channel model of transceivers for multiterminal secret key agreement," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*.   IEEE, Oct. 2020, pp. 412–416, Copyright© 2020 IEICE. [Online]. Available: https://ieeexplore.ieee.org/document/9366098

219

[37] A. Poostindouz and R. Safavi-Naini, "Secret key agreement in multiterminal channel model of transceivers," *To be submitted to Entropy*, 2022. [Online]. Available: https://arxiv.org/abs/2008.02977

[38] A. Poostindouz and R. Safavi-Naini, "Secret key capacity of wiretapped Polytree-PIN," in *The 2021 IEEE Information Theory Workshop (ITW2021)*. IEEE, Oct. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9611419

[39] A. Papoulis, *Probability, random variables, and stochastic processes.* Boston: McGraw-Hill, 2002.

[40] R. W. Yeung, *Information Theory and Network Coding.* Boston, MA: Springer US, 2008. [Online]. Available: http://link.springer.com/10.1007/978-0-387-79234-7http://www.amazon.com/dp/1441946306

[41] M. Mitzenmacher and E. Upfal, *Probability and Computing.* Cambridge: Cambridge University Press, Jan. 2005. [Online]. Available: https://doi.org/10.1017/CBO9780511813603

[42] R. G. Gallager, *Stochastic processes: theory for applications.* Cambridge University Press, 2013.

[43] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed., ser. Wiley Series in Probability and Statistics. Wiley, 1982, vol. 2.

[44] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6773024

[45] A. El Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge: Cambridge University Press, 2011. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9781139030687

[46] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973. [Online]. Available: http://ieeexplore.ieee.org/document/1055037/

[47] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Hoboken, NJ, USA: John Wiley {\&} Sons, Inc., Sep. 2005. [Online]. Available: http://doi.wiley.com/10.1002/047174882X

[48] T. S. Han and S. Verdu, "Approximation theory of output statistics," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 752–772, May 1993. [Online]. Available: http://ieeexplore.ieee.org/document/256486/

[49] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009. [Online]. Available: http://ieeexplore.ieee.org/document/5290292/

[50] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5452208/

[51] M. Hayashi, "Semi-finite length analysis for information theoretic tasks," pp. 1–29, Nov. 2018. [Online]. Available: http://arxiv.org/abs/1811.00262

[52] V. Strassen, "Asymptotische abschätzugen in shannon's informationstheorie," *Transactions of the Third Prague Conference on Information Theory etc,*, pp. 689–723, 1962.

[53] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4619–4637, Oct. 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4626060/

[54] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *2006 IEEE International Symposium on Information Theory.* IEEE,

Jul. 2006, pp. 2593–2597. [Online]. Available: http://ieeexplore.ieee.org/document/4036441/

[55] C. Ye and A. Reznik, "Group secret key generation algorithms," in *2007 IEEE International Symposium on Information Theory*, vol. 1, no. 1.   IEEE, Jun. 2007, pp. 2596–2600. [Online]. Available: http://ieeexplore.ieee.org/document/4557610/

[56] D. Jost, U. Maurer, and J. L. Ribeiro, "Information-theoretic secret-key agreement: The asymptotically tight relation between the secret-key rate and the channel quality ratio," in *Theory of Cryptography. TCC 2018. Lecture Notes in Computer Science*, A. Beimel and S. Dziembowski, Eds.   Springer, Cham, 2018, pp. 345–369. [Online]. Available: http://link.springer.com/10.1007/978-3-030-03807-6_13

[57] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6772207

[58] R. Renner, "Security of quantum key distribution," Dec. 2005. [Online]. Available: https://doi.org/10.3929/ethz-a-005115027

[59] H. Tyagi and S. Watanabe, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *Advances in Cryptology – EUROCRYPT 2014*, ser. Lecture Notes in Computer Science, P. Q. Nguyen and E. Oswald, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, vol. 8441, pp. 369–386. [Online]. Available: http://link.springer.com/10.1007/978-3-642-55220-5_21

[60] S. Nitinawarat, C. Ye, A. Barg, P. Narayan, and A. Reznik, "Secret key generation for a pairwise independent network model," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6482–6489, Dec. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5625626/

[61] H. Tyagi, "Common randomness principles of secrecy," Ph.D. Dissertation, University of Maryland (College Park, Md.), 2013. [Online]. Available: http://drum.lib.umd.edu/handle/1903/14670

[62] C. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, 1995. [Online]. Available: http://ieeexplore.ieee.org/document/476316/

[63] H. Tyagi and A. Vardy, "Universal hashing for information-theoretic security," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1781–1795, Oct. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7270413/

[64] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, Jan. 2008. [Online]. Available: http://epubs.siam.org/doi/10.1137/060651380

[65] M. Hayashi, "Tight exponential analysis of universally composable privacy amplification and its applications," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7728–7746, Nov. 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6613554/

[66] M. Hayashi and T. Tsurumaru, "More efficient privacy amplification with less random seeds via dual universal hash function," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2213–2232, Apr. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7399404/

[67] R. Shaltiel, "Recent developments in explicit constructions of extractors," in *Current Trends in Theoretical Computer Science*. World Scientific, Apr. 2004, pp. 189–228. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/9789812562494_0013

[68] D. Elkouss, J. Martinez-mateo, and V. Martin, "Information reconciliation for quantum key distribution," *Quantum Information & Computation*, vol. 11, no. 3, p. 226–238, Mar. 2011.

[69] D. Elkouss, "Information reconciliation methods in secret-key distribution," Ph.D. dissertation, Universidad Politecnica de Madrid, 2011.

[70] M. Sudan, H. Tyagi, and S. Watanabe, "Communication for generating correlation: A unifying survey," *IEEE Transactions on Information Theory*, vol. 66, no. 1, pp. 5–37, Jan. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8863950/

[71] T. A. Courtade and T. R. Halford, "Coded cooperative data exchange for a secret key," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3785–3795, Jul. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7428916/

[72] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, "On the optimality of secret key agreement via omniscience," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2371–2389, Apr. 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8281549/

[73] H. Tyagi and S. Watanabe, "Secret key capacity for multipleaccess channel with public feedback," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, Oct. 2013, pp. 1–7. [Online]. Available: http://ieeexplore.ieee.org/document/6736497/

[74] P. Narayan and H. Tyagi, "Multiterminal secrecy by public discussion," *Foundations and Trends® in Communications and Information Theory*, vol. 13, no. 2-3, pp. 129–275, 2016. [Online]. Available: http://www.nowpublishers.com/article/Details/CIT-072

[75] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals—part ii: Channel model," *IEEE Transactions on Information Theory*,

vol. 56, no. 8, pp. 3997–4010, Aug. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5508612/

[76] T. Holenstein and R. Renner, "One-way secret-key agreement and applications to circuit polarization and immunization of public-key encryption," in *Crypto 2005*, V. Shoup, Ed., vol. LNCS 3621.   Springer, Berlin, Heidelberg, 2005, pp. 478–493. [Online]. Available: http://link.springer.com/10.1007/11535218_29

[77] J. Wullschleger, "Oblivious-transfer amplification," in *EUROCRYPT 2007*, M. Naor, Ed., vol. 4515 LNCS.   Springer, Berlin, Heidelberg, 2007, pp. 555–572. [Online]. Available: http://link.springer.com/10.1007/978-3-540-72540-4_32

[78] T. Holenstein, "Strengthening key agreement using hard-core sets," Ph.D. dissertation, ETH ZURICH, 2006. [Online]. Available: https://doi.org/10.3929/ethz-a-005205852

[79] J. M. Renes, R. Renner, and D. Sutter, "Efficient one-way secret-key agreement and private channel coding via polarization," in *ASIACRYPT 2013*, K. Sako and P. Sarkar, Eds., vol. LNCS 8269.   Springer, Berlin, Heidelberg, 2013, pp. 194–213. [Online]. Available: http://link.springer.com/10.1007/978-3-642-42033-7_11

[80] R. A. Chou, M. R. Bloch, and E. Abbe, "Polar coding for secret-key generation," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 6213–6237, Nov. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7217814/

[81] J. Muramatsu and S. Miyake, "Construction of codes for the wiretap channel and the secret key agreement from correlated source outputs based on the hash property," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 671–692, 2012.

[82] M. Tomamichel, J. Martinez-Mateo, C. Pacher, and D. Elkouss, "Fundamental finite key limits for information reconciliation in quantum key distribution," in *2014 IEEE International Symposium on Information Theory*.   IEEE, Jun. 2014,

pp. 1469–1473. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6875077

[83] T. Holenstein and R. Renner, "On the randomness of independent experiments," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1865–1871, Apr. 2011. [Online]. Available: http://ieeexplore.ieee.org/document/5730579/

[84] A. C. Berry, "The accuracy of the gaussian approximation to the sum of independent variates," *Transactions of the American Mathematical Society*, vol. 49, no. 1, p. 122, Jan. 1941. [Online]. Available: https://www.jstor.org/stable/1990053?origin=crossref

[85] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, Jul. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8665906/

[86] M. Hayashi, "Semi-finite length analysis for secure random number generation," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Jul. 2019, pp. 952–956. [Online]. Available: https://ieeexplore.ieee.org/document/8849241/

[87] M. Bellare and S. Tessaro, "Polynomial-time, semantically-secure encryption achieving the secrecy capacity," *Arxiv preprint*, Jan. 2012. [Online]. Available: http://arxiv.org/abs/1201.3160

[88] M. H. Yassaee, M. R. Aref, and A. Gohari, "Non-asymptotic output statistics of random binning and its applications," *IEEE International Symposium on Information Theory - Proceedings*, no. 88114, pp. 1849–1853, 2013.

[89] N. Tavangaran, H. Boche, and R. Schaefer, "Secret-key generation using compound sources and one-way public communication," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 1–1, 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7572155/

[90] I. Csiszár and J. Körner, *Information Theory.* Cambridge: Cambridge University Press, 2011. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9780511921889

[91] U. M. Maurer and S. Wolf, "Unconditionally secure key agreement and the intrinsic conditional information," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 499–514, Mar. 1999. [Online]. Available: http://ieeexplore.ieee.org/document/748999/

[92] R. Renner and S. Wolf, "New bounds in secret-key agreement: The gap between formation and secrecy extraction," in *Advances in Cryptology—EUROCRYPT 2003*, 2003, pp. 562–577. [Online]. Available: http://link.springer.com/10.1007/3-540-39200-9_35

[93] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals—part i," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5508611/

[94] A. A. Gohari, O. Gunlu, and G. Kramer, "On achieving a positive rate in the source model key agreement problem," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Jun. 2018, pp. 2659–2663. [Online]. Available: https://ieeexplore.ieee.org/document/8437749/

[95] M. Pinsker, *Information and information stability of random variables and processes*, 1st ed., ser. Holden-Day series in time series analysis. San Francisco: Holden-Day, 1964.

[96] S. H. Hassani, K. Alishahi, and R. L. Urbanke, "Finite-length scaling for polar codes," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 5875–5898, Oct. 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6866198/

[97] M. Mondelli, S. H. Hassani, and R. L. Urbanke, "Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors," *IEEE Transactions*

on Information Theory, vol. 62, no. 12, pp. 6698–6712, Dec. 2016. [Online]. Available: https://ieeexplore.ieee.org/document/7589109/

[98] A. Fazeli, H. Hassani, M. Mondelli, and A. Vardy, "Binary linear codes with optimal scaling: Polar codes with large kernels," *IEEE Transactions on Information Theory*, pp. 1–1, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9261460/

[99] N. Hussami, S. B. Korada, and R. Urbanke, "Performance of polar codes for channel and source coding," in *2009 IEEE International Symposium on Information Theory*. IEEE, Jun. 2009, pp. 1488–1492. [Online]. Available: http://ieeexplore.ieee.org/document/5205860/

[100] S. B. Korada and R. Urbanke, "Polar codes for slepian-wolf, wyner-ziv, and gelfand-pinsker," in *IEEE Information Theory Workshop 2010 (ITW 2010)*. IEEE, Jan. 2010, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/document/5503220/

[101] Vu Thi Thuy Trang, J. W. Kang, M. Jang, Jong-hwan Kim, and S.-H. Kimy, "The performance of polar codes in distributed source coding," in *2012 Fourth International Conference on Communications and Electronics (ICCE)*. IEEE, Aug. 2012, pp. 196–199. [Online]. Available: http://ieeexplore.ieee.org/document/6315897/

[102] C. Yaacoub and M. Sarkis, "Distributed compression of correlated sources using systematic polar codes," in *2016 9th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*. IEEE, Sep. 2016, pp. 96–100. [Online]. Available: http://ieeexplore.ieee.org/document/7593084/

[103] M. Tomamichel and A. Leverrier, "A largely self-contained and complete security proof for quantum key distribution," *Quantum*, vol. 1, p. 14, Jul. 2017. [Online]. Available: http://quantum-journal.org/papers/q-2017-07-14-14/

[104] P. K. Vippathalla, C. Chan, N. Kashyap, and Q. Zhou, "Secret key agreement and secure omniscience of Tree-PIN source with linear wiretapper," in *2021 IEEE Inter-*

national *Symposium on Information Theory (ISIT)*. IEEE, Jul. 2021, pp. 1624–1629. [Online]. Available: https://ieeexplore.ieee.org/document/9518075/

[105] C. Chan, "Linear perfect secret key agreement," in *2011 IEEE Information Theory Workshop*. IEEE, Oct. 2011, pp. 723–726. [Online]. Available: http://ieeexplore.ieee.org/document/6089530/

[106] S. Nitinawarat and P. Narayan, "Perfect omniscience, perfect secrecy, and steiner tree packing," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6490–6500, Dec. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5625644/

[107] N. Kashyap, M. Mukherjee, and Y. Sankarasubramaniam, "On the secret key capacity of the harary graph pin model," in *2013 National Conference on Communications, NCC 2013*. IEEE, Feb. 2013, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/document/6487950/

[108] P. Xu, Z. Ding, and X. Dai, "The private key capacity of a cooperative pairwise-independent network," in *2015 IEEE International Symposium on Information Theory (ISIT)*, vol. 2015-June. IEEE, Jun. 2015, pp. 286–290. [Online]. Available: http://ieeexplore.ieee.org/document/7282462/

[109] C. Chan and L. Zheng, "Mutual dependence for secret key agreement," in *2010 44th Annual Conference on Information Sciences and Systems (CISS)*, vol. 2. IEEE, Mar. 2010, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/5464805/

[110] Q. Zhou and C. Chan, "Secrecy capacity under limited discussion rate for minimally connected hypergraphical sources," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Jun. 2018, pp. 2664–2668. [Online]. Available: https://ieeexplore.ieee.org/document/8437565/

[111] H. Tyagi and S. Watanabe, "Universal multiparty data exchange and secret key agree-

ment," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4057–4074, Jul. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7902138/

[112] H. Tyagi, "Common information and secret key capacity," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5627–5640, Sep. 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6517479/

[113] M. Mukherjee and N. Kashyap, "The communication complexity of achieving sk capacity in a class of pin models," in *2015 IEEE International Symposium on Information Theory (ISIT)*, vol. 2015-June.   IEEE, Jun. 2015, pp. 296–300. [Online]. Available: http://ieeexplore.ieee.org/document/7282464/

[114] M. Mukherjee, C. Chan, N. Kashyap, and Q. Zhou, "Bounds on the communication rate needed to achieve sk capacity in the hypergraphical source model," in *2016 IEEE International Symposium on Information Theory (ISIT)*, vol. 2016-Augus.   IEEE, Jul. 2016, pp. 2504–2508. [Online]. Available: http://ieeexplore.ieee.org/document/7541750/

[115] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, "Secret key agreement under discussion rate constraints," in *2017 IEEE International Symposium on Information Theory (ISIT)*.   IEEE, Jun. 2017, pp. 1519–1523. [Online]. Available: http://ieeexplore.ieee.org/document/8006783/

[116] H. Tyagi and S. Watanabe, "Converses for secret key agreement and secure computing," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4809–4827, Sep. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7161366/

[117] R. Diestel, *Graph Theory*, ser. Graduate Texts in Mathematics.   Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, vol. 173. [Online]. Available: http://link.springer.com/10.1007/978-3-662-53622-3

[118] R. Ahlswede and J. Körner, "Appendix: On common information and related charac-teristics of correlated information sources," in *General Theory of Information Transfer and Combinatorics*, R. Ahlswede, L. Bäumer, N. Cai, H. Aydinian, V. Blinovsky, C. Deppe, and H. Mashurian, Eds. Springer Berlin Heidelberg, 2006, pp. 664–677. [Online]. Available: http://link.springer.com/10.1007/11889342_41

[119] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[120] H. Tyagi and S. Watanabe, "Secret key capacity for multipleaccess channel with pub-lic feedback," in *2013 51st Annual Allerton Conference on Communication, Con-trol, and Computing (Allerton)*. IEEE, Oct. 2013, pp. 1–7. [Online]. Available: http://ieeexplore.ieee.org/document/6736497/

[121] C. Chan, "Compressed secret key agreement:maximizing multivariate mutual infor-mation per bit," *Entropy*, vol. 19, no. 10, p. 545, Oct. 2017. [Online]. Available: http://www.mdpi.com/1099-4300/19/10/545

[122] G. Bassi, P. Piantanida, and S. Shamai (Shitz), "The secret key capacity of a class of noisy channels with correlated sources," *Entropy*, vol. 21, no. 8, p. 732, Jul. 2019. [Online]. Available: https://www.mdpi.com/1099-4300/21/8/732

[123] A. Goldsmith, *Wireless Communications*. Cambridge: Cambridge University Press, 2005. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9780511841224

[124] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 4–21, Jan. 2003. [Online]. Available: http://ieeexplore.ieee.org/document/1159758/

[125] A. Gohari and G. Kramer, "An upper bound for wiretap multi-way channels," pp. 1–9, Sep. 2020. [Online]. Available: http://arxiv.org/abs/2009.14814

[126] A. Mukherjee, "Physical-layer security in the internet of things: Sensing and communication confidentiality under resource constraints," *Proceedings of the IEEE*, vol. 103, no. 10, pp. 1747–1761, Oct. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7258313/

[127] D. P. Bertsekas, *Convex optimization theory*. Nashua, NH, U.S.A.: Athena Scientific, 2009.

[128] Y. Chen and A. J. Han Vinck, "Wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, 2008. [Online]. Available: https://ieeexplore.ieee.org/document/4418466

[129] A. Zibaeenejad, "Key generation over wiretap models with non-causal side information," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 7, pp. 1456–1471, 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7047879

[130] Z. Goldfeld, P. Cuff, and H. H. Permuter, "Wiretap channels with random states non-causally available at the encoder," 2019. [Online]. Available: https://arxiv.org/abs/1608.00743

Appendix A

Copyright Permissions

Permission for Paper [34]

Permission for Papers [32, 36]

IEICE holds the copyright of papers [32] and [36]. Please refer to the copyright information included below which is available online via https://www.ieice.org/eng/copyright/copyright_faq.html – please see highlighted text.

Copyright FAQ

Copyright FAQ

This is a list of frequently- asked questions (FAQ) relating to the use of authored works in IEICE publications, including the application procedure. To find out more about the IEICE's application standards, you should also see the document "Explanation of IEICE Provisions on Copyright".

INDEX

1. Use by the author himself or herself (or the organization to which the author belongs)

1-1. When using article/paper, etc. published in the IEICE

1-2. When using article paper, etc. that have been submitted to the IEICE and have not been published

2. Use by third parties (Other than the author / organization to which the author belongs)

3. Use of article/paper, etc. appearing in other publications

4. Miscellaneous questions and answers

1. Use by the author himself or herself (or the organization to which the author belongs)

>Top

1-1. When using article/paper, etc. published in the IEICE

Q 1: Can an author use his or her own article/paper, etc. whose copyrights have been transferred to the IEICE, without request permission to the IEICE? What kind of case is that?

A 1:

Permission requests for use are unnecessary, in the case of using it for "non-commercial purposes" and "the use does not unfairly infringe on the IEICE's interests." However, the following conditions must be met.

[Conditions for no permission request (must meet all)]
Copyright notices (e.g., copyright©2020 IEICE)
Indication of source (e.g., author name, title, magazine name, volume, issue,

page, year of publication, etc.)
Author's consent (When using the organization to which the author belongs / when there are some co-authors)
Publication of publisher's version PDF (When using the whole article/paper)

【Permitted medium】
Author's own personal server, server of an organization to which the author belongs (*2), preprint server(*3), electronic media such as DVDs, paper medium author's bachelor's/master's/doctoral dissertation, bulletin of the university/college/school to which the author belongs and any publications published by the organization to which the author belongs.

【Permitted period】
In the case of using the whole article/paper, in principle, it can be used after publication (*1). Please use the publisher's version (PDF).

* 1
In the IEICE Transactions that has "advanced publication", "advanced publication version PDF" is available from advanced publication to until publication of publisher's version.
(Advance publication version PDF means a PDF to which minimum level of alteration has been performed by IEICE, just sufficient to indicate that the version is an early publication version.)
However, advance publication version must be replaced when publisher's version is published.

* 2
Author's own personal server: A server to or from which the author can upload or delete material without any permission from others (e.g., a blog or the server of a university department).
server of an organization that author belongs to: A public website wholly managed and administered by the institution as an organization (e.g., an institutional repository, a Homepage of organization, company or university/college/school intranets).

* 3
See Q9 for using the preprint server.

Q 2: Dose an author need the IEICE's permission to write his or her dissertation based on or copied in a substantial amount (or in whole) from a paper published in IEICE?
Is it possible for the university/college to use the thesis freely available in the university's repository? How should I quote it?

A 2: For dissertations, application is not required, under the conditions of "Indication of source and copyright".
Please refer to the following for the citation method.

＜Using the whole paper or copying a substantial amount from a paper＞
Please refer to the IEICE publications and indicate the source and copyright in a footnote.
Example of annotation in a footnote：
This dissertation is based on "Title" [1], by the same author, which appeared in the Proceedings of ****, Copyright(C)2020 IEICE.
The material in this paper was presented in part at the Proceedings of **** [1], and all the figures of this paper are reused from [1] under the permission of the IEICE.

＜Using the figures＞
Please refer to the IEICE publications and indicate the copyright in a captions of the figures.
Example of annotation:

Permission for Paper [33]

Permission for Paper [38]